Publication design with incentives in mind

Ravi Jagadeesan
Stanford

Davide Viviano
Harvard

"Economists are quick to assume opportunistic behavior in almost every walk of life other than our own. Our empirical methods are based on assumptions of human behavior that would not pass muster in any of our models." (Glaeser, 2006)

Selective publication

- suppose an *editor* is deciding which findings to publish
  - published studies may inform the public about the state of the world
  - the public (*audience*) will take a decision after observing published studies
  - the editor wants to minimize the audience's loss

Selective publication

- suppose an *editor* is deciding which findings to publish
  - published studies may inform the public about the state of the world
  - the public (*audience*) will take a decision after observing published studies
  - the editor wants to minimize the audience's loss

- if publication is costly (e.g., cognitive costs for the audience),
  optimal policy is to publish a result if and only if it is sufficiently "surprising"
  - Frankel and Kasy (2022)

Selective publication

- suppose an *editor* is deciding which findings to publish
  - published studies may inform the public about the state of the world
  - the public (*audience*) will take a decision after observing published studies
  - the editor wants to minimize the audience's loss

- if publication is costly (e.g., cognitive costs for the audience), optimal policy is to publish a result if and only if it is sufficiently "surprising"
  - Frankel and Kasy (2022)

- but if *researchers* are interested in publishing, selective publication affects their incentives about what studies to conduct and how to implement them

Selective publication

- suppose an *editor* is deciding which findings to publish
  - published studies may inform the public about the state of the world
  - the public (*audience*) will take a decision after observing published studies
  - the editor wants to minimize the audience's loss

- if publication is costly (e.g., cognitive costs for the audience), optimal policy is to publish a result if and only if it is sufficiently "surprising"
  - Frankel and Kasy (2022)

- but if *researchers* are interested in publishing, selective publication affects their incentives about what studies to conduct and how to implement them
  - e.g., may not run a costly large-scale experiment w/low chance of publishing
  - e.g., may manipulate results to increase chance of finding significant results

This paper

taking researcher's incentives to publish into account, we ask:

1. which research designs should be incentivized more when?
2. what form (if any) of selective publication is optimal?

This paper

taking researcher's incentives to publish into account, we ask:

1. which research designs should be incentivized more when?
2. what form (if any) of selective publication is optimal?

we formulate a model of optimal publication decisions that takes into account

- researcher's incentives about what studies to conduct (verifiable design)
- researcher's incentives to manipulate the findings (non verifiable design)

$\Rightarrow$ mechanism design problem with limited transfers

This paper

taking researcher's incentives to publish into account, we ask:

1. which research designs should be incentivized more when?
2. what form (if any) of selective publication is optimal?

we formulate a model of optimal publication decisions that takes into account

- researcher's incentives about what studies to conduct (verifiable design)
- researcher's incentives to manipulate the findings (non verifiable design)

$\Rightarrow$ mechanism design problem with limited transfers

some takeaways:

1. optimal publication is biased towards studies that are cheaper for researchers to do
2. less surprising results, and manipulated results, are sometimes published
3. even if planner can enforce non manipulable designs, this can be sub-optimal

Related literature

- economic analysis of statistics [Chassang et al. (2012); Tetenov (2016); Spiess (Forthcoming, 2025); Henry and Ottaviani (2019); Di Tillio et al. (2017); Viviano et al. (2025); Kasy and Spiess (2023)]
  - we study choosing between different study designs (and manipulation)
- modeling scientific approval and communication [Frankel and Kasy (2022); Andrews and Shapiro (2021); Glaeser (2006); Manski (2015)]
  - we provide model that incorporates researchers' incentives
- treatment effect literature with selection bias/external validity [e.g. Meager (2019); Allcott (2015); Beets et al. (2020); Rosenzweig and Udry (2016)]
  - we study how these issues interact with researcher's incentives
- work on decision theory and hypothesis testing [e.g., Wald (1950); Storey (2003); Efron (2008); Manski and Tetenov (2016); Manski (2004); McCloskey and Michaillat (Forthcoming, 2024)]
  - we provide an economic model with incentives for publication rules

Model setup

Model overview

- three agents: an editor, an audience, and a researcher
- state of the world $\theta \sim \mathcal{N}(\mu, \eta^2)$; without loss $\mu = 0$

Model overview

- three agents: an editor, an audience, and a researcher
- state of the world $\theta \sim \mathcal{N}(\mu, \eta^2)$; without loss $\mu = 0$

1. editor pre-commit to a publication rule $p(\cdot)$
2. researcher chooses study design Δ (associated with bias and variance)
    - to maximize chance of publication, net of cost C_Δ of executing Δ

Model overview

- three agents: an editor, an audience, and a researcher
- state of the world $\theta \sim \mathcal{N}(\mu, \eta^2)$; without loss $\mu = 0$

1. editor pre-commit to a publication rule $p(\cdot)$
2. researcher chooses study design Δ (associated with bias and variance)
    - to maximize chance of publication, net of cost C_Δ of executing Δ
3. researcher reports results $X(\Delta)$

Model overview

- three agents: an editor, an audience, and a researcher
- state of the world $\theta \sim \mathcal{N}(\mu, \eta^2)$; without loss $\mu = 0$

1. editor pre-commit to a publication rule $p(\cdot)$
2. researcher chooses study design Δ (associated with bias and variance)
    - to maximize chance of publication, net of cost C_Δ of executing Δ
3. researcher reports results $X(\Delta)$
4. if published, audience action $a^*(X)$ to minimize expected loss $[(a - \theta)^2 | X]$

Model overview

- three agents: an editor, an audience, and a researcher
- state of the world $\theta \sim \mathcal{N}(\mu, \eta^2)$; without loss $\mu = 0$

1. editor pre-commit to a publication rule $p(\cdot)$
2. researcher chooses study design Δ (associated with bias and variance)
   - to maximize chance of publication, net of cost C_Δ of executing Δ
3. researcher reports results $X(\Delta)$
4. if published, audience action $a^*(X)$ to minimize expected loss $[(a-\theta)^2|X]$

- editor minimizes audience's loss net of cost c_A per publication

Model: two cases

1. verifiable: $p(\cdot)$ can depend on X and Δ, researcher does not know θ
   - choosing betw/ experiments with different precisions: $X(\Delta) \sim \mathcal{N}(\theta, S^2_\Delta)$
   - cost varies by design
2. non verifiable: researcher chooses Δ as function of data and $p(\cdot)$ only depends on X
   - researcher can introduce bias in the study
   - consider (reputational) cost for data manipulation

Verifiable design

Model: verifiable design

Δ	$X \sim \mathcal{N}(\theta, S_\Delta^2)$	$p(X, \Delta)$	$a^*(X)$
Design	Execution	Evaluation	Action
(researcher)		(editor)	(audience)

Model: verifiable design

Δ	$X \sim \mathcal{N}(\theta, S_\Delta^2)$	$p(X, \Delta)$	$a^*(X)$
Design	Execution	Evaluation	Action
(researcher)		(editor)	(audience)

- researcher: $\max_\Delta b\mathbb{E}_X[p(X, \Delta)] - C_\Delta$ [without loss $b = 1$]

Model: verifiable design

Δ	$X \sim \mathcal{N}(\theta, S_\Delta^2)$	$p(X, \Delta)$	$a^*(X)$
Design	Execution	Evaluation	Action
(researcher)		(editor)	(audience)

- researcher: $\max_\Delta b\mathbb{E}_X[p(X,\Delta)] - C_\Delta$ [without loss $b = 1$]
- editor's objective:

$$\mathcal{L}_\Delta = \mathbb{E}\Big[\underbrace{p(X,\Delta)}_{\text{publication}}(a^*(X) - \theta)^2\Big]$$

Model: verifiable design

Δ	$X \sim \mathcal{N}(\theta, S_\Delta^2)$	$p(X, \Delta)$	$a^*(X)$
Design	Execution	Evaluation	Action
(researcher)		(editor)	(audience)

- researcher: $\max_\Delta b\mathbb{E}_X[p(X,\Delta)] - C_\Delta$ [without loss $b = 1$]
- editor's objective:

$$\mathcal{L}_\Delta = \mathbb{E}\Big[\underbrace{p(X,\Delta)(a^*(X) - \theta)^2}_{\text{publication}} + \underbrace{(1 - p(X,\Delta))(\theta - \mu)^2}_{\text{status quo}}$$

Model: verifiable design

Δ	$X \sim \mathcal{N}(\theta, S_\Delta^2)$	$p(X, \Delta)$	$a^*(X)$
Design	Execution	Evaluation	Action
(researcher)		(editor)	(audience)

- researcher: $\max_\Delta b\mathbb{E}_X[p(X,\Delta)] - C_\Delta$ [without loss $b = 1$]
- editor's objective:

$$\mathcal{L}_\Delta = \mathbb{E}\Big[\underbrace{p(X,\Delta)(a^*(X) - \theta)^2}_{\text{publication}} + \underbrace{(1 - p(X,\Delta))(\theta - \mu)^2}_{\text{status quo}} + \underbrace{p(X,\Delta)c_A}_{\text{cost}}\Big]$$

Which cheap studies to publish? ($C_O = 0$)

if the editor is constrained to implement $\Delta = O$ with $C_O = 0$,

then optimal publication decision rules satisfy (Frankel and Kasy, 2022)

$$p(X, O) = \begin{cases} 1 & \text{if } |X| > X_O^* \\ 0 & \text{if } |X| < X_O^* \end{cases},$$

where

$$X_O^* = \frac{S_O^2 + \eta^2}{\eta^2} \sqrt{c_A}$$

Which cheap studies to publish? ($C_O = 0$)

if the editor is constrained to implement $\Delta = O$ with $C_O = 0$,

then optimal publication decision rules satisfy (Frankel and Kasy, 2022)

$$p(X, O) = \begin{cases} 1 & \text{if } |X| > X_O^* \\ 0 & \text{if } |X| < X_O^* \end{cases},$$

where

$$X_O^* = \frac{S_O^2 + \eta^2}{\eta^2} \sqrt{c_A}$$

- intuition: publish results that move a enough to be worth paying c_A

Which expensive studies to publish? $C_E > 0$

if the editor is constrained to implement $\Delta = E$ with $C_E > 0$,

then optimal publication decision rules satisfy

$$p(X, E) = \begin{cases} 1 & \text{if } |X| > X_E^* \\ 0 & \text{if } |X| < X_E^* \end{cases},$$

where

$$X_E^* = \min\Big\{ \frac{S_E^2 + \eta^2}{\eta^2} \sqrt{c_A},$$

Which expensive studies to publish? $C_E > 0$

if the editor is constrained to implement $\Delta = E$ with $C_E > 0$,

then optimal publication decision rules satisfy

$$p(X, E) = \begin{cases} 1 & \text{if } |X| > X_E^* \\ 0 & \text{if } |X| < X_E^* \end{cases},$$

where

$$X_E^* = \min\left\{ \frac{S_E^2 + \eta^2}{\eta^2} \sqrt{c_A}, \quad \Phi^{-1}(1 - C_E/2)\sqrt{S_E^2 + \eta^2} \right\}$$

- intuition: need to make $\mathbb{E}[p(X, E)]$ large enough to implement E

Which expensive studies to publish? $C_E > 0$

if the editor is constrained to implement $\Delta = E$ with $C_E > 0$,

then optimal publication decision rules satisfy

$$p(X, E) = \begin{cases} 1 & \text{if } |X| > X_E^* \\ 0 & \text{if } |X| < X_E^* \end{cases},$$

where

$$X_E^* = \min\Big\{ \frac{S_E^2 + \eta^2}{\eta^2} \sqrt{c_A}, \quad \Phi^{-1}(1 - C_E/2)\sqrt{S_E^2 + \eta^2} \Big\}$$

- intuition: need to make $\mathbb{E}[p(X, E)]$ large enough to implement E
- relevant if the researcher's IR constraint binds for $\Delta = E$

Implications

- suppose two designs are available:
  - cheaper but less accurate study O, with $C_O > 0$
  - more expensive but more accurate study E, with $C_E > C_O$
  - Define $\mathrm{PostVar}(\Delta) = \mathbb{V}(\theta | X(\Delta))$ the posterior variance
- should the planner publish results from O? or from E alone?

Implications

- suppose two designs are available:
  - cheaper but less accurate study O, with $C_O > 0$
  - more expensive but more accurate study E, with $C_E > C_O$
  - Define $\mathrm{PostVar}(\Delta) = \mathbb{V}(\theta | X(\Delta))$ the posterior variance
- should the planner publish results from O? or from E alone?

Prop If both C_E, C_O are sufficiently small (but non-zero), and $S_E < S_O$ the editor prefers E

Implications

- suppose two designs are available:
  - cheaper but less accurate study O, with $C_O > 0$
  - more expensive but more accurate study E, with $C_E > C_O$
  - Define $\mathrm{PostVar}(\Delta) = \mathbb{V}(\theta | X(\Delta))$ the posterior variance

- should the planner publish results from O? or from E alone?

Prop If both C_E, C_O are sufficiently small (but non-zero), and $S_E < S_O$ the editor prefers E

Prop for O entailing non-trivial costs (IR is binding for O) planner prefers E over O iff

$$\mathrm{PostVar}(O) - \mathrm{PostVar}(E) \geq (C_E - C_O)c_A + O(\epsilon)$$

with $\epsilon = (1 - C_O)^3$ (exact expressions in the paper)

Implications

- suppose two designs are available:
  - cheaper but less accurate study O, with $C_O > 0$
  - more expensive but more accurate study E, with $C_E > C_O$
  - Define $\text{PostVar}(\Delta) = \mathbb{V}(\theta | X(\Delta))$ the posterior variance

- should the planner publish results from O? or from E alone?

Prop If both C_E, C_O are sufficiently small (but non-zero), and $S_E < S_O$ the editor prefers E

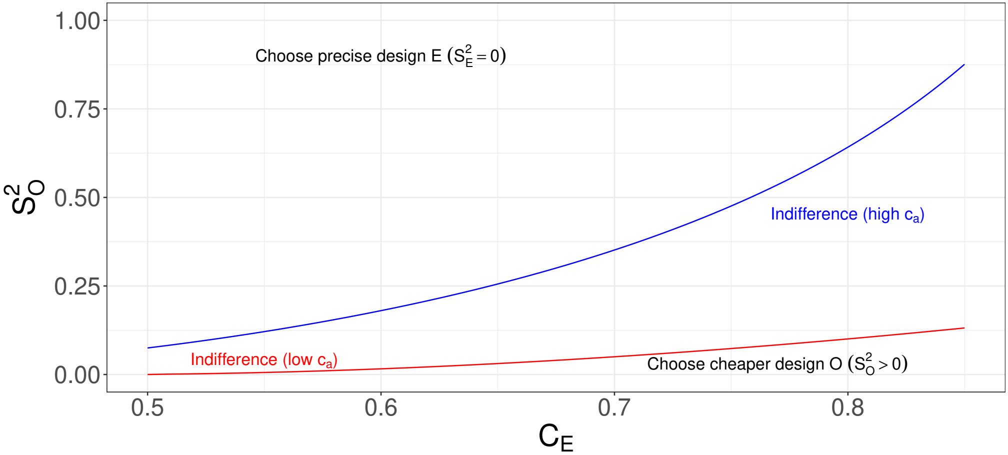
Prop for O entailing non-trivial costs (IR is binding for O) planner prefers E over O iff

$$\text{PostVar}(O) - \text{PostVar}(E) \geq (C_E - C_O)c_A + O(\epsilon)$$

with $\epsilon = (1 - C_O)^3$ (exact expressions in the paper)

$\Rightarrow$ larger attention cost shifts preference towards less expensive design due to supply effect

Graphical illustration



Non-verifiable design

Model: asymmetric info case

$\Delta(\theta, \varepsilon)$	$X = \theta + \beta_\Delta + \varepsilon$	$p(X)$	$a(X)$
Design	Execution	Evaluation	Action
(researcher)		(editor)	(audience)

- suppose that researcher knows selects the design knowing $\theta + \varepsilon$
  - here, $\varepsilon \sim \mathcal{N}(0, S^2)$ is common sampling uncertainty (with $S^2 = 1$ interpret X as t-stat)

Model: asymmetric info case

$\Delta(\theta, \varepsilon)$	$X = \theta + \beta_\Delta + \varepsilon$	$p(X)$	$a(X)$
Design	Execution	Evaluation	Action
(researcher)		(editor)	(audience)

- suppose that researcher knows selects the design knowing $\theta + \varepsilon$
  - here, $\varepsilon \sim \mathcal{N}(0, S^2)$ is common sampling uncertainty (with $S^2 = 1$ interpret X as t-stat)
- nontrivial potential designs $\Delta \in \mathbb{R}$ parameterized by bias $\beta_\Delta \in \mathbb{R}$

Model: asymmetric info case

$\Delta(\theta, \varepsilon)$	$X = \theta + \beta_\Delta + \varepsilon$	$p(X)$	$a(X)$
Design	Execution	Evaluation	Action
(researcher)		(editor)	(audience)

- suppose that researcher knows selects the design knowing $\theta + \varepsilon$
  - here, $\varepsilon \sim \mathcal{N}(0, S^2)$ is common sampling uncertainty (with $S^2 = 1$ interpret X as t-stat)
- nontrivial potential designs $\Delta \in \mathbb{R}$ parameterized by bias $\beta_\Delta \in \mathbb{R}$
- researcher chooses Δ to max $p(X) - C_\Delta$, where $C_\Delta = c_M|\beta_\Delta|$ [fixed costs in the paper]

Model: asymmetric info case

$\Delta(\theta, \varepsilon)$	$X = \theta + \beta_\Delta + \varepsilon$	$p(X)$	$a(X)$
Design	Execution	Evaluation	Action
(researcher)		(editor)	(audience)

- suppose that researcher knows selects the design knowing $\theta + \varepsilon$
  - here, $\varepsilon \sim \mathcal{N}(0, S^2)$ is common sampling uncertainty (with $S^2 = 1$ interpret X as t-stat)
- nontrivial potential designs $\Delta \in \mathbb{R}$ parameterized by bias $\beta_\Delta \in \mathbb{R}$
- researcher chooses Δ to max $p(X) - C_\Delta$, where $C_\Delta = c_M|\beta_\Delta|$ [fixed costs in the paper]
- audience forms posterior under naive assumption of no bias β_Δ

Optimal publication rule

Optimal publication rule minimizes audience's loss accounting for researcher's best action.

Thm: optimal publication rule takes the form

$$p(X) = \begin{cases} 0 & \text{if } |X| \leq X^\star - \frac{1}{c_M} \\ 1 - c_M(X^\star - |X|) & \text{if } X^\star - \frac{1}{c_M} < |X| < X^\star \\ 1 & \text{otherwise} \end{cases}$$

where $X^\star > t^\star$, where $t^\star$ is optimal threshold with no manipulation

Optimal publication rule

Optimal publication rule minimizes audience's loss accounting for researcher's best action.

Thm: optimal publication rule takes the form

$$p(X) = \begin{cases} 0 & \text{if } |X| \leq X^\star - \frac{1}{c_M} \\ 1 - c_M(X^\star - |X|) & \text{if } X^\star - \frac{1}{c_M} < |X| < X^\star \\ 1 & \text{otherwise} \end{cases}$$

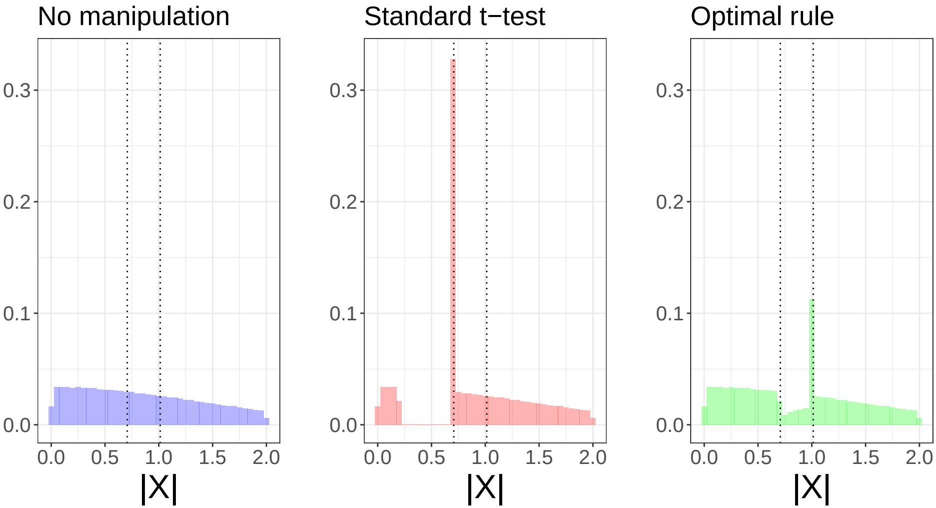
where $X^\star > t^\star$, where $t^\star$ is optimal threshold with no manipulation

- publication mitigates manipulation, but does not eliminate it
- it raises the threshold for guaranteed publication
- it randomizes publication (just) below the threshold
- it publishes some results with $|X| < t^\star$

Implications

Publication rule	Testable observation	Published results	Manipulation
Optimal cutoff rule ignoring manipulation	Large bunching	Only results with $\lvert X\rvert \geq t^\star$	Large
Add randomization below cutoff	No bunching	Many results with $\lvert X\rvert < t^\star$	None
Optimal rule (Randomize + raise cutoff)	Some bunching	Some results with $\lvert X\rvert < t^\star$	Some

Implications



An empirical illustration to medical studies

Application to medical publications

- Pub/ in top medical journals signal for marketing and credibility (Modi et al., 2023)
- However, clinical trials are often expensive and costs are burnt privately

$\Rightarrow$ Moved researchers and FDA to a debate on the use of synthetic or external control groups

Q: publication rule in the absence of a pre-specified (manipulable) experiment?

Application to medical publications

- Pub/ in top medical journals signal for marketing and credibility (Modi et al., 2023)
- However, clinical trials are often expensive and costs are burnt privately
  $\Rightarrow$ Moved researchers and FDA to a debate on the use of synthetic or external control groups
  Q: publication rule in the absence of a pre-specified (manipulable) experiment?

- Head et al. (2015) collect $\sim$ 800,000 p-values from PubMed for medical and pharmaceutical journals. We invert $X_i = \Phi^{-1}(1 - p_i/2)$ to obtain t-stat
- we model $X_i = \theta_i + \beta_i + \varepsilon_i$, with $\varepsilon_i \sim \mathcal{N}(0, 1)$ and $\theta_i \sim \mathcal{N}(0, \eta^2)$
  [zero mean consistent with existing meta-analysis (Bartoš et al., 2023)]

Application to medical publications

- Pub/ in top medical journals signal for marketing and credibility (Modi et al., 2023)
- However, clinical trials are often expensive and costs are burnt privately
  - $\Rightarrow$ Moved researchers and FDA to a debate on the use of synthetic or external control groups
  - Q: publication rule in the absence of a pre-specified (manipulable) experiment?

- Head et al. (2015) collect $\sim$ 800,000 p-values from PubMed for medical and pharmaceutical journals. We invert $X_i = \Phi^{-1}(1 - p_i/2)$ to obtain t-stat
- we model $X_i = \theta_i + \beta_i + \varepsilon_i$, with $\varepsilon_i \sim \mathcal{N}(0, 1)$ and $\theta_i \sim \mathcal{N}(0, \eta^2)$
  [zero mean consistent with existing meta-analysis (Bartoš et al., 2023)]
- Some remarks:
  - 36% of study never published (Vorland et al., 2024); we will assume all these are non-significant for simplicity – also consistent with Showell et al. (2024)
  - 27% of studies in PubMed 2002-15 are pre-specified exp/ (Lamberink et al., 2022).

Application to medical publications

- Pub/ in top medical journals signal for marketing and credibility (Modi et al., 2023)
- However, clinical trials are often expensive and costs are burnt privately
  - $\Rightarrow$ Moved researchers and FDA to a debate on the use of synthetic or external control groups
  - Q: publication rule in the absence of a pre-specified (manipulable) experiment?

- Head et al. (2015) collect $\sim$ 800,000 p-values from PubMed for medical and pharmaceutical journals. We invert $X_i = \Phi^{-1}(1 - p_i/2)$ to obtain t-stat
- we model $X_i = \theta_i + \beta_i + \varepsilon_i$, with $\varepsilon_i \sim \mathcal{N}(0, 1)$ and $\theta_i \sim \mathcal{N}(0, \eta^2)$
  [zero mean consistent with existing meta-analysis (Bartoš et al., 2023)]

- Some remarks:
  - 36% of study never published (Vorland et al., 2024); we will assume all these are non-significant for simplicity – also consistent with Showell et al. (2024)
  - 27% of studies in PubMed 2002-15 are pre-specified exp/ (Lamberink et al., 2022).
- task: calibrate (η^2, c_A, c_M) [sunk fixed costs, can be relaxed]

Calibration

- For publication rules $1\{|X| \geq 1.96\}$, we should observe bunching at 1.96. Therefore take

$$\Phi(1.96) - \Phi\left(1.96 - \frac{1}{c_M}\right) = \underbrace{s_{1.96}}_{\text{share t-stat around 1.96}} \times \underbrace{(1 - 0.36)}_{\textit{Share published}} \times \underbrace{\frac{1}{1 - 0.27}}_{\textit{Share manipulable}}$$

Calibration

- For publication rules $1\{|X| \geq 1.96\}$, we should observe bunching at 1.96. Therefore take

$$\Phi(1.96) - \Phi\left(1.96 - \frac{1}{c_M}\right) = \underbrace{s_{1.96}}_{\text{share t-stat around 1.96}} \times \underbrace{(1-0.36)}_{\textit{Share published}} \times \underbrace{\frac{1}{1-0.27}}_{\textit{Share manipulable}}$$

- For η^2 use 95^{th} quantile (adjusting for pub bias) equal to $3.43(\gg 1.96) \Rightarrow \eta^2 = 1.94$.
- For c_A, choose $t^\star = 1.96 \Rightarrow \sqrt{c_A \frac{1+\eta^2}{\eta^2}} = 1.96$, the standard 5%-critical value

Calibration

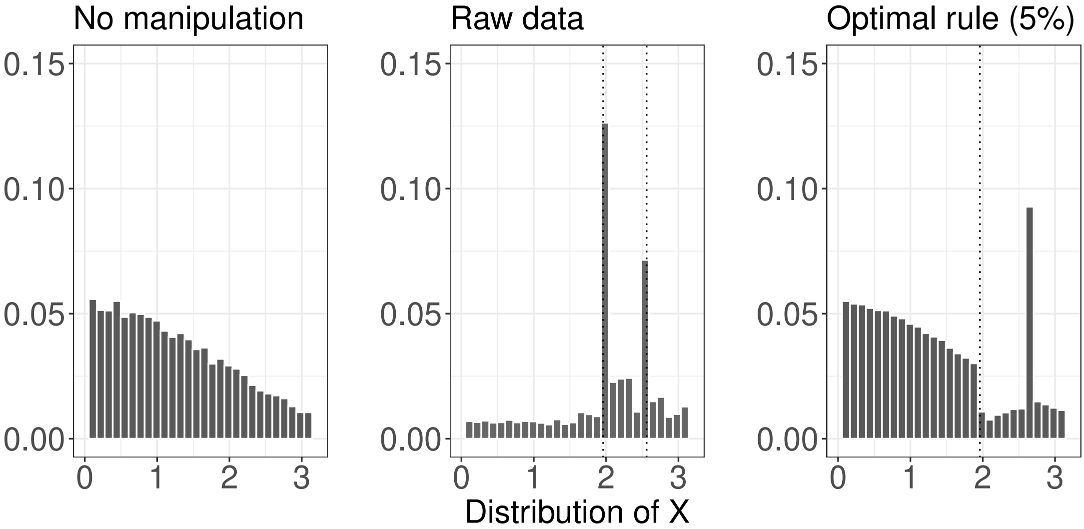
- For publication rules $1\{|X| \geq 1.96\}$, we should observe bunching at 1.96. Therefore take

$$\Phi(1.96) - \Phi\left(1.96 - \frac{1}{c_M}\right) = \underbrace{s_{1.96}}_{\text{share t-stat around 1.96}} \times \underbrace{(1 - 0.36)}_{\textit{Share published}} \times \underbrace{\frac{1}{1 - 0.27}}_{\textit{Share manipulable}}$$

- For η^2 use 95^{th} quantile (adjusting for pub bias) equal to $3.43 (\gg 1.96) \Rightarrow \eta^2 = 1.94$.
- For c_A, choose $t^\star = 1.96 \Rightarrow \sqrt{c_A \frac{1+\eta^2}{\eta^2}} = 1.96$, the standard 5%-critical value

- In the paper, same analysis also with $t^\star = 2.56$.

Distribution of X in equilibrium



Optimal rule vs standard threshold

Publication rule	% Published	Within published findings		
		% $\lvert X \rvert < 1.96$	% Manipulated	Average Bias $\lvert \beta \rvert$
t-test rule $1\{\lvert X \rvert \geq 1.96\}$ (without manipulation)	25%	0%	—	—
t-test rule $1\{\lvert X \rvert \geq 1.96\}$ (with manipulation)	58%	0%	56%	0.31
Optimal rule ($X^\star = 2.64$)	25%	5%	45%	0.11

Choosing an experiment or observational study?

- recent debates for FDA for use of external or "matched" controls vs pre-specified experiments (Food and Drug Administration, 2023) $\Rightarrow$ lower cost but manipulability

[...] In an externally controlled trial, outcomes in participants receiving the test treatment according to a protocol are compared to outcomes in a group of people external to the trial who had not received the same treatment. The external control arm can be a group of people, treated or untreated, from an earlier time (historical control), or it can be a group of people, treated or untreated, during the same time period (concurrent control) but in another setting.

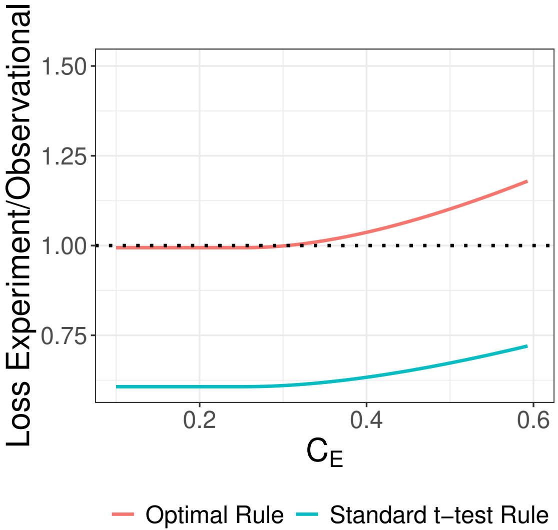
Choosing an experiment or observational study?

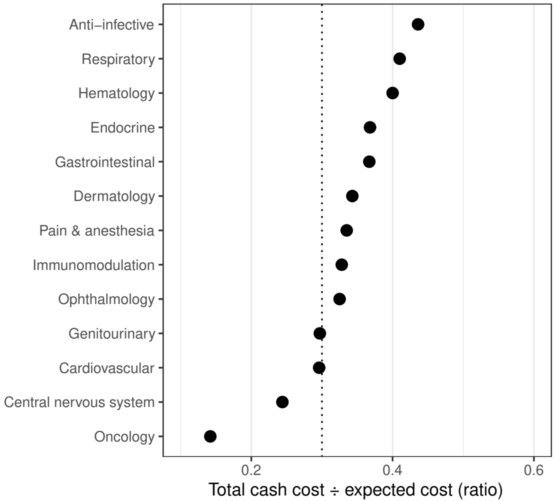
- recent debates for FDA for use of external or "matched" controls vs pre-specified experiments (Food and Drug Administration, 2023) $\Rightarrow$ lower cost but manipulability

  [...] In an externally controlled trial, outcomes in participants receiving the test treatment according to a protocol are compared to outcomes in a group of people external to the trial who had not received the same treatment. The external control arm can be a group of people, treated or untreated, from an earlier time (historical control), or it can be a group of people, treated or untreated, during the same time period (concurrent control) but in another setting.

- takeaways:
  - in some therapeutic areas obs studies may be preferred when exp cost is high
  - but only true if adopt different publication rules for obs studies and exp/

Calibration with our model





Conclusions

Conclusions

- the design of scientific communication shapes research process
  - with a verifiable design planner's preference must depend on research and attention costs
  - with non-verifiable design, optimal publication rule
    - publishes some results that would not be published in absence of manipulation
    - publishes some manipulated findings
    - increases the threshold for guaranteed publication

Conclusions

- the design of scientific communication shapes research process
  - with a verifiable design planner's preference must depend on research and attention costs
  - with non-verifiable design, optimal publication rule
    - publishes some results that would not be published in absence of manipulation
    - publishes some manipulated findings
    - increases the threshold for guaranteed publication
  - application to medical study illustrates that
    - incentives for manipulation can significantly change standard t-test rules
    - choosing synthetic control groups requires different publication standards

Conclusions

- the design of scientific communication shapes research process
  - with a verifiable design planner's preference must depend on research and attention costs
  - with non-verifiable design, optimal publication rule
    - publishes some results that would not be published in absence of manipulation
    - publishes some manipulated findings
    - increases the threshold for guaranteed publication
  - application to medical study illustrates that
    - incentives for manipulation can significantly change standard t-test rules
    - choosing synthetic control groups requires different publication standards

Open questions:

- more complex decisions of planner and researcher
- general models of manipulation
- application to other forms of decisions/loss functions
- ...

Thanks very much, questions?

References

Allcott, H., 2015. Site selection bias in program evaluation. Quarterly Journal of Economics 130, 1117–1165.

Andrews, I., Shapiro, J.M., 2021. A model of scientific communication. Econometrica 89, 2117–2142.

Bartoš, F., Otte, W.M., Gronau, Q.F., Timmers, B., Ly, A., Wagenmakers, E.J., 2023. Empirical prior distributions for Bayesian meta-analyses of binary and time to event outcomes. arXiv preprint arXiv:2306.11468 .

Beets, M.W., Weaver, R.G., Ioannidis, J.P., Geraci, M., Brazendale, K., Decker, L., Okely, A.D., Lubans, D., Van Sluijs, E., Jago, R., et al., 2020. Identification and evaluation of risk of generalizability biases in pilot versus efficacy/effectiveness trials: a systematic review and meta-analysis. International Journal of Behavioral Nutrition and Physical Activity 17, 1–20.

Chassang, S., Padro I Miquel, G., Snowberg, E., 2012. Selective trials: A principal-agent approach to randomized controlled experiments. American Economic Review 102, 1279–1309. URL: `https://www.aeaweb.org/articles?id=10.1257/aer.102.4.1279`, doi:`10.1257/aer.102.4.1279`.

Di Tillio, A., Ottaviani, M., Sørensen, P.N., 2017. Persuasion bias in science: Can economics help? Economic Journal 127, F266–F304.

Efron, B., 2008. Simultaneous inference: when should hypothesis testing problems be combined? Annals of Applied Statistics 2, 197–223.

Food and Drug Administration, 2023. Considerations for the design and conduct of externally controlled trials for drug and biological products. Available at https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-design-and-conduct-externally-controlled-trials-drug-and-biological-products.

Frankel, A., Kasy, M., 2022. Which findings should be published? American Economic Journal: Microeconomics 14, 1–38.

Glaeser, E.L., 2006. Researcher incentives and empirical methods. Working paper, National Bureau of Economic Research.

Head, M.L., Holman, L., Lanfear, R., Kahn, A.T., Jennions, M.D., 2015. The extent and consequences of p-hacking in science. PLoS Biology 13, e1002106.

Henry, E., Ottaviani, M., 2019. Research and the approval process: the organization of persuasion. American Economic Review 109, 911–55.

Kasy, M., Spiess, J., 2023. Optimal pre-analysis plans: Statistical decisions subject to implementability Working paper, University of Oxford and Stanford University.

Lamberink, H.J., Vinkers, C.H., Lancee, M., Damen, J.A., Bouter, L.M., Otte, W.M., Tijdink, J.K., 2022. Clinical trial registration patterns and changes in primary outcomes of randomized clinical trials from 2002 to 2017. JAMA Internal Medicine 182, 779–782.

Manski, C., 2004. Statistical treatment rules for heterogeneous populations. Econometrica 72, 1221–1246.

Manski, C.F., 2015. Randomizing regulatory approval for adaptive diversification and deterrence. The Journal of Legal Studies 44, S367–S385.

Manski, C.F., Tetenov, A., 2016. Sufficient trial size to inform clinical practice. Proceedings of the National Academy of Sciences 113, 10518–10523.

McCloskey, A., Michaillat, P., Forthcoming, 2024. Critical values robust to p-hacking. Review of Economics and Statistics .

Meager, R., 2019. Understanding the average impact of microcredit expansions: A Bayesian hierarchical analysis of seven randomized experiments. American Economic Journal: Applied Economics 11, 57–91.

Modi, N.D., Kichenadasse, G., Hoffmann, T.C., Haseloff, M., Logan, J.M., Veroniki, A.A., Venchiarutti, R.L., Smit, A.K., Tuffaha, H., Jayasekara, H., et al., 2023. A 10-year update to the principles for clinical trial data sharing by pharmaceutical companies: perspectives based on a decade of literature and policies. BMC Medicine 21, 400.

Rosenzweig, M., Udry, C., 2016. External validity in a stochastic world. Technical Report. National bureau of economic research.

Showell, M.G., Cole, S., Clarke, M.J., DeVito, N.J., Farquhar, C., Jordan, V., 2024. Time to publication for results of clinical trials. Cochrane Database of Systematic Reviews .

Spiess, J., Forthcoming, 2025. Optimal estimation when researcher and social preferences are misaligned. Econometrica .

Storey, J.D., 2003. The positive false discovery rate: a Bayesian interpretation and the q-value. The Annals of Statistics 31, 2013–2035.

Tetenov, A., 2016. An economic theory of statistical testing. Working Paper, University of Bristol.

Viviano, D., Wuthrich, K., Niehaus, P., 2025. A model of multiple hypothesis testing arXiv:2104.13367.

Vorland, C.J., Brown, A.W., Kilicoglu, H., Ying, X., Mayo-Wilson, E., 2024. Publication of results of registered trials with published study protocols, 2011-2022. JAMA Network Open 7, e2350688–e2350688.

Wald, A., 1950. Statistical Decision Functions. Wiley.