Publication Design with Incentives in Mind

Ravi Jagadeesan
Stanford

Davide Viviano
Harvard

July, 2024
(Preliminary, comments welcome)

"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

- if publication is costly (e.g., cognitive costs for the audience), optimal policy is to publish a result iff 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 iff 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 iff 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 designs to increase chance of finding significant results

Model in this paper: brief overview

- three agents: an editor, an audience, and a researcher
- state of the world $\theta \sim \mathcal{N}(0, \eta^2)$,

Model in this paper: brief overview

- three agents: an editor, an audience, and a researcher
- state of the world $\theta \sim \mathcal{N}(0, \eta^2)$,

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 $C(\Delta)$ of executing Δ

Model in this paper: brief overview

- three agents: an editor, an audience, and a researcher
- state of the world $\theta \sim \mathcal{N}(0, \eta^2)$,

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 $C(\Delta)$ of executing Δ
3. researcher obtains results $X(\Delta)$

Model in this paper: brief overview

- three agents: an editor, an audience, and a researcher
- state of the world $\theta \sim \mathcal{N}(0, \eta^2)$,

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 $C(\Delta)$ of executing Δ
3. researcher obtains results $X(\Delta)$
4. if published, audience action $a^*(X)$ to minimize expected loss $\mathbb{E}[(a-\theta)^2|X]$

Model in this paper: brief overview

- three agents: an editor, an audience, and a researcher
- state of the world $\theta \sim \mathcal{N}(0, \eta^2)$,

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 $C(\Delta)$ of executing Δ
3. researcher obtains results $X(\Delta)$
4. if published, audience action $a^*(X)$ to minimize expected loss $\mathbb{E}[(a-\theta)^2 | X]$

- editor maximizes audience's welfare net of publication cost c_p per publication

Model: two cases

1. symmetric info: $p(\cdot)$ can depend on X and Δ, researcher does not know θ
   - choosing betw/ experiments with different precisions: $X(\Delta) \sim \mathcal{N}(\theta, S_\Delta^2)$
   - biased vs unbiased experiment $X(\Delta) \sim \mathcal{N}(\theta + \beta_\Delta, S_\Delta^2)$
   - cost varies by design

2. asymmetric info: researcher chooses X and $p(\cdot)$ only depends on X
   - consider (reputational) cost for data manipulation

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
- researcher's information (e.g., no private info vs private info)

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
- researcher's information (e.g., no private info vs private info)

main results: in equilibrium under the editor's optimal publication decision rule:

1. publication is biased towards low cost studies
2. less surprising results, and manipulated results, are sometimes published (with randomization)

Related literature

- economic analysis of statistics [Chassang et al. (2012); Tetenov (2016); Spiess (2024); Henry and Ottaviani (2019); Di Tillio et al. (2017); Viviano et al. (2024); Kasy and Spiess (2023)]
  - we consider the problem of choosing between different study designs
- modeling scientific approval and communication [Frankel and Kasy (2022); Andrews and Shapiro (2021); Glaeser (2006); Manski (2015)]
  - we provide a formal model to choose publication rules that incorporate 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 (2024)]
  - we provide an economic model with incentives for publication rules

Outline

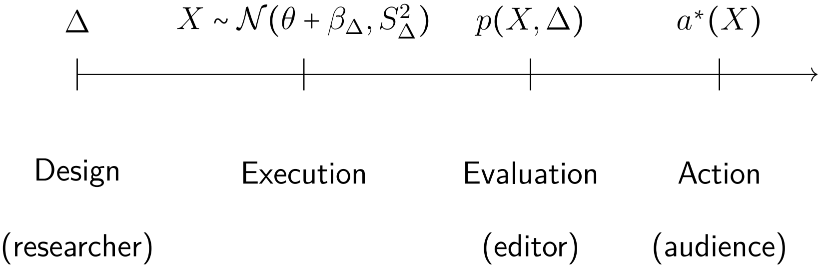
1. which research designs should be incentivized more?
    - Symmetric information case
2. what form (if any) of selective publication is optimal?
    - Allow for asymmetric information case

Publication rules with symmetric information

Model: symmetric info case

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

Model: symmetric info case



- researcher chooses Δ to maximize $\mathbb{E}_X[p(X,\Delta)] - C(\Delta)$

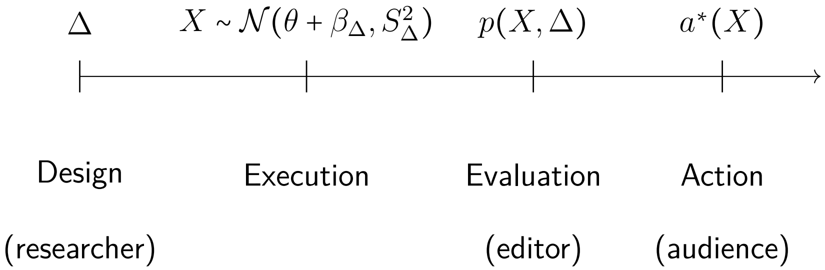
Model: symmetric info case

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

- researcher chooses Δ to maximize $\mathbb{E}_X[p(X,\Delta)] - C(\Delta)$
- editor's objective is

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

Model: symmetric info case



- researcher chooses Δ to maximize $\mathbb{E}_X[p(X,\Delta)] - C(\Delta)$
- editor's objective is

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

Model: symmetric info case

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

- researcher chooses Δ to maximize $\mathbb{E}_X[p(X,\Delta)] - C(\Delta)$
- editor's objective is

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

Examples of design and audience action

- potential designs $\Delta \in \{E\text{xperiment}, O\text{bservational}\}$ ($\varnothing$ possible but omitted)

Examples of design and audience action

- potential designs $\Delta \in \{E\text{xperiment}, O\text{bservational}\}$ ($\varnothing$ possible but omitted)
  - experiment has $\beta_E = 0$, $C(E) > 0$ (e.g. Tetenov, 2016; Viviano et al., 2024)

Examples of design and audience action

- potential designs $\Delta \in \{E\text{xperiment}, O\text{bservational}\}$ ($\varnothing$ possible but omitted)
  - experiment has $\beta_E = 0$, $C(E) > 0$ (e.g. Tetenov, 2016; Viviano et al., 2024)
  - obs study has $\beta_O \sim \mathcal{N}(0, \sigma_O^2), C(O) = 0$ (e.g. Rhys Bernard et al., 2024)

Examples of design and audience action

- potential designs $\Delta \in \{E\text{xperiment}, O\text{bservational}\}$ ($\varnothing$ possible but omitted)
  - experiment has $\beta_E = 0$, $C(E) > 0$ (e.g. Tetenov, 2016; Viviano et al., 2024)
  - obs study has $\beta_O \sim \mathcal{N}(0, \sigma_O^2), C(O) = 0$ (e.g. Rhys Bernard et al., 2024)
- Audience action $a^*(X)$:
  - Naive audience: assume $\beta_\Delta = 0$ for all Δ
  - Sophisticated audience: incorporate info about distribution of β_Δ

Examples of design and audience action

- potential designs $\Delta \in \{E\text{xperiment}, O\text{bservational}\}$ ($\varnothing$ possible but omitted)
    - experiment has $\beta_E = 0$, $C(E) > 0$ (e.g. Tetenov, 2016; Viviano et al., 2024)
    - obs study has $\beta_O \sim \mathcal{N}(0, \sigma_O^2), C(O) = 0$ (e.g. Rhys Bernard et al., 2024)
- Audience action $a^*(X)$:
    - Naive audience: assume $\beta_\Delta = 0$ for all Δ
    - Sophisticated audience: incorporate info about distribution of β_Δ
- For now, focus on $\sigma_O^2 = 0$, i.e., $\beta_O = 0$ and comment as we go throughout

Which observational studies to publish?

(Frankel and Kasy, 2022)

proposition

if the editor is constrained to implement $\Delta = O$bservational, and $\beta_O = 0$
then optimal publication decision rules satisfy

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

where

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

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

Which experiments studies to publish?

proposition

if the editor is constrained to implement $\Delta = E$xperiment,
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

$$X_E^* = \frac{S_E^2 + \eta^2}{\eta^2}\sqrt{c_p}$$

Which experiments studies to publish?

proposition

if the editor is constrained to implement $\Delta = E$xperiment,
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

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

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

Which experiments studies to publish?

proposition

if the editor is constrained to implement $\Delta = E$xperiment,
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

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

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

Experiments vs. observational studies: cheap experiments

- classify experiments based on whether the researcher's IR constraint binds

definition

the experiment is **cheap** if $C_E \leq t^*$, **expensive** otherwise,

Experiments vs. observational studies: cheap experiments

- classify experiments based on whether the researcher's IR constraint binds

definition

the experiment is **cheap** if $C_E \leq t^*$, **expensive** otherwise, $t^* = 2\Phi\Big(-\frac{\sqrt{c_p(S_E^2+\eta^2)}}{\eta^2}\Big)$.

Experiments vs. observational studies: cheap experiments

- classify experiments based on whether the researcher's IR constraint binds

definition

the experiment is **cheap** if $C_E \le t^*$, **expensive** otherwise, $t^* = 2\Phi\Big(-\frac{\sqrt{c_p(S_E^2+\eta^2)}}{\eta^2}\Big)$.

proposition

if the experiment is cheap and $\beta_O = 0$, then optimal publication rules implement $\Delta = O$ iff

$$S_O^2 < S_E^2$$

Experiments vs. observational studies: cheap experiments

- classify experiments based on whether the researcher's IR constraint binds

definition

the experiment is **cheap** if $C_E \le t^*$, **expensive** otherwise, $t^* = 2\Phi\Big(-\frac{\sqrt{c_p(S_E^2+\eta^2)}}{\eta^2}\Big)$.

proposition

if the experiment is cheap and $\beta_O = 0$, then optimal publication rules implement $\Delta = O$ iff

$$S_O^2 < S_E^2$$

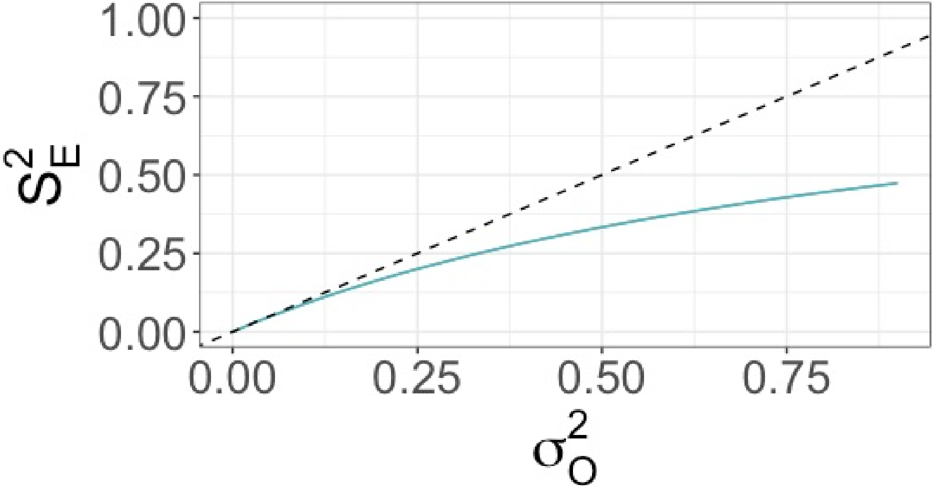
- **implication**: cost does not matter for choice of design with cheap exp

Observational studies with bias vs cheap experiments

- for $\beta_O \neq 0$ need to compare σ_O^2 (variance of the bias) vs S_E^2

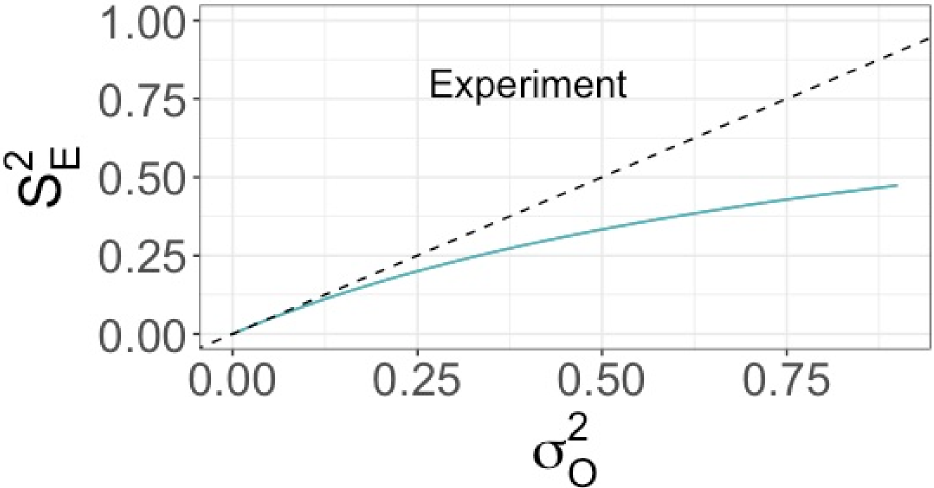
Observational studies with bias vs cheap experiments

- for $\beta_O \neq 0$ need to compare σ_O^2 (variance of the bias) vs S_E^2



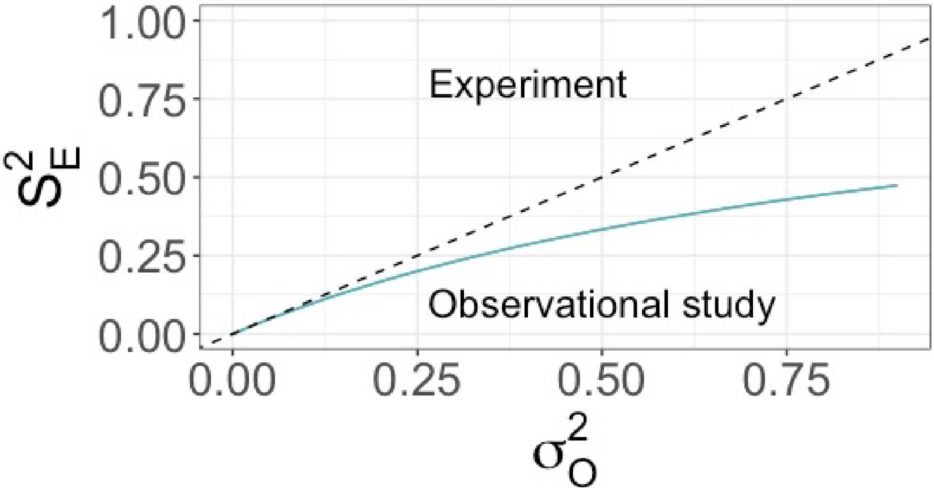
Observational studies with bias vs cheap experiments

- for $\beta_O \neq 0$ need to compare σ_O^2 (variance of the bias) vs S_E^2



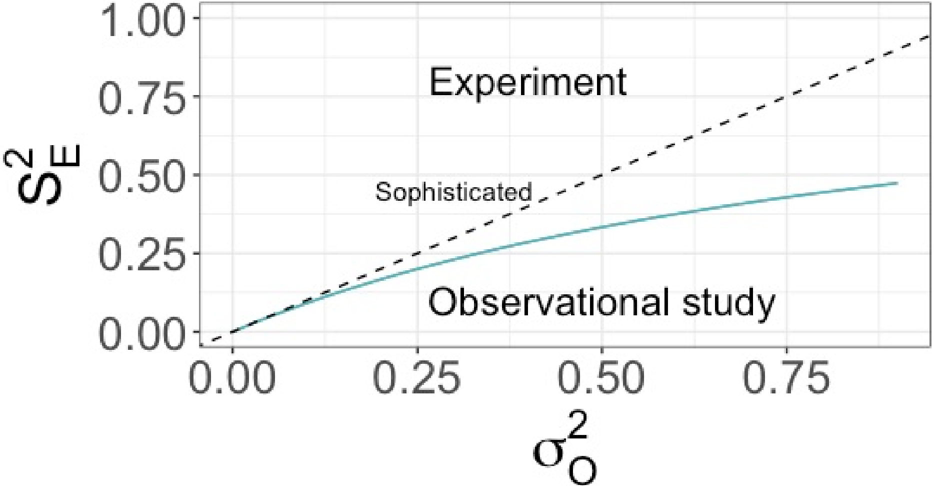
Observational studies with bias vs cheap experiments

- for $\beta_O \neq 0$ need to compare σ_O^2 (variance of the bias) vs S_E^2



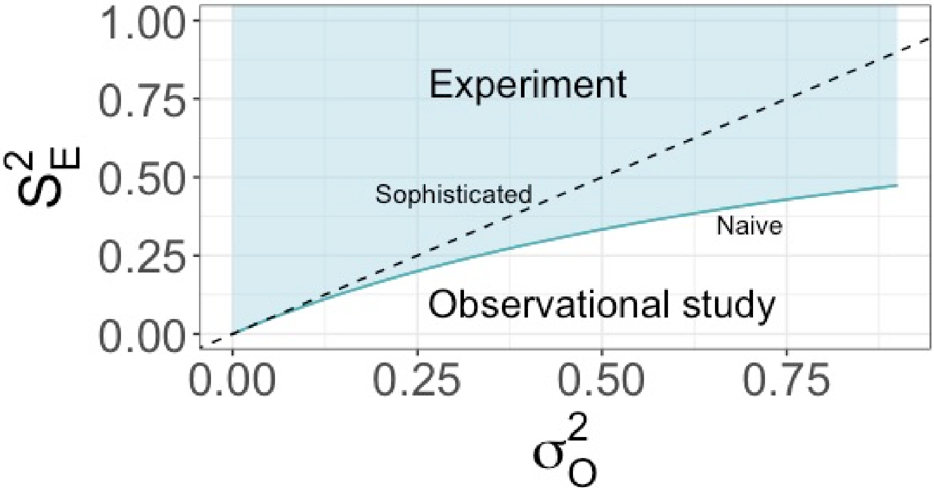
Observational studies with bias vs cheap experiments

- for $\beta_O \neq 0$ need to compare σ_O^2 (variance of the bias) vs S_E^2



Observational studies with bias vs cheap experiments

- for $\beta_O \neq 0$ need to compare σ^2_O (variance of the bias) vs S^2_E



Experiment vs obs study: expensive experiments

- denote $p_O^\star = \mathbb{E}_X[p(X, O)]$ the probability of publishing obs study

Experiment vs obs study: expensive experiments

- denote $p_O^\star = \mathbb{E}_X[p(X, O)]$ the probability of publishing obs study

proposition

Consider an expensive experiment and obs study with $\beta_O = 0$. Let $\Delta \in \{E, O\}$. Then $\Delta = O$ if and only if

$$\underbrace{\frac{S_O^2}{S_O^2 + \eta^2} - \frac{S_E^2}{S_E^2 + \eta^2}}_{\text{Variance comparison}} <$$

Experiment vs obs study: expensive experiments

- denote $p_O^\star = \mathbb{E}_X[p(X,O)]$ the probability of publishing obs study

proposition

Consider an expensive experiment and obs study with $\beta_O = 0$. Let $\Delta \in \{E, O\}$. Then $\Delta = O$ if and only if

$$\underbrace{\frac{S_O^2}{S_O^2+\eta^2} - \frac{S_E^2}{S_E^2+\eta^2}}_{\text{Variance comparison}} < \underbrace{\frac{c_p}{\eta^2}}_{\text{cost of publication}} \times \underbrace{(C_E - p_O^\star)}_{\geq 0}$$

Experiment vs obs study: expensive experiments

- denote $p_O^\star = \mathbb{E}_X[p(X, O)]$ the probability of publishing obs study

proposition

Consider an expensive experiment and obs study with $\beta_O = 0$. Let $\Delta \in \{E, O\}$. Then $\Delta = O$ if and only if

$$\underbrace{\frac{S_O^2}{S_O^2 + \eta^2} - \frac{S_E^2}{S_E^2 + \eta^2}}_{\text{Variance comparison}} < \underbrace{\frac{c_p}{\eta^2}}_{\text{cost of publication}} \times \underbrace{(C_E - p_O^\star)}_{\geq 0} + \underbrace{\mathcal{O}(r)}_{\text{smaller order}}$$

where $r = \frac{c_p^{3/2}}{\eta^3} + (C_E - 1)^3$.

Experiment vs obs study: expensive experiments

- denote $p_O^\star = \mathbb{E}_X[p(X,O)]$ the probability of publishing obs study

proposition

Consider an expensive experiment and obs study with $\beta_O = 0$. Let $\Delta \in \{E, O\}$. Then $\Delta = O$ if and only if

$$\underbrace{\frac{S_O^2}{S_O^2+\eta^2} - \frac{S_E^2}{S_E^2+\eta^2}}_{\text{Variance comparison}} < \underbrace{\frac{c_p}{\eta^2}}_{\text{cost of publication}} \times \underbrace{(C_E - p_O^\star)}_{\geq 0} + \underbrace{\mathcal{O}(r)}_{\text{smaller order}}$$

where $r = \frac{c_p^{3/2}}{\eta^3} + (C_E - 1)^3$.

- Costly experiment are published less even if with smaller variance/MSE

Experiment vs obs study: expensive experiments

- denote $p_O^\star = \mathbb{E}_X[p(X,O)]$ the probability of publishing obs study

proposition

Consider an expensive experiment and obs study with $\beta_O = 0$. Let $\Delta \in \{E, O\}$. Then $\Delta = O$ if and only if

$$\underbrace{\frac{S_O^2}{S_O^2+\eta^2} - \frac{S_E^2}{S_E^2+\eta^2}}_{\text{Variance comparison}} < \underbrace{\frac{c_p}{\eta^2}}_{\text{cost of publication}} \times \underbrace{(C_E - p_O^\star)}_{\geq 0} + \underbrace{\mathcal{O}(r)}_{\text{smaller order}}$$

where $r = \frac{c_p^{3/2}}{\eta^3} + (C_E - 1)^3$.

- Costly experiment are published less even if with smaller variance/MSE
- Larger c_p tilts preference towards obs studies

Illustration: experiment with $S_E = 0$ vs obs study

Illustration: experiment with $S_E = 0$ vs obs study

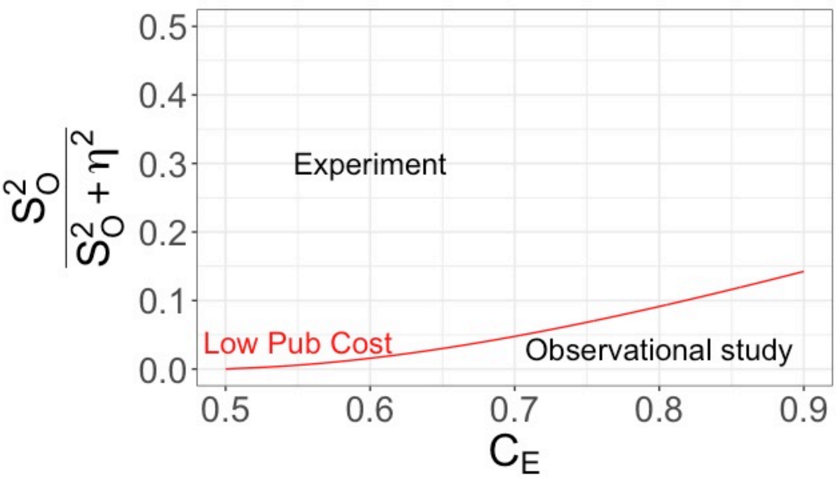
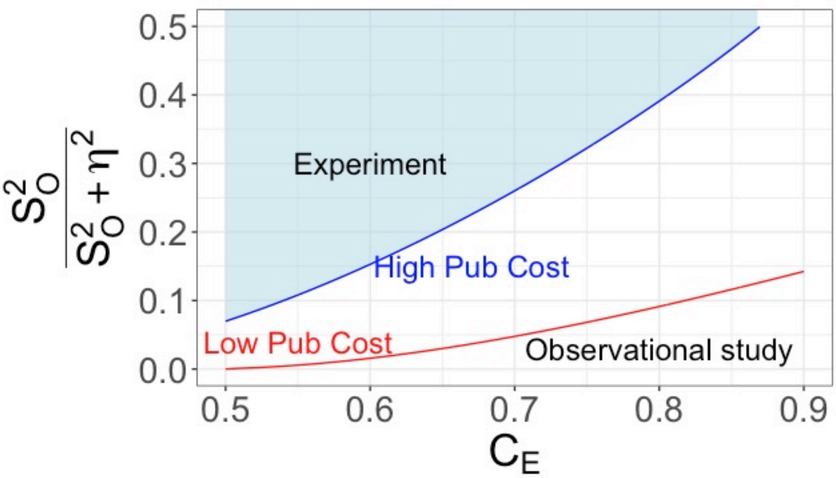


Illustration: experiment with $S_E = 0$ vs obs study



Some takeaways

- for small costs, comparison solely on variance/bias
  - small costs do not affect incentives
- when costs bind editor must reward costly experiments
  - less surprising results must be published
  - when the cost of publication increases, welfare loss can be substantial
  - editor may prefer not to publish costly experiments at all when cost is high

Selective publication rules with asymmetric information

Model: asymmetric info case

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

Model: asymmetric info case

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

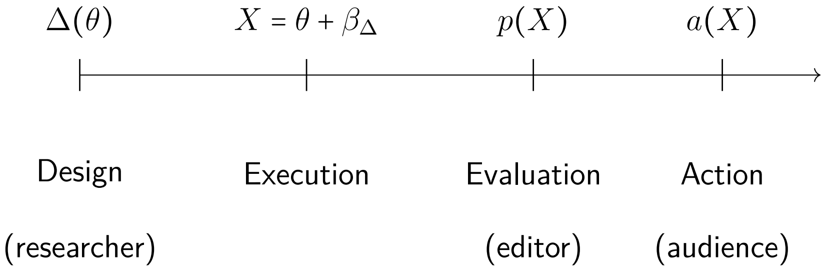
- potential designs $\Delta \in \{\varnothing\} \cup \mathbb{R}$ parameterized by $\beta_\Delta \in \mathbb{R}$

Model: asymmetric info case

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

- potential designs $\Delta \in \{\varnothing\} \cup \mathbb{R}$ parameterized by $\beta_\Delta \in \mathbb{R}$
- researcher chooses Δ to maximize $p(X) - C(\Delta)$ and knows θ

Model: asymmetric info case



- potential designs $\Delta \in \{\varnothing\} \cup \mathbb{R}$ parameterized by $\beta_\Delta \in \mathbb{R}$
- researcher chooses Δ to maximize $p(X) - C(\Delta)$ and knows θ
- for tractability: suppose that $S_\Delta = 0$ and that $C(\Delta) = c_d|\beta_\Delta|$ (can be extended with fixed costs)

Interpretations of manipulation of research design

- researcher has private info about θ through a pilot study
  - knows that effects are larger for certain subgroups, or in certain villages
  - under $S_\Delta = 0$, our framework is equivalent to assuming researcher knows X (worst-case over researcher's knowledge)

Interpretations of manipulation of research design

- researcher has private info about θ through a pilot study
  - knows that effects are larger for certain subgroups, or in certain villages
  - under $S_\Delta = 0$, our framework is equivalent to assuming researcher knows X (worst-case over researcher's knowledge)

- researcher observes the data before deciding which results to report
  - p-hacking: often problematic in scientific research (Elliott et al., 2022)
  - e.g., suppose researchers can choose from a large set of specifications
  - if they have access to large data $D(\theta)$ can learn $\mathbb{E}[\theta|D(\theta)] \approx \theta$

Interpretations of manipulation of research design

- researcher has private info about θ through a pilot study
  - knows that effects are larger for certain subgroups, or in certain villages
  - under $S_\Delta = 0$, our framework is equivalent to assuming researcher knows X (worst-case over researcher's knowledge)

- researcher observes the data before deciding which results to report
  - p-hacking: often problematic in scientific research (Elliott et al., 2022)
  - e.g., suppose researchers can choose from a large set of specifications
  - if they have access to large data $D(\theta)$ can learn $\mathbb{E}[\theta|D(\theta)] \approx \theta$

- manipulation cost $c_d|\beta_\Delta|$ can be interpreted as physical, or reputational

Optimal publication decision rules with a naive audience

- suppose the audience is **naive** in that it assumes $\beta_\Delta = 0$
  - that is, if result X is published, the action is $a = X$
- tie-breaking rule where indifferent researchers maximize editor's utility

Optimal publication decision rules with a naive audience

- suppose the audience is **naive** in that it assumes $\beta_\Delta = 0$
  - that is, if result X is published, the action is $a = X$
- tie-breaking rule where indifferent researchers maximize editor's utility

theorem

there exists $X^* \in \left(\sqrt{c_p}, \sqrt{c_p} + \frac{1}{c_d}\right)$ such that optimal publication rules satisfy

$$p(X) = \begin{cases} 0 & \text{if } |X| \leq X^* - \frac{1}{c_d} \\ & \\ & \end{cases}$$

Optimal publication decision rules with a naive audience

- suppose the audience is **naive** in that it assumes $\beta_\Delta = 0$
  - that is, if result X is published, the action is $a = X$
- tie-breaking rule where indifferent researchers maximize editor's utility

theorem

there exists $X^* \in \left(\sqrt{c_p}, \sqrt{c_p} + \frac{1}{c_d}\right)$ such that optimal publication rules satisfy

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

Optimal publication decision rules with a naive audience

- suppose the audience is **naive** in that it assumes $\beta_\Delta = 0$
  - that is, if result X is published, the action is $a = X$
- tie-breaking rule where indifferent researchers maximize editor's utility

theorem

there exists $X^* \in \left(\sqrt{c_p}, \sqrt{c_p} + \frac{1}{c_d}\right)$ such that optimal publication rules satisfy

$$p(X) = \begin{cases} 0 & \text{if } |X| \leq X^* - \frac{1}{c_d} \\ 1 - c_d(X^* - |X|) & \text{if } X^* - \frac{1}{c_d} < |X| < X^* \\ 1 & \text{if } |X| \geq X^* \end{cases}$$

Intuition of the result: what if editor ignores p-hacking

- suppose we consider a publication rule editor ignores p-hacking

Intuition of the result: what if editor ignores p-hacking

- suppose we consider a publication rule editor ignores p-hacking
  - $X \geq X^*, X^* = \sqrt{c_p}$ is optimal rule

Intuition of the result: what if editor ignores p-hacking

- suppose we consider a publication rule editor ignores p-hacking
  - $X \geq X^*, X^* = \sqrt{c_p}$ is optimal rule
  - researchers with $|\theta| > \sqrt{c_p} - \frac{1}{c_d}$ will p-hack with $|\beta_\Delta| = \sqrt{c_p} - |\theta|$

Intuition of the result: what if editor ignores p-hacking

- suppose we consider a publication rule editor ignores p-hacking
  - $X \geq X^*, X^* = \sqrt{c_p}$ is optimal rule
  - researchers with $|\theta| > \sqrt{c_p} - \frac{1}{c_d}$ will p-hack with $|\beta_\Delta| = \sqrt{c_p} - |\theta|$
- we will observe bunching at $\sqrt{c_p}$
- whenever c_d is not too big, this can be harmful for social welfare!

Intuition of the result: editor's best action

- suppose now we add randomization *around* $(X^* - \frac{1}{c_d}, X^*), X^* = \sqrt{c_p}$
  - here researchers indifferent betw/ p-hacking and not p-hacking
  - but we still publish some non-surprising results

Intuition of the result: editor's best action

- suppose now we add randomization *around* $(X^* - \frac{1}{c_d}, X^*), X^* = \sqrt{c_p}$
  - here researchers indifferent betw/ p-hacking and not p-hacking
  - but we still publish some non-surprising results
- raise X^*
  - stop publishing some non-surprising
  - raise X^* just enough so that loss from p-hacking is second order

Intuition of the result: editor's best action

- suppose now we add randomization *around* $(X^* - \frac{1}{c_d}, X^*), X^* = \sqrt{c_p}$
  - here researchers indifferent betw/ p-hacking and not p-hacking
  - but we still publish some non-surprising results
- raise X^*
  - stop publishing some non-surprising
  - raise X^* just enough so that loss from p-hacking is second order
- some researchers just below X^* will p-hack
  - some p-hacking can improve welfare on the margin

Some takeaways

- we should increase critical thresholds X^* with p-hacking
  - the increase should not be "too large"
  - some p-hacking can be of second order when results are indeed surprising
- below X^*, and above the threshold with no p-hacking, we should randomize
  - some results with "non-significant" effects should get published
  - randomization will avoid harmful p-hacking on the margin

Conclusions

- we study publication decisions taking into account researcher's incentives
- relevant to compare different study designs or allow for p-hacking

Conclusions

- we study publication decisions taking into account researcher's incentives
- relevant to compare different study designs or allow for p-hacking

- with symmetric info:
  - choice between different designs should take into account not only MSE-type comparisons but also costs of the study
  - cost of the experiment matters for publication when these are binding

Conclusions

- we study publication decisions taking into account researcher's incentives
- relevant to compare different study designs or allow for p-hacking

- with symmetric info:
  - choice between different designs should take into account not only MSE-type comparisons but also costs of the study
  - cost of the experiment matters for publication when these are binding

- with asymmetric info:
  - publication rules should increase critical threshold for publication
  - randomize below such a threshold
  - some p-hacking may occur in equilibrium when its effects are second order

Conclusions

- we study publication decisions taking into account researcher's incentives
- relevant to compare different study designs or allow for p-hacking

- with symmetric info:
  - choice between different designs should take into account not only MSE-type comparisons but also costs of the study
  - cost of the experiment matters for publication when these are binding

- with asymmetric info:
  - publication rules should increase critical threshold for publication
  - randomize below such a threshold
  - some p-hacking may occur in equilibrium when its effects are second order

- Thanks much! Draft soon online, for questions email us.

References

Allcott, H. (2015). Site selection bias in program evaluation. *Quarterly Journal of Economics 130*(3), 1117–1165.

Andrews, I. and J. M. Shapiro (2021). A model of scientific communication. *Econometrica 89*(5), 2117–2142.

Beets, M. W., R. G. Weaver, J. P. Ioannidis, M. Geraci, K. Brazendale, L. Decker, A. D. Okely, D. Lubans, E. Van Sluijs, R. Jago, 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., G. Padro I Miquel, and E. Snowberg (2012). Selective trials: A principal-agent approach to randomized controlled experiments. *American Economic Review 102*(4), 1279–1309.

Di Tillio, A., M. Ottaviani, and P. N. Sørensen (2017). Persuasion bias in science: Can economics help? *The Economic Journal 127*(605), F266–F304.

Efron, B. (2008). Simultaneous inference: when should hypothesis testing problems be combined? *Annals of Applied Statistics 2*(1), 197–223.

Elliott, G., N. Kudrin, and K. Wüthrich (2022). Detecting p-hacking. *Econometrica 90*(2), 887–906.

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

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

Henry, E. and M. Ottaviani (2019). Research and the approval process: the organization of persuasion. *American Economic Review 109*(3), 911–55.

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

Manski, C. (2004). Statistical treatment rules for heterogeneous populations. *Econometrica 72*(4), 1221–1246.

Manski, C. F. (2015). Randomizing regulatory approval for adaptive diversification and deterrence. *The Journal of Legal Studies 44*(S2), S367–S385.

Manski, C. F. and A. Tetenov (2016). Sufficient trial size to inform clinical practice. *Proceedings of the National Academy of Sciences 113*(38), 10518–10523.

McCloskey, A. and P. Michaillat (2024). Critical values robust to p-hacking. *Review of Economics and Statistics*, 1–35.

Meager, R. (2019). Understanding the average impact of microcredit expansions: A bayesian hierarchical analysis of seven randomized experiments. *American Economic Journal: Applied Economics 11*(1), 57–91.

Rhys Bernard, D., G. Bryan, S. Chabé-Ferret, J. De Quidt, J. Fliegner, and R. Rathelot (2024). How much should we trust observational estimates? accumulating evidence using randomized controlled trials with imperfect compliance. Working Paper, Toulouse School of Economics.

Rosenzweig, M. and C. Udry (2016). External validity in a stochastic world. Technical report, National bureau of economic research.

Spiess, J. (2024). Optimal estimation when researcher and social preferences are misaligned. Working paper, Stanford University.

Storey, J. D. (2003). The positive false discovery rate: a Bayesian interpretation and the q-value. *The Annals of Statistics 31*(6), 2013–2035.

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

Viviano, D., K. Wuthrich, and P. Niehaus (2024). A model of multiple hypothesis testing. arXiv:2104.13367.

Wald, A. (1950). *Statistical Decision Functions*. Wiley.