Publication Design with Incentives in Mind*

Ravi Jagadeesan† Davide Viviano‡

July 14, 2025

Abstract

The publication process both determines which research receives the most attention, and influences the supply of research through its impact on the researcher's private incentives. We introduce a framework to study optimal publication decisions when researchers can choose (i) whether or how to conduct a study and (ii) whether or how to manipulate the research findings (e.g., via selective reporting or data manipulation). When manipulation is not possible, but research entails substantial private costs for the researchers, it may be optimal to incentivize cheaper research designs even if they are less accurate. When manipulation is possible, it is optimal to publish some manipulated results, as well as results that would have not received attention in the absence of manipulability. Even if it is possible to deter manipulation, such as by requiring pre-registered experiments instead of (potentially manipulable) observational studies, it is suboptimal to do so when experiments entail high research costs.

Our model calibrated to medical studies shows that optimal publication rules taking the researchers' best response into account, should increase by about 0.3 the standard cutoffs for which results are considered "significant", but also increase substantially the share of published non-significant findings.

1 Introduction

Publication decisions shape the process of scientific communication. By selecting what to publish, journals affect which findings receive the most attention and can inform the public about the state of the world. The design of publication rules has therefore motivated recent debates on how statistical significance should affect publication when the goal is to direct attention to the most informative results (e.g., Abadie, 2020; Frankel and Kasy, 2022).

*The first version of this paper was circulated in October, 2024. We thank Isaiah Andrews, Arun Chandrasekhar, Kevin Chen, Matt Gentzkow, Max Kasy, Toru Kitagawa, Paul Niehaus, Marco Ottaviani, Jesse Shapiro, Jann Spiess, Elie Tamer, Alex Tetenov, and Kaspar Wüthrich for helpful comments. We thank Karthik Seetharaman for exceptional research assistance.

†Department of Economics, Stanford University. Email address: ravi.jagadeesan@gmail.com.

‡Department of Economics, Harvard University. Email address: dviviano@fas.harvard.edu.

However, the publication process also affects the supply of research by influencing researcher's incentives about how to conduct research. Researchers have many degrees of freedom about how to conduct their research, such as how and where to run an experiment (e.g., Allcott, 2015; Gechter and Meager, 2022); the size, cost, and effort associated with the study (e.g., Thompson and Barber, 2000; Grabowski et al., 2002); and which findings to report from a given study (Brodeur et al., 2020; Elliott et al., 2022). We refer broadly to their choices about each of these aspects of a study broadly as a *research design*.

Researcher's private incentives may influence how they choose their design. Yet, "while economists assiduously apply incentive theory to the outside world, we use research methods that rely on the assumption that social scientists are saintly automatons" (Glaeser, 2006). This raises the questions of when and how researchers' incentives should impact the design of publication processes, and, more broadly, the optimal allocation of attention to research.

This paper studies optimal publication decisions when researchers chooses the research design based on private costs and benefits. We frame this question as a mechanism design problem: a social planner (principal) optimizes a publication rule, taking into account the incentives of the researcher (agent). The social planner aims to use the publication process to allocate the attention of the audience to the most informative research findings.

More concretely, as in Frankel and Kasy (2022) (building in turn on Wald (1950)), we suppose that research results impact the decisions of an audience who has limits on how much attention they can devote to research. The social planner seeks to publish results that are most important for the audience, net of the cost of (or taking into account constraints on) publishing or attention. Due to attention constraints, not all results will be published, which leaves the planner with a non-trivial trade-off about which results and designs to publish. We then introduce a model in which publication decisions affect the supply of research in the first place (Section 2): given the publication rule, the researcher chooses the design that maximizes her value for publication (or other attention) net of research costs.

As a concrete example, consider a medical journal seeking to decide whether to publish results from a clinical study. The journal wants to convey accurate information about the drug's efficacy in the study (e.g., DeMaria, 2022; Ana, 2004) and direct the audience's attention to the most effective drugs. This objective is motivated by scientific practice. For example, the stated mission of the New England Journal of Medicine is

> "to publish the best research and information at the intersection of biomedical science and clinical practice and to present this information in understandable, clinically useful formats that inform health care practice and improve patient outcomes."

However, researchers may respond in how they conduct their studies through the size, length, cost of the studies, composition of the control groups (Thorlund et al., 2020), and—in some cases—in which specific findings to report (e.g., Riveros et al., 2013; Shinohara et al., 2015).

We draw a dichotomy between researchers' incentives about (i) whether or how to conduct a study and (ii) which findings to report (e.g., via data manipulation or selective reporting).

We first focus on (i) and abstract from data manipulation and selective reporting; we therefore suppose that the research design is observable and verifiable (Section 3). For example, a researcher could choose between experiments with different mean-squared errors and costs and, by pre-committing to an analysis plan, truthfully report an unbiased estimate of a treatment effect. The planner can direct the attention of the audience to any executed study by publishing it; publication can depend on a study's design and results.

Returning to our example, after providing a new drug to a treatment group, scientists can evaluate its efficacy by comparing the treatment group to either an experimental placebo group or using a pre-specified synthetic control group obtained from historical medical records (Popat et al., 2022; Yin et al., 2022). Using a synthetic control group can have large impact on the supply of medical research and new drugs by decreasing research costs (Jahanshahi et al., 2021; Food and Drug Administration, 2023; Wong et al., 2014). However, using a synthetic control group may increase the estimate's mean-squared error due to lack of randomization (see Raices Cruz et al., 2022; Rhys Bernard et al., 2024), raising the question of whether a social planner should allow or incentivize their use. (A closely related application is choosing between two experiments with different numbers of participants.)

We thus suppose that the publication process affects the supply of research through an individual rationality constraint: for a researcher to be willing to conduct a study, they must be compensated with a large enough *ex ante* publication probability. Without these individual rationality constraints, the first-best publication decision rule would ignore researchers' costs and publish only studies conducted with the lowest possible mean-squared error; due to attention constraints or publication costs, not all such studies would be published. When the feasible designs are inexpensive for the researcher to conduct, the first-best policy can be feasibly implemented by the planner; the individual rationality constraint does not bind. Thus, small research costs do not affect the planning problem.

However, when the research cost of accurate designs is above a tipping point, the first-best publication rule does not compensate the researcher enough to incentivize them to choose such a design. As a result, the planner faces a trade-off between providing attention to results that are not valuable enough to deserve attention, and rewarding the researcher enough to make them willing to use a costly design in the first place.

This trade-off has implications for which designs the planner should publish. For example, suppose that clinical studies can be conducted based on either quasiexperiments based on synthetic controls, which are less expensive but less informative, or full experiments, which are costly but more informative. Incentivizing researchers to conduct the costly experiment would require publishing results from such a study, even when doing so misdirects attention toward treatments with negligible effects on patients. In this case, the planner may prefer publishing studies with potentially larger mean-squared error to avoid having to publish too many results from the experiment; researchers will then choose not to use costly experiments in equilibrium. This analysis suggests that due to the interaction between attention costs and supply effects, publishing medical studies with synthetic control groups can be desirable when it is sufficiently costly for researchers to use experimental control groups.

We next turn in Section 4 to scenarios where the researcher, *after* observing the study's results, can report a biased statistic from such a study. The bias is chosen by the researcher and unobserved by the planner, but incurs a cost to the researcher that is increasing in the bias (e.g., capturing reputational costs). The audience is unaware of possible manipulations of published findings, taking results at face value. We think of this model as a stylized description of settings with potential data manipulation or selective reporting. For instance, in the absence of a precise pre-specification, we may be concerned that researchers select a synthetic control group from medical records based on their observed outcomes.

Each publication rule generates a different degree of manipulation. For example, suppose that the planner used the publication rule that would be optimal without manipulation—i.e., publishing if the reported statistic is above a cutoff. Then, researchers whose results would be close to the cutoff would manipulate their data or analyses to reach the cutoff.[1] Publishing substantially manipulated results would incur a substantial loss for the audience.

One approach that has been proposed is to completely deter manipulation. An extreme form of this approach would be to make publication dependent only on the design, and not on results.[2] However, this approach generally incurs substantial costs by directing the audience's attention to results that would not affect their decisions, and is not optimal.

By contrast, we show that the optimal publication rule under manipulability has three key features. First, it increases the cutoff for which findings always get published compared to settings without data manipulation. Second, just below this cutoff, it randomizes publication chances, making the researcher indifferent about whether to manipulating their results. As a result, with nonzero probability, the planner publishes findings that are not truly valuable enough to deserve attention, and would not be published without the possibility of manipulation. Third, some manipulation *does* occur in equilibrium.

[1] This manipulation would lead to bunching at the cutoff, which is consistent with bunching that has been documented in settings with p-hacking (e.g., Elliott et al., 2022).

[2] In practice, this approach can be implemented by committing to publication based on pre-analysis plans, as the *Journal of Development Economics* and the *Journal of Clinical Epidemiology* do (among others).

To gain some intuition for these features of the optimum, consider first the optimal publication rule without manipulation. The planner can eliminate the researcher's incentive to manipulate by publishing results below the cutoff with a positive probability. However, some results that do not affect the audience's decision do get published as a result, which is undesirable in the presence of attention costs or constraints. To counteract this effect, it then is optimal for the planner to increase the cutoff to guarantee publication.

With a more stringent cutoff, some results that would be published without manipulation are no longer published with probability one. The planner then encourages researchers with results in this range to engage in a limited degree of manipulation to increase their publication chances. The loss from publishing some slightly manipulated studies is second-order relative to the (first-order) gain from publishing results that do affect the audience's decisions.

To prove that the optimal publication rule takes the described form, we formulate the social planner's problem as a mechanism design problem with "false moral hazard" due to the researcher choosing a manipulation after learning the true results. The absence of direct transfers and the inability to reward the researcher with a publication probability above one make the mechanism design problem effectively one with limited transfers. As a result, standard methods as in Mirrlees (1971) and Myerson (1981) do not apply. We solve the mechanism design problem by identifying the precise pattern of binding incentive constraints.

In Section 5, we bring our model to the data and study optimal publication rules under manipulation (p-hacking). Using about 800,000 t-statistics collected across medical and pharmaceutical journals by Head et al. (2015), we estimate the researcher's cost of manipulation and other relevant parameters in the model. Estimation of the cost of manipulation leverages discontinuity in the distribution of the t-statistics at relevant cutoffs such as 2.33. Using these data, we then compute optimal publication rules taking into account the researcher's best response.

We find that the threshold at which results should be published with probability one increases from 2.33 (1.96) for 1% (5%)-α test to 2.63 (2.26). Second, despite this larger threshold, *more* results must be published in equilibrium, including "non-significant" results. Third, the share of results which are published is above 20% for a cost of publication corresponding to 1%-significance test. This share is substantially larger than what we would observe under the naive significance test at 1%-level (equal to 4% in the presence of manipulation).

These results provide practical guidance for the optimal rule: the optimal publication rule should decrease both p-hacking and publication bias in equilibrium than naive α-tests, but it does not reduce these to zero. For practice, we derive an easy-to-compute "publication score" that defines relevance of the findings.

As a final exercise in Section 6 we then combine these two models to ask whether the planner should mandate researchers to send a costly signal that deters them from manipulating data. For example, the planner could mandate that researchers run costly experiments that adhere to pre-analysis plans, rather than also allowing for (cheaper) observational studies. Without accounting for the effects of making research more costly on the supply of research, the planner would always require such signals to ensure that results are unbiased (Spiess, 2025; Kasy and Spiess, 2023). For example, the planner would not publish observational studies if equally informative experiments are feasible, no matter their costs.

However, to ensure that the researcher is actually willing to conduct the experiment, the planner needs to reward the researcher with enough attention through a high enough publication probability. With large attention costs (or binding publication constraints), this makes incentivizing experiments costly enough for the planner, that they prefer to publish observational studies, even though manipulation may occur. Therefore, taking into account the supply effects of the publication process, the planner may prefer to drive attention towards studies whose results may be manipulable, and hence possibly biased.

Related literature. This paper connects to a growing literature that studies economic models to analyze statistical protocols. In the context of scientific communication, Andrews and Kasy (2019), Abadie (2020), Andrews and Shapiro (2021), Kitagawa and Vu (2023), and (most closely related to our paper) Frankel and Kasy (2022) have analyzed how research findings are or should be reported to inform the public. Our analysis builds on this literature by introducing a model that incorporates researcher's incentives and studying how these incentives shape the optimal design of the optimal publication process. None of these references account for the researcher's private incentives in the design of the study.

We connect to a broad literature on statistical decision theory (e.g., Wald, 1950; Savage, 1951; Manski, 2004; Hirano and Porter, 2009; Tetenov, 2012; Kitagawa and Tetenov, 2018), focusing in particular on settings with private researcher incentives. Other work in this line includes Chassang et al. (2012), Tetenov (2016), Manski and Tetenov (2016), Banerjee et al. (2017), Henry and Ottaviani (2019), Banerjee et al. (2020), Di Tillio et al. (2021), Williams (2021), Bates et al. (2022, 2023), Libgober (2022), Yoder (2022), Kasy and Spiess (2023), McCloskey and Michaillat (2024), Spiess (2025) and Viviano et al. (2024). Different from these papers, we analyze settings where the researcher may choose the research design absent private information, and settings in which the researcher can choose the design and manipulate reported findings with private information. This allows us to formally study ideas such as when/whether unsurprising results should be published, and whether manipulation should occur in equilibrium (Glaeser, 2006).

In particular, an important distinction from some of these models studying approval decisions, such as Tetenov (2016), Bates et al. (2022, 2023), and Viviano et al. (2024), is that these papers assume that researchers truthfully report the statistics sampled from their study and abstract from questions about data manipulation or optimal design of the experiment studied in this paper, and assume no attention costs. Spiess (2025) and Kasy and Spiess (2023) study models of scientific communication, without, however, focusing on optimal publication rules studied here. Different from these references, here we incorporate the researcher's costs of the design (and misreporting), which, we show, leads to qualitatively different solutions in the amount of misreporting. Di Tillio et al. (2021) study the different question of selective sampling, and Henry and Ottaviani (2019) study the question of decisions with continuous and sequential access to the data, different from the question of selective design choice studied here. McCloskey and Michaillat (2024) and Andrews and Kasy (2019) propose statistical adjustments for p-hacking or publication bias without allowing researchers to best-respond to changes in such adjustments. Here we study optimal publication rules in equilibrium, taking into account the researcher's best response.

Finally, a large empirical and econometric literature has documented several aspects of the research process, including selective reporting, data manipulation, specification search, as well as site selection bias and observational studies' bias (e.g., Allcott, 2015; Gechter and Meager, 2022; Rosenzweig and Udry, 2020; Elliott et al., 2022; Brodeur et al., 2020; Miguel, 2021; Olken, 2015; Banerjee et al., 2020; Rhys Bernard et al., 2024). Our contribution here is to provide a formal theoretical model that studies how incentives interact with some of these choices, shedding light on qualitative aspects of optimal publication rules.

2 Setup

Consider three agents: a researcher, a (representative) audience, and a social planner. The audience and social planner are interested in learning a parameter $\theta_0 \in \mathbb{R}$. All agents share a common prior $\theta_0 \sim \mathcal{N}(0, \eta^2)$, whose mean is normalized to 0 without loss of generality.

A researcher conducts a study to inform the audience about θ_0. A study is summarized by (X, Δ), where Δ denotes the design and

$$X(\Delta)|\theta_0 \sim \mathcal{N}\Big(\theta_0 + \beta_\Delta, S_\Delta^2\Big) \tag{1}$$

the results. Here, β_Δ and S_Δ^2 are the average bias, and variance, of design Δ, respectively.

If a study is conducted, it will be evaluated according to a publication rule $p(X, \Delta)$ with values in $[0, 1]$. Here, $p(X, \Delta)$ represents the probability of publishing the study, which is assumed to be a Borel measurable function of (X, Δ).

Conditional on publication, the audience forms posterior beliefs about θ_0 using Bayes' Rule assuming $\beta_\Delta = 0$. Conditional on non-publication, the audience's posterior mean equals its prior mean (zero), as would arise from Bayes' Rule with $p(\cdot)$ symmetric in X and β_Δ symmetric around 0. Based on these beliefs, the audience takes an action to minimize mean-squared distance from θ_0; thus, the audience's action $a_p^\star(X, \Delta)$ is the posterior mean

$$a_p^\star(X, \Delta) = \begin{cases} \frac{X\eta^2}{S_\Delta^2 + \eta^2} & \text{if the study is published} \\ 0 & \text{otherwise} \end{cases}.$$

Given results $X = X(\Delta)$ for a design Δ, and a parameter θ_0, the planner incurs a loss

$$\mathcal{L}_p(X, \Delta, \theta_0) = \mathbb{E}_p\left[\left(\theta_0 - a_p^\star(X, \Delta)\right)^2\right] - c_p p(X, \Delta), \tag{2}$$

where $\mathbb{E}_p$ denotes expectation with respect to any stochasticity in the publication decision rule. Thus, the planner minimizes the expected loss of the audience, net of a cost of publication or attention. The quantity c_p captures the publication or attention costs or constraints.

Given a design Δ, a publication rule p, and results X, the researcher's expected payoff is

$$v_p(X, \Delta) = p(X, \Delta) - C_\Delta \tag{3}$$

where $C_\Delta \leq 1$ is the researcher's cost of executing design Δ.[3] (Here, we normalize the value of publication to 1.) As is standard, whenever the researcher is indifferent between two designs, we implicitly assume she chooses the design that minimizes the planner's expected loss.

3 Publication rules under verifiable designs

This section studies optimal publication rules when the planner can observe the research design, and condition publication on it. We define a design Δ *unbiased on average* if $\beta_\Delta = 0$, and focus on such designs for simplicity in this section (and defer designs that are not unbiased on average to the following section).

Our analysis proceeds in three steps. We first characterize the optimal publication rule subject to the constraint of incentivizing the researcher to implement a particular design Δ. We then characterize which designs are worth incentivizing relative to an outside option. Last, we characterize the optimal publication rule that chooses between multiple designs.

[3] We assume $C_\Delta \leq 1$ simply to rule out trivial cases in which design Δ is never chosen by the researcher.

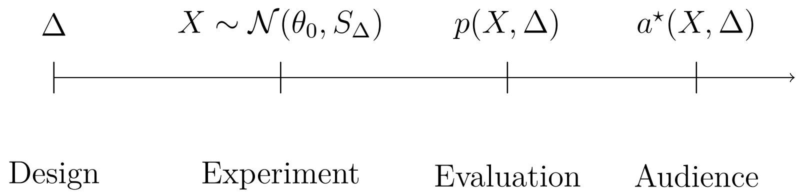

Figure 1: Illustration of the variables in the model under observable and verifiable designs. First, researchers pre-specify the population of interest. Second, they run an experiment and draw a statistic X. Finally, a social planner evaluates the experiment based on a decision rule $p(X, \Delta) \in [0, 1]$. Finally, the audience forms a posterior about the estimand of interest.

3.1 Preliminary analysis for implementing a particular design

As a first step, we characterize the constrained optimal publication when the planner must make implementing a particular design Δ individually rational for the researcher, i.e., when the planner must guarantee that the researcher always conducts the study.

Definition 1. A *constrained optimal publication rule* for a design Δ is a publication rule $p^\star_\Delta$ that minimizes $\mathbb{E}\left[\mathcal{L}_p(X(\Delta), \Delta, \theta_0)\right]$ subject to $\mathbb{E}[v_p(X(\Delta), \Delta)] \geq 0$.

Intuitively, a constrained publication rule is a publication rule that always guarantees that research is conducted. Our first result shows that the constrained optimal publication rule then takes a threshold form, where the threshold $t^\star_\Delta$ for publication depends on the prior, the mean squared error and research cost of the design Δ, and the publication cost.

Proposition 1 (Constrained optimal publication rule)**.** *If Δ is unbiased on average, then a constrained optimal publication rule for Δ is the threshold rule $p^\star_\Delta(X) = 1\left\{|X| \geq t^\star_\Delta\right\}$, where*

$$t^\star_\Delta = \min\left\{\frac{S^2_\Delta + \eta^2}{\eta^2}\sqrt{c_p}, \left|\Phi^{-1}(C_\Delta/2)\right|\sqrt{S^2_\Delta + \eta^2}\right\}.$$

Here, we write Φ for the cumulative distribution function of a standard normal.

The proof is in Appendix B.1. To understand the intuition behind Proposition 1, first suppose that the research cost is $C_\Delta = 0$, so there is no individual rationality constraint for the researcher. Then, as in Frankel and Kasy (2022), as the planner's publication cost c_p is nonzero, the planner will publish results that move the audience's optimal action enough to justify the attention/publication costs c_p: i.e., results $|X| \geq \gamma^\star_\Delta$, where[4]

$$\gamma^\star_\Delta = \frac{S^2_\Delta + \eta^2}{\eta^2}\sqrt{c_p}.$$

When this cutoff rule guarantees an *ex ante* publication probability of at least C_Δ, the individual rationality constraint does not bind. In this case, we say the design is cheap.

Definition 2. Let Δ be a design that is unbiased on average. Design Δ is *cheap* if $C_\Delta < \mathbb{P}\left(|X(\Delta)| \geq \gamma^\star_\Delta\right)$ and *expensive* otherwise.

Note that higher publication costs c_p, and lower prior variances η^2, both raise the threshold $\gamma^\star_\Delta$ and hence make designs more likely to be expensive. Whether a design is cheap depends on how informative it is (relative to publication or attention costs).

For example, if we take an experiment with unit variance (e.g., $X(\Delta)$ denotes the standard t-test), and further approximates $1/\eta^2 \approx 0$ as for a diffuse prior, $P(|X(\Delta)| \geq \gamma^\star_\Delta) = 5\%$ for standard t-tests with 5% size control. In this case, cheap experiments corresponds to $C_\Delta < 5\%$, where C_Δ captures costs to benefit ratio of publications.

For expensive designs, the cutoff rule from Frankel and Kasy (2022) does not provide a large enough *ex ante* publication probability to entice the researcher to conduct research in the first place. Hence, the planner needs to commit to publishing more results in order to satisfy the researcher's individual rationality constraint. It is best for the planner to publish results that move the audience's action the most, even if these results do not move the audience's action enough to justify the attention cost c_p. Hence, the planner sets a cutoff that ensures an *ex ante* publication chance of C_Δ—i.e., a cutoff of

$$\left|\Phi^{-1}(C_\Delta/2)\right| \sqrt{S^2_\Delta + \eta^2}.$$

This second cutoff is below $\gamma^\star_\Delta$ for (and only for) expensive designs, and is the optimal cutoff for such designs. Thus, if research costs are large enough that the implemented design is expensive, the researcher's incentives play a central role in determining the optimal publication rule, unlike in Frankel and Kasy (2022): the publication process should become less stringent as research becomes more costly.

The following corollary summarizes and formalizes the preceding discussion.

[4]Our framework directly extends to the case in which $\theta_0 \sim \mathcal{N}(\mu, \eta^2)$ for a prior mean μ. Here, μ is the audience's action in the absence of publication. In this case, the optimal constrained publication rule takes the form $|X - \mu| \geq t^\star_\Delta$ for the same threshold $t^\star_\Delta$ as in Proposition 1. Therefore, as in Frankel and Kasy (2022), one should interpret surprising results as ones that move the audience away from its default action.

Corollary 1. *(a) If* Δ *is a cheap design, then the cutoff* $t^\star_\Delta$ *for a constrained optimal publication rule is* $t^\star_\Delta = \gamma^\star_\Delta$.

(b) If Δ *is an expensive design, then* $t^\star_\Delta = |\Phi^{-1}(C_\Delta/2)|\sqrt{S^2_\Delta + \eta^2}$ *is the cutoff for a constrained optimal publication rule.*

3.2 Which designs are ever worth incentivizing

As a second step building up towards the main results of this section, we study when a design is ever worth incentivizing, relative to the outside option of no study (i.e., relative to not making it individually rational for the researcher to conduct research with design Δ).

Let $\mathcal{L}^\star_\Delta = \mathbb{E}\left[\mathcal{L}_{p^\star_\Delta}(X(\Delta), \Delta, \theta_0)\right]$ denote the optimal expected loss for the planner once implementing design Δ. (Here $p^\star_\Delta$ is a constrained optimal publication rule for Δ.) The expected loss if no research is published is the prior variance η^2. Comparing these two quantities determines whether a design is worth incentivizing in the first place.

Definition 3. A design Δ is *worthwhile* if $\mathcal{L}^\star_\Delta \leq \eta^2$.

We next characterize which designs are worthwhile. If a design is cheap, then the planner can selectively publish only results that move the audience's beliefs enough to justify the publication or attention cost. Thus, the *ex post* loss from incentivizing the design (under the constrained optimal publication rule) is always lower than η^2, and similarly its *ex ante* loss.

Proposition 2 (When are cheap designs worthwhile?)**.** *Every cheap design* Δ *is worthwhile.*

The proof is in Appendix B.2. Whereas cheap design are always worthwhile (regardless of their cost), for expensive designs, the situation is more delicate. Incentivizing the researcher to implement a design requires committing to publish results that the planner would *ex post* prefer not to publish. When attention costs c_p are large enough, the costs of publishing these marginal results outweigh the benefits of publishing surprising results. How large c_p needs to be for this to occur depends on the design's cost and variance.

To formalize this intuition, it will be convenient to express our results in terms of the difference between the posterior and prior variances conditional on publication of the results of a design Δ, which we denote by

$$\text{PostVarRed}(\Delta) := \eta^2 - \frac{S^2_\Delta \eta^2}{S^2_\Delta + \eta^2} = \frac{\eta^4}{S^2_\Delta + \eta^2}.$$

This quantity is a measure of the informativeness of a design: it represents how much learning the results of the design improves the expected utility of a Bayesian audience with L^2 loss.

Note that $\text{PostVarRed}(\Delta)$ is increasing in η^2 and decreasing in S^2_Δ.

Proposition 3 (When are expensive designs worthwhile?)**.** *Let Δ be a design that is expensive and unbiased on average.*

(a) If $\text{PostVarRed}(\Delta) \geq C_\Delta c_p + \frac{\pi}{6}\eta^2(1 - C_\Delta)^3$*, then Δ is worthwhile.*

(b) If $\text{PostVarRed}(\Delta) < C_\Delta c_p$*, then Δ is not worthwhile.*

The proof is in Appendix B.4. To understand this result, in the case of large research costs, suppose that $C_\Delta \approx 1$. A design is worthwhile only if the posterior variance reduction exceeds the product of the attention and research cost (up to a small remainder).[5] A design instead is *not* worthwhile if the posterior variance reduction relative to the cost of the design is smaller than the publication cost. That is, as C_Δ or c_p increase, the posterior variance reduction must increase proportionally for the design to be worthwhile.

Example 1. Consider a design Δ that is unbiased on average, and suppose $C_\Delta = \frac{c_v}{S_\Delta^2} + c_f$, for a variable cost c_v and fixed cost c_f. If $\eta^2 > c_p c_f$ and $S_\Delta^2 > \frac{c_p c_v}{\eta^2 - c_p c_f}$, then Δ is not worthwhile.

3.3 Choosing which design to incentivize

We next study the optimal publication rule when there is more than one possible design. Without loss of generality, we suppose that both designs are worthwhile, and that the one with a lower mean squared error has a higher research cost, so the planner faces a non-trivial problem about which design to incentivize.

Assumption 1. Researchers can choose between two designs E, O that are unbiased on average and worthwhile. The designs have mean squared errors $S_E^2 < S_O^2$ and costs $C_E > C_O$.

We think of $\Delta = E$ as a possibly expensive experiment and $\Delta = O$ as a lower-cost experiment (or, in some cases observational study). To understand this case, let us return to our example of the choice of control in a clinical trial from the introduction. Here, $\Delta = E$ represents a possibly expensive (large) medical experiment with a treatment and placebo control group, and $\Delta = O$ represents a lower-cost medical study such as a clinical trial with a smaller sample size or with a pre-specified synthetic control group. Design O has smaller cost, since under that design, researchers have to recruit fewer participants. For simplicity, we suppose that the bias of both studies E and O is unknown with mean zero; as we discuss below, we can then capture the bias as part of the mean-squared error.

Example 2 (Low-cost and costly experiment)**.** Suppose that E corresponds to an experiment with costly implementation C_E, where O is an experiment with larger variance (fewer

[5] Proposition 3 is sharp up to the term $\frac{\pi}{6}\eta^2(1 - C_\Delta)^3$, and we provide explicit expressions in the Appendix.

participants), but smaller cost of implementation. In this case, we have $\mathbb{E}[X(E)|\theta_0] = \mathbb{E}[X(O)|\theta_0] = \theta_0$, but the two experiments may have different costs (and variances). □

Example 3 (Experiment versus pre-specified synthetic control group)**.** Let $X(E) = \theta_0 + \varepsilon_E, X(O) = \theta_0 + b_O + \varepsilon_O$, where $\varepsilon_E \sim \mathcal{N}(0, S_E^2)$ denote the estimation noise from an experiment, $\varepsilon_O \sim \mathcal{N}(0, \sigma_O^2)$ denotes the idiosyncratic noise from an observational study and $b_O|\varepsilon_O \sim \mathcal{N}(0, \sigma_B^2)$ denotes the observational study's random effect which capture unobserved bias drawn from a fixed (Gaussian) distribution. For example, Rhys Bernard et al. (2024) empirically investigates the distribution of b_O through a meta-analysis, where σ_B^2 captures the impact of the random bias in statistics' distribution.[6] We then have that $X(O) \sim \mathcal{N}(\theta_0, S_O^2)$, where the mean-squared error $S_O^2 = \sigma_O^2 + \sigma_B^2$ includes both sampling uncertainty σ_O^2 and irreducible error σ_B^2 arising the variance of the bias. □

We next use Proposition 1 to study the optimal choice between the two designs. Because Δ is observable by the planner (and verifiable as part of the publication process), the planner can incentivize their preferred design by setting

$$p^\star(X, \Delta) = 1\{\Delta = \Delta^{\text{planner}}\} p^\star_\Delta(X) \qquad \text{where} \qquad \Delta^{\text{planner}} \in \underset{\Delta \in \{E,O\}}{\arg\min}\, \mathcal{L}^\star_\Delta. \tag{4}$$

For instance, the planner may only accept experiments with a minimum level of precision. It is immediate that $p^\star(X, \Delta)$ minimizes the planner's expected loss. We therefore study the optimal design choice by comparing the minimized loss of the social planner when implementing the experiment versus implementing the observational study. More generally, we can use similar logic to compare the effectiveness of any two designs.

Definition 4. Design Δ is *planner-preferred* to design Δ' if $\mathcal{L}^\star_\Delta < \mathcal{L}^\star_{\Delta'}$.

It is immediate that if a design is planner-preferred to a worthwhile design Δ', then Δ is worthwhile. In particular, Proposition 2 implies that if a design Δ is planner-preferred to a cheap, unbiased design Δ', then Δ is worthwhile.

If the more precise experiment E is cheap, then its higher research cost is irrelevant to the planner. Therefore, the experiment is planner-preferred to O.

Proposition 4. *Under Assumption 1, if E is cheap, then E is planner-preferred to O.*

The proof is in Appendix B.5. This result implies that it suffices to compare the mean-squared error of two cheap studies to select the planner preferred.

[6] Whenever b_O has non zero expectation we can think of X recentered by its expectation, that can be learned through meta-analyses (see, e.g., Rhys Bernard et al. (2024)).

When the experiment E is expensive, the situation is more delicate. Implementing E requires committing to publish more results, which may be costly for a planner. When the publication or attention costs c_p are large enough, the costs of publishing more results outweighs the benefits of a more precise design.

Proposition 5. *Under Assumption 1, if E, O are expensive, then there exists a threshold $c_p^\star(E, O, \eta) > 0$ such that E is planner-preferred to O if and only if $c_p < c_p^\star(E, O, \eta)$, where*

$$c_p^\star(E, O, \eta) = \frac{\mathrm{PostVarRed}(E) - \mathrm{PostVarRed}(O)}{C_E - C_O} - \eta^2 \frac{O\big((1 - C_E)^3\big) - O\big((1 - C_O)^3\big)}{C_E - C_O}.$$

The proof is in Appendix B.6. Proposition 5 shows that, assuming both E and O have high research costs ($(1 - C_O)^3 \approx 0$), it suffices to compare

$$\mathrm{PostVarRed}(E) - \mathrm{PostVarRed}(O) \geq (C_E - C_O)c_p \tag{5}$$

to choose between a more precise and more expensive experiment E or a less precise design O. That is, we must compare the difference in the posterior variance reduction to the difference in costs, adjusted by the attention cost c_p. A larger c_p shifts the preference of the planner towards *less costly* designs. The following theorem formalizes these intuitions.

Theorem 1 (Comparing two designs)**.** *Under Assumption 1, suppose that E, O is expensive.*

(a) If $\mathrm{PostVarRed}(E) - \mathrm{PostVarRed}(O) \geq \big(1 - \frac{C_O}{C_E}\big)c_p$, then E is planner-preferred to O.

(b) If $\mathrm{PostVarRed}(E) - \mathrm{PostVarRed}(O) \leq \big(C_E - \frac{1+2C_O}{3}\big)c_p$, then O is planner-preferred to E.

If instead O is cheap, then (a) and (b) hold with $P(|X(O)| \geq \gamma_O^\star)$ in lieu of C_O.

The proof is in Appendix B.7. Intuitively the comparison between two studies must depend on the posterior variance reduction of each study (which itself depends on their "quality" captured through their mean-squared error) and the *costs* of each study.

As the cost of attention c_p increases, the planner's preference shifts from a more accurate design to a less accurate design with a smaller cost. This is because, for costly studies, the planner must internalize not only the effect of the mean-squared error on the audience's loss function, but also the research cost associated with the study. Whenever she can publish fewer results (c_p is larger), more costly experiments impose more stringent constraints on the publication rules, making those undesiderable for the planner. This is suggestive that medical studies with synthetic control groups may be preferred over placebo groups when the cost of the placebo group is sufficiently large.

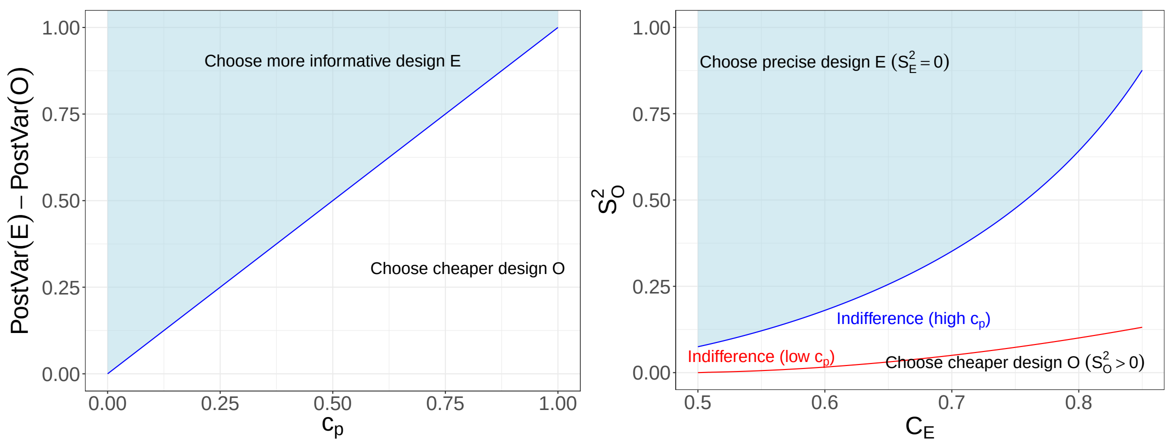

Figure 2: Comparison of two designs. The left hand side makes the comparison using the approximation in Equation (5) with E being twice as expensive than O. On the y-axis it reports the difference in posterior variance reduction and on the x-axis reports the cost of attention. As the cost of attention increases, the more costly experiment must lead to larger posterior variance reduction for this to be published. The right-hand-side reports different choices using the exact expression for $\mathcal{L}^\star_\Delta$. The right-hand-side figure reports on the x-axis the cost of an expensive experiment with no variance ($S_E^2 = 0$), and on the y-axis the rescaled squared error S_O^2/η^2 of a cheap design (e.g., placebo study or cheaper experiment) using the exact expression. The red line denotes the frontier of values for cost of publication $c_p/\eta^2 = 0.5$, and the blue lines for the higher cost of publication $c_p/\eta^2 = 1$. The region above the blue line denotes the set of values under which an experiment is preferred for a large cost of publication, and the area above the red line denotes the low cost of publication. The figure shows that a cheaper design may be preferred over a more expensive experiment, even if its mean squared error is larger, whenever the experiment E is sufficiently costly. As the cost of publication increases, the planner prefers more the cheaper design due to the effect on the supply of research.

Theorem 1 provides bounds (instead of the exact expression) to enhance interpretability. However, explicit expressions can be obtained directly from our calculations in the Appendix. Using such expressions, Figure 2 reports the indifference curves between two experiments with different mean-squared errors and costs. Figure 2 shows how the planner's preference shifts towards cheaper (and noisier) designs even when the experiment has *no* error.

Remark 1 (Explicit recommendation for practice)**.** In Appendix A we use data from distributions of t-statistics both in economics (Brodeur et al., 2023) and medical studies (Head et al., 2015) to calibrate the model under a verifiable design (in addition to our main application with manipulation in Section 5). The two datasets return similar results. Specifically, once we calibrate c_p to mimic significance threshold at 5% levels, we find that we should prefer an expensive study E over a less costly study O whenever

$$\frac{1}{S_E^2} - \frac{1}{S_O^2} \geq 1.34(1 - C_O/C_E). \tag{6}$$

Note that (6) only depends on the relative research costs $\frac{C_O}{C_E}$, so is independent of the absolute benefits of publication to the researcher (which we have normalized to 1 for simplicity).

4 Publication rules under non-verifiable designs

In this section, we turn to settings where researchers may manipulate the findings through, e.g., p-hacking. To this end, we investigate optimal publication rules when researchers choose the research design Δ using information of the statistics drawn in the experiment. Here, the design Δ (and its corresponding bias β_Δ) are unknown to the social planner but not to the researcher. Specifically, we consider the following model.

Assumption 2. Consider a class of designs $\Delta \in \mathbb{R}$, with $\beta_\Delta = \Delta$ known to the researcher but not verifiable for the planner, and $S_\Delta^2 = S_E^2$. Writing $X(\Delta) = \theta_0 + \beta_\Delta + \varepsilon$ (where $\varepsilon|\theta_0 \sim \mathcal{N}(0, S_E^2)$), the researcher observes $\theta_0 + \varepsilon$ and chooses Δ to maximizes her realized payoff $v_p(X(\Delta), \Delta)$. Let $C_\Delta = c_d|\beta_\Delta| + C_0$, where $c_d < \infty$ and $C_0 < 1$. The social planner chooses $p(X)$ as a function of X only (and not Δ).

We assume that the variance of the residual noise ε is independent of Δ and equal to S_E^2. Our results will be valid for any S_E^2 (including $S_E^2 \approx 0$ as in large studies). We interpret this assumption as stating that standard errors are verifiable as part of the publication process; we therefore focus on manipulation that introduces unverifiable bias in the reported results.

Figure 3 illustrates the model: the researcher chooses deterministically the bias of the reported statistic. They, however, pay a cost C_Δ increasing in the bias. We think of the researcher's action as a stylized description of data manipulation or selective reporting. In particular, researchers, after looking at the data, can change their specification by, e.g., changing the covariates in a regression, winsorizing the data in particular ways, or making other design choices functions of the statistics. Our stylized description captures these features by defining X as the sum of $\theta_0 + \varepsilon$ plus a bias arising from manipulation. The component $c_d|\beta_\Delta|$ of the cost C_Δ captures reputational or computational costs associated with the manipulation, assumed to be increasing and linear in the magnitude $|\beta_\Delta|$ of the bias. The researcher observes $\theta_0 + \varepsilon$, and hence, she maximizes her *realized* utility conditional on the observed statistics when choosing Δ.

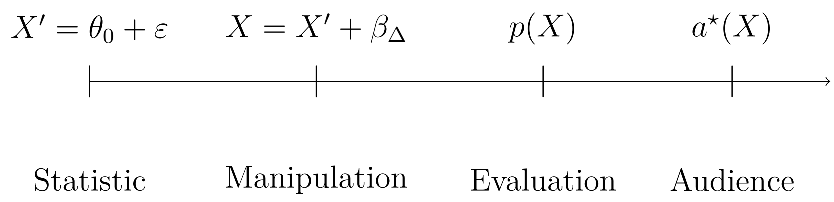

Figure 3: Illustration of the variables in the model. First, researchers observe the vector of statistics. They then manipulate the design by introducing a bias into the statistics and maximize their private utility. The social planner does not observe the bias, and evaluates the study based on a publication function $p(X)$ that only depends on the statistics X.

As we discuss in Section 2, the audience (but not the planner) is unaware of the possibility that published findings may have some data manipulation.

4.1 Optimal publication rule under manipulation

The planner cannot observe or verify β_Δ, knows S_E^2, and minimizes expected loss over (θ_0, ε) taking into account the researcher's (endogeneous) incentives to manipulate their results. That is, we define an optimal publication rule as

$$p^\star \in \arg\max_{p \in \mathcal{P}} \mathbb{E}_{X,\theta_0} \left[\mathcal{L}_p(X(\Delta_p^\star), \Delta_p^\star, \theta_0)\right] \quad \text{with} \quad \Delta_p^\star \in \arg\max_{\Delta} v_p(X(\Delta), \Delta), \tag{7}$$

where $\mathcal{P}$ is the set of all Borel measurable functions $p(X, \Delta)$ constant in Δ (i.e., do not depend on the design), which we write without loss as $p(X)$ (note that p may implicitly depend on S_E^2). Here, X satisfies Assumption 2 and $\Delta_p^\star$ denotes the optimal researcher's response. The publication rule only depends on X but not on Δ, as this is chosen privately by the researcher. Before introducing our next theorem we introduce the following definition.

Definition 5. A *linearly smoothed cutoff rule* with cutoff $X^\star$ and slope m is defined by

$$p_{X^\star,m}(X) = \begin{cases} 0 & \text{if } |X| \leq X^\star - \frac{1}{m} \\ 1 - m(X^\star - |X|) & \text{if } X^\star - \frac{1}{m} < |X| < X^\star \\ 1 & \text{if } |X| \geq X^\star \end{cases}.$$

A linearly smoothed cutoff rule considered here is a deterministic publication rule above and below thresholds $(X^\star - \frac{1}{m}, X^\star)$ for given m and it randomizes the publication chances between these two values, with publication probability increasing in the value of the reported statistic $|X|$. The special case of slope $m = \infty$ and threshold $X^\star = \gamma_E^\star = \frac{S_E^2+\eta^2}{\eta^2}\sqrt{c_p}$. corresponds to a publication rule in Frankel and Kasy (2022), i.e., a publication rule for cheap experiments without manipulation.

We next characterize the optimal publication probability in contexts with manipulation.

Theorem 2 (Optimal publication rule under non-verifiable designs)**.** *Under Assumption 2:*

(a) There exists a cutoff $X^\star \in \left(\gamma_E^\star, \gamma_E^\star + \frac{1-C_0}{c_d}\right)$ *such that the linearly smoothed cutoff rule* $p_{X^\star,c_d}$ *is optimal.*

(b) For each optimal publication rule p*, there exists* $X^\star \in \left(\gamma_E^\star, \gamma_E^\star + \frac{1-C_0}{c_d}\right)$ *such that* $p^\star(X) = p_{X^\star,c_d}(X)$ *(resp.* $p^\star(X) \leq C_0$*) for almost all* $X \geq 0$ *with* $p_{X^\star,c_d}(X) > C_0$ *(resp.* $p_{X^\star,c_d}(X) \leq C_0$*).*

The proof is in Appendix C.1. Theorem 2 characterizes the optimal publication rule under manipulation. The rule is a linearly smoothed cutoff rule that: (i) does not publish results below a certain cutoff $X^\star - \frac{1}{c_d}$; (ii) randomizes whenever X is above the $X^\star - \frac{1}{c_d}$ but below $X^\star$; and (iii) publishes with probability 1 for results $X \geq X^\star$. Although the optimal cutoff $X^\star$ does not admit a simple closed-form expression, it can be computed numerically as we show in Figure 4.

4.2 Interpretation and implications for published findings

To gain further intuition about the optimal publication rule in Theorem 2, it is instructive to compare it with the optimal publication rule for a *cheap* experiment in Proposition 1. For ease of exposition, we abstract from fixed research costs and take $C_0 = 0$. Consider first a scenario in which the social planner ignores manipulation. Then we would observe *bunching* around the cutoff to publish $\gamma_E^\star$, as researchers with $\theta_0 + \varepsilon \in (\gamma_E^\star - \frac{1}{c_d}, \gamma_E^\star)$ would introduce a bias to publish. Researchers with $\theta_0 + \varepsilon < \gamma_E^\star - \frac{1}{c_d}$ would find it nonprofitable to introduce any bias (as the cost would not compensate the benefits) and therefore would not publish. The first line of Table 1 and the first two panels of Figure 4 summarize this discussion.

Next, suppose that the planner introduces randomization in the publication rule whenever $X \in (\gamma_E^\star - \frac{1}{c_d}, \gamma_E^\star)$ as in a linearly smoothed cutoff rule with cutoff $\gamma_E^\star$. This randomization makes the researcher indifferent between manipulating and not manipulating the data, at the cost of publishing some unsurprising results. However, this publication rule is still sub-optimal as too many unsurprising results are published in the randomization regime. The second line of Table 1 summarizes this discussion.

Publication rule	Testable observation	Published findings	Manipulation
Optimal cutoff rule ignoring manipulation	Large bunching	Only surprising findings are published	Large
Add randomization below cutoff	No bunching	Many unsurprising findings are published	None
Optimal (Increase cutoff + randomize below)	Some bunching	Some unsurprising findings are published	Some

Table 1: Comparisons between three different publication rules. The first row corresponds to naive publication rule that ignores manipulation and chooses the optimal cutoff assuming no manipulation. In this case we observe large bunching at the cutoff. The second row corresponds to a publication rules that removes manipulation altogether by introducing randomization below such a cutoff. This rule is sub-optimal since too many unsurprising findings get published. The last row corresponds to the optimal publication rule.

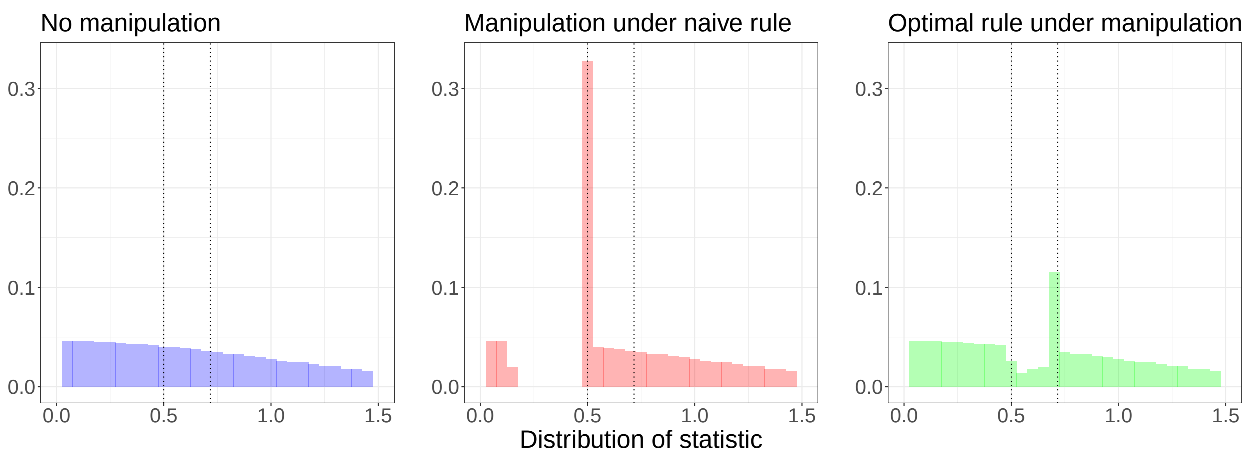

Figure 4: Distribution of $|t| = \frac{|X+\beta_{\Delta^\star}|}{\sqrt{S_E^2+\eta^2}}$ under different publication regimes. In context where no manipulation is possible (first panel)), we see no bunching on the distribution of $|t|$. When instead we consider a naive publication rule that ignores manipulation, we observe a large bunching around the critical cutoff $\gamma_E^\star$ (second panel). With the optimal publication regime under manipulation (third panel), the critical threshold for which findings are published with probability one is higher, whereas in the neighborhood of $X^\star$ are published with positive probability. We observe some bunching at $X^\star$ although this is less extreme than the one we observe under the naive rule. We consider as parameters $c_d = 2, \eta^2 = 2, S_E^2 = 0$ and $c_p = 0.5$. The first dotted line corresponds to naive optimal cutoff $\gamma_E^\star$ and the second dotted line corresponds to the cutoff $X^\star$ for the optimal rule.

The last step is to increase the threshold $X^\star$ to lower the loss from publishing unsurprising results. However, a consequence is that some surprising results are not published.

Proposition 6 (Some unsurprising results are published, and some surprising results are not)**.** *Under Assumption 2, for any optimal publication rule* $p^\star$*:*

(a) for some values of $|\theta_0 + \varepsilon| < \gamma_E^\star$*, we have* $p^\star(X) > C_0$ *but* $\beta_{\Delta^\star_{p^\star}} = 0$*; and*

(b) *for some values of* $|\theta_0 + \varepsilon| > \gamma_E^\star$, *we have* $p^\star(X) < 1$.

The proof is in Appendix C.2. Proposition 6 shows that some unsurprising results that would have not been published without manipulation do get published under the optimal publication rule with manipulation. This feature is not due to manipulation of these results, but rather to deter manipulation. It is in stark contrast with the case without manipulation, as in Frankel and Kasy (2022). Moreover, some surprising results that would have been published without manipulation do not get published under the optimal publication rule with manipulation. This is again to deter manipulation.

Given that results are only guaranteed publication if they cross a higher threshold than $\gamma_E^\star$, some manipulation can be beneficial to the planner to increase the publication rate of surprising findings. Therefore, in the planner's preferred equilibrium, researchers below the cutoff $X^\star$ and above $\gamma_E^\star$ engage in some manipulation. This form of manipulation is distinct from the manipulation that researchers would engage in under a (non-smoothed) cutoff rule, which involves results that should not be published and therefore hurts the planner.

Proposition 7 (Manipulation and bunching in equilibrium)**.** *Under Assumption 2, consider any optimal publication rule, and let* $X^\star$ *be as in Theorem 2(b).*

(a) *for almost all* $|\theta_0 + \varepsilon| \in (\gamma_E^\star, X^\star)$, *we have* $|\beta_{\Delta^\star_{p^\star}}| > 0$; *and*

(b) *there exists* $\zeta > 0$ *such that* $|\theta_0 + \varepsilon + \beta_{\Delta^\star_{p^\star}}| = X^\star$ *for almost all* $|\theta_0 + \varepsilon| \in (X^\star - \zeta, X^\star)$.

The proof is in Appendix C.2. The third line of Table 1 and the third panel of Figure 4 summarize the implications of the proposition.

4.3 Role of fixed experimental costs

As the last exercise, we study the role of fixed costs in the presence of manipulation.

Proposition 8 (Implementation costs)**.** *Under the model in Assumption 2, for* $C_0 \geq 1 - c_d\gamma_E^\star$, *the set of optimal cutoffs* $X^\star$ *in Theorem 2 is decreasing in* C_0 *in the strong set order.*

The proof is in Appendix C.3. Proposition 8 shows that for sufficiently costly studies, the social planner *lowers* the cutoff as the fixed cost increases. This result is suggestive that less surprising results may be published more when the study is sufficiently costly. It differs from what we found in Theorem 1 in the absence of manipulation, where the planner may force the researcher to run cheap over expensive studies. Fixed costs imply a lower chance of manipulation, motivating increasing the chance of publication for such studies.

5 An empirical application to medical studies: optimal rules under p-hacking

In this section we study the implications of our model under non-verifiable designs focusing on applications for medical and pharmaceutical studies. Head et al. (2015) collected p-values via text-mining the PubMed database across several disciplines. We focus here on the corresponding (about $800,000$) t-statistics classified as medical and pharmaceutical studies. Define $t := X(\Delta) - \beta_\Delta$ as the true t-statistics (absent p-hacking) and $X(\Delta)$ as the observed t-statistics. For our calibrations, we will assume for simplicity that observed t statistics are (mostly) representative of results submitted to medical journals (i.e., assuming that any study will eventually published in *some*, possibly low rank, medical journal).

We consider the same model as in the main text, where $\theta \sim \mathcal{N}(0, \eta^2), X(\Delta)|\theta \sim \mathcal{N}(\theta, 1)$, with $X(\Delta)$ denoting the estimated effect divided by its standard error, and $a_p^\star(X, \Delta) = \frac{\eta^2}{\eta^2+1}X$. Here, η^2 captures any prior uncertainty. Our model depends on parameters (c_d, c_p, η^2), assuming fixed costs are sunk (and therefore $C_0 = 0$).[7] We use the data from Head et al. (2015) to calibrate the model.

Calibration of c_d We calibrate c_d as follows: we first note that for many journals, standard ($1, 5, 10\%$ significance) levels may (approximately) correspond to publication rules of the form $p^\star(t) = 1\{|t| \geq q_{\alpha/2}\}$ where $q_{\alpha/2}$ is the critical Gaussian quantile at given significance level α. In this case, we would observe a discontinuity of the distribution of X in equilibrium. Such discontinuity occurs over the interval $(q_{\alpha/2} - \frac{1}{c_d}, q_{\alpha/2})$. We can therefore use the data to learn the lower end of the discontinuity $q_{\alpha/2} - \frac{1}{c_d}$ and therefore learn c_d. For instance, we observe a discontinuity between ~ 2 and 2.3 (the 1%-α test critical threshold), suggesting a value of $c_d \approx \frac{1}{2.3-2} = 3.33$ (see e.g., Figure 5, right-hand side).[8]

Calibration of η^2 Manipulation may change the distribution of X around relevant cutoffs (e.g., $1.96, 2.33$). To construct a measure robust to such distributional changes, we take the 95^{th} quantile of the distribution of X, denoted as $\bar{q}_{95}$. Here, $\bar{q}_{95} = 3.71$ far above the critical value 2.33. Using properties of the Gaussian distribution, $\eta^2 \approx (\bar{q}_{95}/2)^2 - 1 = 2.458$.

Calibration of c_p We choose c_p to mimic critical values we find in practice, since in this model $\sqrt{c_p \frac{1+\eta^2}{\eta^2}}$ defines the critical threshold for standard publication rules. Therefore we

[7] Therefore we should interpret our analysis valid conditional on researchers having conducted the experiment. Our analysis can be extended to non-sunk fixed costs, by using multiple cutoffs in the data (e.g., at 1.96 and 2.33) to learn two parameters C_0, c_d instead of only c_d only.

[8] A similar value can be obtained if we consider a discontinuity around 10%-significance level.

choose $\sqrt{c_p \frac{1+\eta^2}{\eta^2}} \in \{2.33, 1.96\}$, corresponding to the critical value for 1% and 5%-α tests.

Main takeaways In Figure 5 and Table 2 we collect the distribution of t-statistics X observed in equilibrium under the optimal publication rule alongside the empirical distribution of t statistics from data in Head et al. (2015). The panels at the top (bottom) reports optimal rules for critical values $\sqrt{c_p \frac{1+\eta^2}{\eta^2}} = 2.33$ (1.96), with the dotted line corresponding to the critical value $\sqrt{c_p \frac{1+\eta^2}{\eta^2}}$. The left-hand side of Figure 5 collects the distribution of t-statistics reported in papers (without conditioning on those being necessarily published).

The first takeaway is that the optimal publication rule increases the significant cutoff to $X^\star = 2.63$ for 1%-α tests instead of 2.33 and to $X^\star = 2.26$ for 5%-α tests. That is, "significance level" becomes more stringent, as we interpret significance level as the threshold above which results will be accepted almost surely. Just below the cutoff, researchers will introduce some discontinuity due to some p-hacking under the optimal rule. Such discontinuity is also observed (at different points) in the empirical data, although under the optimal rule is less pronounced.

The second key takeaway is that the share of published results is significantly *larger* (in equilibrium) under the optimal publication rule compared to the threshold rule of the form $1\{|X| \geq \sqrt{c_p \frac{1+\eta^2}{\eta^2}}\}$, see Table 2.[9] For instance, under the naive rule $1\{|X| \geq 2.33\}$ (≥ 1.96) only 4.2% (9.7%) are published (taking into account manipulation). On the other hand, more than 20% of studies are published under the optimal rule.

Taken together, these results suggest that publication rules in medical journals should increase the share of published findings, but increase the "significance" cutoff at which findings are published with probability one. This will reduce (although not eliminate) manipulation, taking into account the researchers' best response to changes in the publication rule.

In practice, we recommend editors asking for a "significance score"

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

The score equals one if effects are above $X^\star$, and is proportional to $|X|$ otherwise. Under our calibration $c_d = 3.33$, and $X^\star = 2.26$ for $\sqrt{c_p \frac{1+\eta^2}{\eta^2}} = 1.96$.

[9] Here we compute the probability of publishing under the naive rule $1\{|X| \geq \sqrt{c_p \frac{1+\eta^2}{\eta^2}}\}$ as $\mathbb{E}[1\{|X| \geq \sqrt{c_p \frac{1+\eta^2}{\eta^2}}\}] = \mathbb{E}[1\{|t| \geq \sqrt{c_p \frac{1+\eta^2}{\eta^2}} - 1/c_d\}]$ since for $X \in (\sqrt{c_p \frac{1+\eta^2}{\eta^2}} - 1/c_d, \sqrt{c_p \frac{1+\eta^2}{\eta^2}})$ the researcher would engage in manipulation under the naive rule $1\{|X| \geq \sqrt{c_p \frac{1+\eta^2}{\eta^2}}\}$.

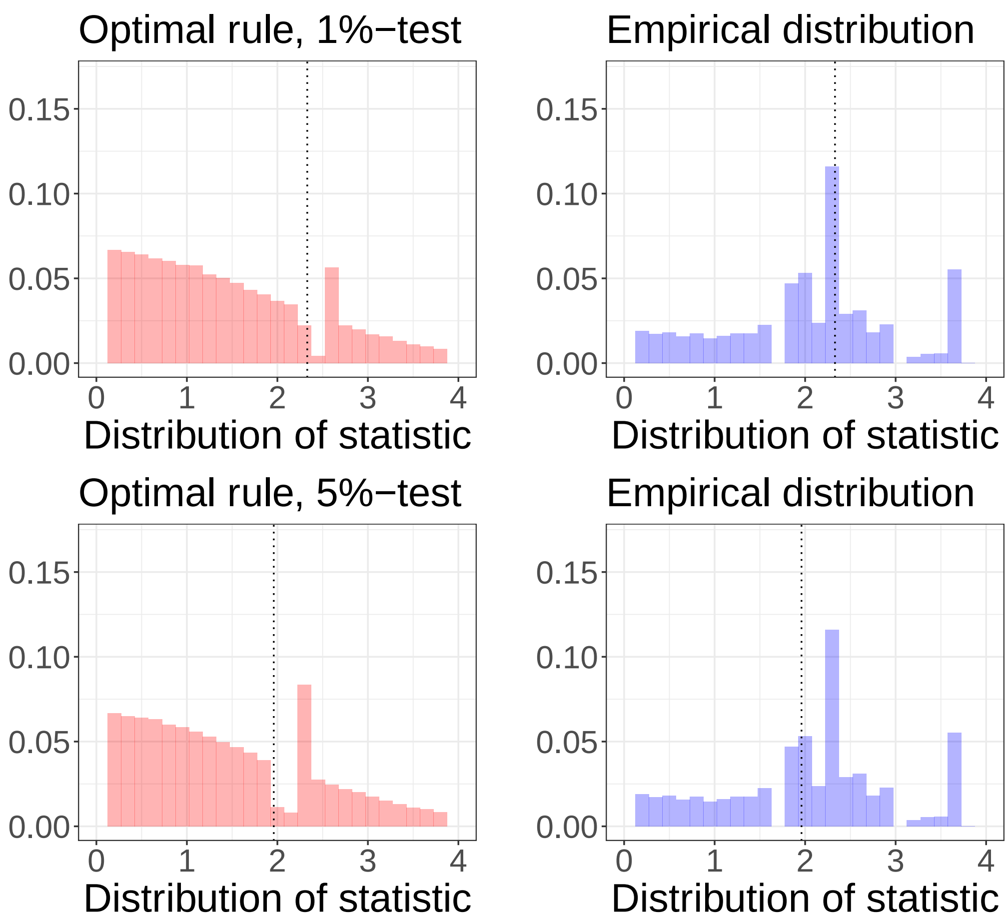

Figure 5: Simulated statistics under optimal publication rules (left-hand side) and empirical distribution of t-statistics in medical and pharmaceutical studies using data from Head et al. (2015) (right-hand side). The left-hand side panels report the distribution of the statistics $|X|$ in equilibrium, under the calibrated optimal publication rule with $c_d = 3.33, \eta^2 = 2.458, \sqrt{c_p}\frac{1+\eta^2}{\eta^2} \in \{2.33, 1.96\}$. The dotted line corresponds to the $\sqrt{c_p}\frac{1+\eta^2}{\eta^2}$ critical value.

Table 2: Each row corresponds to different choices of $\sqrt{c_p}$. The second column (titled $X^\star$) denotes the threshold at which results are published under the optimal rule with probability one. The third column reports the share of results published under the optimal publication rule (in equilibrium). The last column reports the share of published studies in equilibrium under the naive rule $1\{|X| \geq \sqrt{c_p}\}$.

$c_d = 3.33, \eta^2 = 2.458$	$\|X\| \geq X^\star$ (Significance cutoff)	$\mathbb{E}[p^\star(X)]$ (with manipulation)	$\mathbb{E}[1\{\|X\| \geq \sqrt{c_p}\frac{1+\eta^2}{\eta^2}\}]$ (with manipulation)
$\sqrt{c_p}\frac{1+\eta^2}{\eta^2} = 2.33$ (1%-test)	2.631	0.207	0.042
$\sqrt{c_p}\frac{1+\eta^2}{\eta^2} = 1.96$ (5%-test)	2.261	0.287	0.097

6 Implications for pre-specified experiments

In this section we study the implications of our results for incentivizing more observational studies with possible data-manipulation or costly experiments with no data manipulation—e.g., when enforcing pre-analysis plans. We combine the analyses with and without manipulation from the previous two sections.

Definition 6 (Pre-analysis vs possible manipulation)**.** Consider the following two possible families of designs, all of which are unbiased on average and have variance S_E^2.

(A) Experiment with pre-analysis plan: Researchers cannot manipulate their findings (as in Section 3), truthfully report $X = \theta_0 + \varepsilon$, and pay a research cost C_E. The social planner chooses a publication rule $p_E^\star$ as in Definition 1 and incurs a loss $\mathcal{L}_E^\star$.

(B) Possible manipulation (observational studies): Assumption 2 holds, so there is a family of designs $\Delta \in \mathbb{R}$. The manipulation cost is $c_d < \infty$, and researchers can manipulate their findings after observing $\theta_0 + \varepsilon$. The fixed research cost is $C_0 = 0$. The social planner chooses a publication rule $p^\star$ as in Equation (7) and incurs corresponding expected loss $\mathcal{L}_M^\star = \mathbb{E}_{X,\theta_0}\left[\mathcal{L}_{p^\star}(X(\Delta_{p^\star}^\star), \Delta_{p^\star}^\star, \theta_0)\right]$ under the planner's preferred equilibrium.

Scenario (A) states that the researcher cannot manipulate her findings but pays a fixed cost C_E, interpreted as the cost that guarantees no bias in the study. This setting may correspond to the cost of conducting an experiment (including time for grant approval, research assistants, etc.).[10] Scenario (B) allows for manipulation of the findings, with the researcher not required to write a pre-analysis plan. In this case, we assume for simplicity no fixed costs $C_E = 0$, but possible (reputational) costs associated with manipulation. We think of (B) as scenarios where an experiment or observational study is already available to the researcher (and therefore, its cost is sunk).

To simplify notation, we introduce the following quantitative measure of how much loss the expensiveness of a design entails; this measures is useful for our subsequent analysis.

Definition 7 (Incentive costs)**.** Denote $\mathrm{IC}(E) = \mathcal{L}_E^\star - \mathcal{L}_{E'}^\star$ where E' is a cheap design with $C_{E'} = 0$ and variance $S_{E'}^2 = S_E^2$. Namely, $\mathrm{IC}(E)$ denotes the difference in the planner's loss (under the planner's optimal action) for a design E and the same design (with same variance) E' but no implementation costs $C_{E'} = 0$.

It follows that $\mathrm{IC}(E)$ measures the costs on the social planner of private research costs C_E, which are a form of *incentive costs*. Whenever E is a cheap design, $\mathrm{IC}(E) = 0$; otherwise,

[10] For simplicity, we consider here experiments that are worthwhile.

IC(E) > 0. Also, note that IC(E) can be readily computed from the expressions in the Appendix.

Proposition 9 (Value of pre-analysis plan). *Under (A) and (B) in Definition 6:*

(a) If E is cheap (i.e., IC(E) $= 0$*), then E is planner-preferred to M;*

(b) If IC(E) $> \frac{1+2c_d S_E}{c_d^2}$*, then M is planner preferred to E.*

In Proposition 9, we show that an experiment (with a pre-analysis plan) is preferred over an observational study with possible manipulation (with no pre-analysis plan) if the cost of the experiment is sufficiently low.

When the cost of the experiment is high (and therefore IC(E) is large), the trade-off depends on (i) its cost C_E and (ii) the cost of data manipulation c_d. Clearly, if $c_d = 0$, then an experiment with a pre-analysis plan may be preferred. Intuitively, with a low cost of manipulation, the planner must publish with positive probability a larger set of unsurprising results, at a possibly large publication cost.

Suppose instead $c_d < \infty$. Then, an observational study with possible manipulation is preferred for sufficiently high researcher's cost. This is despite the cost of the experiment being private and paid by the researcher and not by the planner. This is because, a sufficiently large of the experiment C_E increases the loss of the social planner who must publish results that would otherwise not publish with low experimentation costs.

Different from (and complementary to) models where either no experimentation or no reputational costs occur (e.g., Kasy and Spiess, 2023; Spiess, 2025), these results illustrate *trade-offs* in the choice of unbiased vs. biased studies. A larger experiment cost may decrease the supply of research, making the planner prefer possible manipulation.

7 Conclusion

This paper studies how researcher's incentives shape the optimal design of the scientific process. Ignoring the researcher's incentives, it is optimal to publish the most surprising results (Frankel and Kasy, 2022). When researcher's incentives matter, we show that optimal publication rules depend on private costs of research and incentives for research manipulation.

As a first exercise, we show that, in the absence of manipulation, the planner prefers experiments or observational studies with larger mean-squared error) over sufficiently costly experiments. In medical studies, a pre-specified synthetic control group obtained from medical records (Jahanshahi et al., 2021; Food and Drug Administration, 2023) may be preferred over a sufficiently expensive placebo control group, despite lack of randomization.

With manipulation, we show that it is optimal to (i) publish some unsurprising results and (ii) knowlingly allow for manipulation (biased studies) at the margin. Observationally, the optimal policy would reduce the bunching of the findings (e.g., t-tests) around the publication cutoff. However, the optimal policy does not completely remove bunching. Even when the planner can mandate a signal to enforce no manipulation, this may not be her preferred policy when the signal entails a larger research cost.

Our model disentangles the contribution of design choice and data manipulation to optimal publication decisions. Future research should study more complex communication strategies. For example, in contexts with pre-analysis plans, the planner may allow multiple signals to decrease the burden on the authors or allow for the publication of non-prespecified findings. Similarly, researchers may report multiple findings in their study. As we point to in our results in Section 6, studying more complex action spaces can shed light on alternative mechanisms relevant to scientific communication.

This paper opens new questions at the intersection of econometrics and mechanism design. Future research should study the implications of some of these conclusions on empirical research. Our framework relies on parameters that capture costs and benefits for the researcher. In the absence of manipulation, such parameters can be learned using cost data for medical trials (e.g., Tetenov, 2016) and field experiments (e.g., Viviano et al., 2024). With manipulation, and reputational costs associated with it, such parameters may be learned using information from meta-analyses through the distribution of the submitted findings as we show in Section 5. Future research should focus on sharping our understanding of the identification of such parameters.

References

Abadie, A. (2020). Statistical nonsignificance in empirical economics. *American Economic Review: Insights 2*(2), 193–208.

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

Ana, J. (2004). The role of a general medical journal. *BMJ 328*(7439), 591.

Andrews, I. and M. Kasy (2019). Identification of and correction for publication bias. *American Economic Review 109*(8), 2766–2794.

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

Banerjee, A., S. Chassang, and E. Snowberg (2017). Chapter 4 – Decision theoretic approaches to experiment design and external validity. In A. V. Banerjee and E. Duflo (Eds.), *Handbook of Field Experiments*, Volume 1 of *Handbook of Economic Field Experiments*, pp. 141–174. North-Holland.

Banerjee, A., E. Duflo, A. Finkelstein, L. F. Katz, B. A. Olken, and A. Sautmann (2020). In praise of moderation: Suggestions for the scope and use of pre-analysis plans for RCTs in economics. Working paper, National Bureau of Economic Research.

Banerjee, A. V., S. Chassang, S. Montero, and E. Snowberg (2020). A theory of experimenters: Robustness, randomization, and balance. *American Economic Review 110*(4), 1206–1230.

Bates, S., M. I. Jordan, M. Sklar, and J. A. Soloff (2022). Principal-agent hypothesis testing. arXiv:2205.06812.

Bates, S., M. I. Jordan, M. Sklar, and J. A. Soloff (2023). Incentive-theoretic Bayesian inference for collaborative science. arXiv:2307.03748.

Brodeur, A., S. Carrell, D. Figlio, and L. Lusher (2023). Unpacking p-hacking and publication bias. *American economic review 113*(11), 2974–3002.

Brodeur, A., N. Cook, and A. Heyes (2020). Methods matter: P-hacking and publication bias in causal analysis in economics. *American Economic Review 110*(11), 3634–3660.

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.

DeMaria, A. N. (2022). The role of a medical journal. *Structural Heart 6*(4), 100079.

Di Tillio, A., M. Ottaviani, and P. N. Sørensen (2021). Strategic sample selection. *Econometrica 89*(2), 911–953.

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

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

Gechter, M. and R. Meager (2022). Combining experimental and observational studies in meta-analysis: A debiasing approach. Working paper, Pennylvania State University and London School of Economics.

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

Grabowski, H., J. Vernon, and J. A. DiMasi (2002). Returns on research and development for 1990s new drug introductions. *Pharmacoeconomics 20*, 11–29.

Head, M. L., L. Holman, R. Lanfear, A. T. Kahn, and M. D. Jennions (2015). The extent and consequences of p-hacking in science. *PLoS biology 13*(3), e1002106.

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

Hirano, K. and J. R. Porter (2009). Asymptotics for statistical treatment rules. *Econometrica 77*(5), 1683–1701.

Jahanshahi, M., K. Gregg, G. Davis, A. Ndu, V. Miller, J. Vockley, C. Ollivier, T. Franolic, and S. Sakai (2021). The use of external controls in FDA regulatory decision making. *Therapeutic Innovation & Regulatory Science 55*(5), 1019–1035.

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

Kitagawa, T. and A. Tetenov (2018). Who should be treated? Empirical welfare maximization methods for treatment choice. *Econometrica 86*(2), 591–616.

Kitagawa, T. and P. Vu (2023). Optimal publication rules for evidence-based policy. Working paper, Brown University.

Libgober, J. (2022). False positives and transparency. *American Economic Journal: Microeconomics 14*(2), 478–505.

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

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 (Forthcoming, 2024). Critical values robust to p-hacking. *Review of Economics and Statistics.*

Miguel, E. (2021). Evidence on research transparency in economics. *Journal of Economic Perspectives 35*(3), 193–214.

Milgrom, P. and I. Segal (2002). Envelope theorems for arbitrary choice sets. *Econometrica 70*(2), 583–601.

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies 38*(2), 175–208.

Myerson, R. B. (1981). Optimal auction design. *Mathematics of Operations Research 6*(1), 58–73.

Olken, B. A. (2015). Promises and perils of pre-analysis plans. *Journal of Economic Perspectives 29*(3), 61–80.

Popat, S., S. V. Liu, N. Scheuer, G. G. Hsu, A. Lockhart, S. V. Ramagopalan, F. Griesinger, and V. Subbiah (2022). Addressing challenges with real-world synthetic control arms to demonstrate the comparative effectiveness of pralsetinib in non-small cell lung cancer. *Nature Communications 13*(1), 3500.

Raices Cruz, I., M. C. Troffaes, J. Lindström, and U. Sahlin (2022). A robust Bayesian bias-adjusted random effects model for consideration of uncertainty about bias terms in evidence synthesis. *Statistics in Medicine 41*(17), 3365–3379.

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.

Riveros, C., A. Dechartres, E. Perrodeau, R. Haneef, I. Boutron, and P. Ravaud (2013). Timing and completeness of trial results posted at clinicaltrials. gov and published in journals. *PLoS Medicine 10*(12), e1001566.

Rosenzweig, M. R. and C. Udry (2020). External validity in a stochastic world: Evidence from low-income countries. *Review of Economic Studies 87*(1), 343–381.

Savage, L. J. (1951). The theory of statistical decision. *Journal of the American Statistical Association 46*(253), 55–67.

Shinohara, K., A. Tajika, H. Imai, N. Takeshima, Y. Hayasaka, and T. A. Furukawa (2015). Protocol registration and selective outcome reporting in recent psychiatry trials: New antidepressants and cognitive behavioural therapies. *Acta Psychiatrica Scandinavica 132*(6), 489–498.

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

Tetenov, A. (2012). Statistical treatment choice based on asymmetric minimax regret criteria. *Journal of Econometrics 166*(1), 157–165.

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

Thompson, S. G. and J. A. Barber (2000). How should cost data in pragmatic randomised trials be analysed? *BMJ 320*(7243), 1197–1200.

Thorlund, K., L. Dron, J. J. Park, and E. J. Mills (2020). Synthetic and external controls in clinical trials–a primer for researchers. *Clinical Epidemiology 12*, 457–467.

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

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

Williams, C. (2021). Preregistration and incentives. Working paper, Durham University.

Wong, H.-H., A. Jessup, A. Sertkaya, A. Birkenbach, A. Berlind, and J. Eyraud (2014). Examination of clinical trial costs and barriers for drug development final. *Office of the Assistant Secretary for Planning and Evaluation, US Department of Health & Human Services*, 1–92.

Yin, X., P. S. Mishra-Kalyan, R. Sridhara, M. D. Stewart, E. A. Stuart, and R. C. Davi (2022). Exploring the potential of external control arms created from patient level data: a case study in non-small cell lung cancer. *Journal of Biopharmaceutical Statistics 32*(1), 204–218.

Yoder, N. (2022). Designing incentives for heterogeneous researchers. *Journal of Political Economy 130*(8), 2018–2054.

A Calibration for a model with verifiable design

In this section we provide a simple calibration for our model with verifiable design. Taking stock, our model and conclusions in Section 3 depend on key parameters (c_p, η^2) in addition to (possibly observable) parameters (S^2_Δ, C^2_Δ). In this section, we illustrate how the parameters of the model can be calibrated in practice to provide practical recommendations. Define $t_\Delta := \frac{X(\Delta)}{S_\Delta}$ as the t-statistic commonly reported in academic papers.

Calibration of η^2/S^2 The first step is to calibrate η^2/S^2 (as a function of S^2). To do so, we can take the distribution of t statistics across submitted journal articles, so that $\mathbb{V}(t|\theta) = 1$ and $\mathbb{V}(t) = 1 + \eta^2/S^2$. Here η^2 captures uncertainty arising from heterogeneity across papers (e.g., papers having different estimands) as well as prior uncertainty about the parameters. Under a Gaussian approximation for $t \sim \mathcal{N}(0, 1 + \eta^2/S^2)$, we take the 95^{th} quantile of the distribution of t across papers, denoted as $\bar{q}_{95}$. The focus on the 95^{th} quantiles may avoid some manipulation that occurs within the support of the t-statistic as we further describe and study in Section 4 and Section 5. For instance, using data about *submitted* papers in economics collected by Brodeur et al. (2023), we estimate $\bar{q}_{95} = 3.39$, so that $\eta^2/S^2 = 1.87$. Using t-statistics from medical articles collected by Head et al. (2015) as described in Section 5, we find similar $\bar{q}_{95} = 3.71$.

Calibration of c_p It follows that for t-statistics t_Δ (having unit variance conditional on θ by construction), the critical threshold $\gamma_\Delta = \sqrt{c_p \frac{1+\eta^2}{\eta^2}}$. Motivated by empirical practice, we set $\gamma_\Delta = 1.96$ as the standard threshold for t-tests with 5% Type I error. Using our tabulation with $\eta^2 = 1.87$ for $S^2 = 1$, $c_p = 1.63$ for tests with 5% Type I error.

Implications for design choice Under such calibration, we can write

$$\text{PostVarRed}(\Delta) = \frac{1.21}{S^2_\Delta}.$$

so that under Proposition 3 an expensive design is not worthwhile if $\frac{1}{S^2_\Delta} \leq 1.34 C_\Delta$. Similarly, under Theorem 1, we prefer an expensive study E over a less costly and precise study O, if

$$\frac{1}{S^2_E} - \frac{1}{S^2_O} \geq 1.34(1 - C_O/C_E).$$

B Proofs omitted from Section 3

Let ϕ denote the probability density function of a standard normal.

B.1 Proof of Proposition 1

The proof uses the following lemma, which provides a simple decomposition of the social planner's loss function conditional on the realized statistic X that we also use in other proofs.

Lemma 1 (Loss function)**.** *Let* Δ *be a design that is unbiased on average.*

$$\mathbb{E}\Big[\mathcal{L}_p(X(\Delta),\Delta,\theta_0)\Big|X(\Delta)\Big] = \Big[c_p - \frac{\eta^4}{(S_\Delta^2+\eta^2)^2}X(\Delta)^2\Big]p(X,\Delta) + \mathbb{E}[\theta_0^2|X(\Delta)]$$

Proof. The lemma follows directly from the fact that $\mathbb{E}[\theta_0|X(\Delta)] = \frac{\eta^2}{S_\Delta^2+\eta^2}X(\Delta)$. □

The individual rationality constraint $\mathbb{E}[v_p(X(\Delta),\Delta)] \geq 0$ can be written equivalently as $\mathbb{E}[p(X(\Delta))] \geq C_\Delta$. Note that the objective and constraints are linear in p. Using Lemma 1 and placing a Lagrange multiplier $\lambda \geq 0$ on this constraint, the Lagrangian of the planner's problem can be written as

$$\mathbb{E}\left[\left(c_p - \lambda - \Big(\frac{\eta^2}{S_E^2+\eta^2}X(\Delta)\Big)^2\right)p\big(X(\Delta)\big)\right] + \lambda C_E + \eta^2$$

We can solve and obtain $p_\lambda(X) = 1\{|X| \geq \frac{S_E^2+\eta^2}{\eta^2}\sqrt{c_p-\lambda}\}$ as the minimizer of the Lagrangian. As $X(\Delta) \sim \mathcal{N}(0, S_E^2+\eta^2))$, the individual rationality constraint can be written as

$$2\Phi\left(-\sqrt{c_p-\lambda}\frac{\sqrt{S_E^2+\eta^2}}{\eta^2}\right) \geq C_E \iff \lambda \geq c_p - \left(\frac{\eta^2}{\sqrt{S_E+\eta^2}}\big|\Phi^{-1}\big(\tfrac{C_E}{2}\big)\big|\right)^2.$$

By complementary slackness, it follows that

$$\lambda = \max\left\{0, c_p - \left(\frac{\eta^2}{\sqrt{S_E+\eta^2}}\big|\Phi^{-1}\big(\tfrac{C_E}{2}\big)\big|\right)^2\right\}.$$

As we then have

$$\frac{S_E^2+\eta^2}{\eta^2}\sqrt{c_p-\lambda} = \min\left\{\frac{S_\Delta^2+\eta^2}{\eta^2}\sqrt{c_p}, \big|\Phi^{-1}(C_\Delta/2)\big|\sqrt{S_\Delta^2+\eta^2}\right\} = t_\Delta^\star,$$

we have that $p_\lambda(X) = p_\Delta^\star(X)$.

B.2 Proof of Proposition 2

Since Δ is cheap, Corollary 1 implies $t^\star_\Delta = \frac{S^2_\Delta+\eta^2}{\eta^2}$. Proposition 1 and Lemma 1 then imply

$$\mathcal{L}^\star_\Delta = \mathbb{E}\left[\left(c_p - \frac{\eta^4}{(S^2_\Delta+\eta^2)^2}X(\Delta)^2\right)1\left\{|X(\Delta)| \geq \frac{S^2_\Delta+\eta^2}{\eta^2}\right\}\right] + \eta^2 < \eta^2.$$

B.3 Preliminary Steps for Analysis of General Designs

Our analysis of general designs relies on the following lemma, which provides an exact characterization of the planner's loss under a threshold rule.

Lemma 2. *If Δ is a design that is unbiased on average, then for any publication rule of the form $p(x) = 1\{|x| \geq t\}$, we have*

$$\mathbb{E}[\mathcal{L}_p(X(\Delta),\Delta,\theta_0)] = \eta^2 + \mathbb{P}(|X(\Delta)| \geq t)c_p - \text{PostVarRed}(\Delta)\Upsilon(\mathbb{P}(|X(\Delta)| \geq t),$$

where we write $\Upsilon(t) = 2\Phi^{-1}\big(1-\frac{t}{2}\big)\phi\big(\Phi^{-1}\big(1-\frac{t}{2}\big)\big) + t$.

Proof. Bt Lemma 1, we have

$$\mathbb{E}\left[\mathcal{L}_p(X(\Delta),\Delta,\theta_0)|X(\Delta)\right] = \mathbb{E}[\theta_0^2|X] + c_p 1\big\{|X| \geq t\big\} - \left(\frac{\eta^2 X(\Delta)}{S^2_\Delta+\eta^2}\right)^2 1\big\{|X| \geq t\big\}.$$

Note that

$$\mathbb{E}\left[\left(\frac{\eta^2 X(\Delta)}{S^2_\Delta+\eta^2}\right)^2 1\big\{|X(\Delta)| \geq t\big\}\right] = \underbrace{\mathbb{E}\left[\left(\frac{\eta^2 X(\Delta)}{S^2_\Delta+\eta^2}\right)^2 \,\middle|\, |X(\Delta)| \geq t\right]}_{(I)} \mathbb{P}\left(|X(\Delta)| \geq t\right).$$

As $X(\Delta) \sim \mathcal{N}(0,\eta^2+S^2_\Delta)$, it follows from the symmetry of the Gaussian distribution and properties of truncated Gaussian distributions that

$$(I) = \mathbb{E}\left[\left(\frac{\eta^2 X(\Delta)}{S^2_\Delta+\eta^2}\right)^2 \,\middle|\, X(\Delta) \geq t\right] = \frac{\eta^4}{S^2_\Delta+\eta^2}\left[\frac{1}{\mathbb{P}\left(X(\Delta) \geq t\right)}\frac{t^\star_\Delta}{\sqrt{S^2_\Delta+\eta^2}}\phi\left(\frac{t}{\sqrt{S^2_\Delta+\eta^2}}\right) + 1\right].$$

Taking expectations and using $\mathbb{P}(|X(\Delta)| \geq t) = 2\mathbb{P}(X(\Delta) \geq t)$, it follows that

$$\begin{aligned}
&\mathbb{E}\left[\mathcal{L}_p(X(\Delta),\Delta,\theta_0)|X(\Delta)\right] \\
&\quad = \eta^2 + c_p\mathbb{P}\left(|X(\Delta)| \geq t\right) - \frac{2\eta^4}{S^2_\Delta+\eta^2}\frac{t^\star_\Delta}{\sqrt{S^2_\Delta+\eta^2}}\phi\left(\frac{t}{\sqrt{S^2_\Delta+\eta^2}}\right) - \frac{\eta^4}{S^2_\Delta+\eta^2}\mathbb{P}\left(|X(\Delta)| \geq t\right) \\
&\quad = \eta^2 + c_p\mathbb{P}\left(|X(\Delta)| \geq t\right) - \text{PostVarRed}(\Delta)\Upsilon\left(\mathbb{P}\left(|X(\Delta)| \geq t\right)\right),
\end{aligned}$$

where the last equality holds as $X(\Delta) \sim \mathcal{N}(0, S_E^2 + \eta^2)$. $\square$

Using Lemma 2, we can characterize the planner's optimal loss.

Lemma 3. *If* Δ *is unbiased on average, then letting* Υ *be as in Lemma 2, we have*

$$\mathcal{L}_\Delta^\star = \begin{cases} \eta^2 + C_\Delta c_p - \text{PostVarRed}(\Delta)\Upsilon(C_\Delta) & \text{if } \Delta \text{ is expensive} \\ \eta^2 + \mathbb{P}(|X(\Delta)| \geq \gamma_\Delta^\star) c_p - \text{PostVarRed}(\Delta)\Upsilon(\mathbb{P}(|X(\Delta)| \geq \gamma_\Delta^\star)) & \text{if } \Delta \text{ is cheap} \end{cases}.$$

Moreover, $\mathcal{L}_\Delta^\star$ *is strictly decreasing in* $\text{PostVarRed}(\Delta)$ *and non-decreasing in* C_Δ*.*

Proof. By Corollary 1, we have $\mathbb{P}(|X(\Delta)| \geq t_\Delta^\star) = \mathbb{P}(|X(\Delta)| \geq \gamma_\Delta^\star)$ if Δ is cheap and $\mathbb{P}(|X(\Delta)| \geq t_\Delta^\star) = C_\Delta$ if Δ is expensive. Hence, the first part of the lemma follows from Lemma 2 and Proposition 1.

Lemma 2 also implies that for all $t \geq 0$, the loss $\mathbb{E}[\mathcal{L}_{1\{|x| \geq t\}}(X(\Delta), \Delta, \theta_0)]$ is strictly decreasing in $\text{PostVarRed}(\Delta)$ and non-decreasing in C_Δ. Proposition 1 also implies that

$$\mathcal{L}_\Delta^\star = \min_{t \geq 0} \left\{ \mathbb{E}[\mathcal{L}_{1\{|x| \geq t\}}(X(\Delta), \Delta, \theta_0)] \right\},$$

and the second part of the lemma follows. $\square$

We next provide a few properties of the function Υ that will be used in the proofs.

Lemma 4. *Let* $\Upsilon(t)$ *be defined as in Lemma 3.*

(i) $\Upsilon(t)$ *is non-decreasing on* $(0, 1)$ *with derivative* $\Upsilon'(t) = \left[\Phi^{-1}(1 - t/2)\right]^2$.

(ii) $1 \geq \Upsilon(t) \geq 1 - \frac{\pi}{6}(1 - t)^3$ *for all* $0 < t < 1$.

(iii) $\Upsilon(t) \geq 1 - \frac{1}{3}(1 - t)\Phi^{-1}\big(1 - \frac{t}{2}\big)^2$ *for all* $0 < t < 1$.

Proof. Define $z(t) = \Phi^{-1}\big(1 - \frac{t}{2}\big)$, so that $\Upsilon(t) = 2z(t)\phi\big(z(t)\big) + t$. By the chain rule, we have

$$\Upsilon'(t) = 2z'(t)\phi\big(z(t)\big) + 2z(t)\phi'\big(z(t)\big)z'(t) + 1.$$

Recalling that $\phi'(x) = -x\phi(x)$, we have

$$\Upsilon'(t) = 2z'(t)\phi\big(z(t)\big) + 2z(t)^2\phi\big(z(t)\big)z'(t) + 1 = 2z'(t)\phi\big(z(t)\big)\left[1 - z(t)^2\right] + 1.$$

The chain rule also implies that

$$z'(t) = -\frac{1}{2\phi(z(t))}, \tag{8}$$

and it follows that

$$\Upsilon'(t) = -\left[1 - z(t)^2\right] + 1 = z(t)^2. \tag{9}$$

As $z(t)^2 \geq 0$, this proves Part (i). As $\Upsilon(1) = 1$, it follows that $\Upsilon(t) \leq 1$ for all $0 < t < 1$.

To bound $\Upsilon(t)$ from below, note that as $\phi(x) \leq \frac{1}{\sqrt{2\pi}}$, we have $\Phi(x) \leq \frac{1}{2} + \frac{x}{\sqrt{2\pi}}$ for $x \geq 0$. Taking $x = \sqrt{\frac{\pi}{2}}(1-t)$, it follows that for $0 < t < 1$, we have

$$\Phi\Big(\sqrt{\frac{\pi}{2}}(1-t)\Big) \leq 1 - \frac{t}{2}.$$

Applying Φ^{-1} to both sides yields that (for $0 < t < 1$)

$$\sqrt{\frac{\pi}{2}}(1-t) \leq \Phi^{-1}\Big(1 - \frac{t}{2}\Big) = z(t)$$

Using (9) then yields that for $0 < t < 1$, we have

$$\Upsilon'(t) = z(t)^2 \leq \frac{\pi}{2}(1-t)^2 = \frac{d}{dt}\Big(1 - \frac{\pi}{6}(1-t)^3\Big).$$

Since $\Upsilon(1) = 1$, we have $\Upsilon(t) \geq 1 - \frac{\pi}{6}(1-t)^3$ for $0 < t < 1$, completing the proof of Part (ii).

To prove Part (iii), note that as ϕ is decreasing on $\mathbb{R}_{\geq 0}$, we have that $\Phi(x) \leq \frac{1}{2} + x\phi(x)$ for $x \geq 0$. Taking $x = z(t)$, it follows that for $0 < t < 1$, we have

$$\frac{1-t}{2} = \Phi\big(z(t)\big) - \frac{1}{2} \leq z(t)\phi\big(z(t)\big).$$

Using (8) then implies that $(1-t)z'(t) \geq z(t)$ for $0 < t < 1$. Using (9) then yields that for $0 < t < 1$, we have

$$\Upsilon'(t) = z(t)^2 \leq \frac{1}{3}z(t)^2 + \frac{2}{3}(1-t)z(t)z'(t) = \frac{d}{dt}\Big(1 - \frac{1}{3}(1-t)z(t)^2\Big).$$

Since $\Upsilon(1) = 1$, Part (iii) follows by the definition of $z(t)$. □

B.4 Proof of Proposition 3

Lemma 3 actually implies the following sharper result.

Lemma 5. *Let Δ be expensive and unbiased on average. Then Δ is worthwhile if and only if*

$$\Upsilon(C_\Delta)\,\mathrm{PostVarRed}(\Delta) \geq C_\Delta c_p.$$

Proof. This is immediate from Lemma 3. □

To derive Proposition 3 from Lemma 5, note that Lemma 4(ii) implies that

$$\begin{aligned}\text{PostVarRed}(\Delta) \geq \Upsilon(C_\Delta)\,\text{PostVarRed}(\Delta) &\geq \text{PostVarRed}(\Delta)\Big(1 - \frac{\pi}{6}(1-C_\Delta)^3\Big)\\ &\geq \text{PostVarRed}(\Delta) - \frac{\pi}{6}\eta^2(1-C_\Delta)^3,\end{aligned}$$

where the last equality holds as $\text{PostVarRed}(\Delta) \geq \eta^2$.

B.5 Proof of Proposition 4

Consider a design E' that is unbiased on average with $S_{E'} = S_E$ and $C_{E'} = 0$. As E' is also cheap, Lemma 3 implies $\mathcal{L}^\star_{E'} = \mathcal{L}^\star_E$. Since $S^2_E \leq S^2_O$ (by Assumption 1), we have $\text{PostVarRed}(E') = \text{PostVarRed}(E) > \text{PostVarRed}(O)$. Hence, Lemma 3 implies $\mathcal{L}^\star_{E'} \leq \mathcal{L}^\star_O$. It follows that $\mathcal{L}^\star_E = \mathcal{L}^\star_{E'} \leq \mathcal{L}^\star_O$.

B.6 Proof of Proposition 5

We first derive an approximation of the planner's optimal loss with large research costs.

Lemma 6. *If Δ is an expensive design that is unbiased on average, then we have*

$$\mathcal{L}^\star_\Delta = \eta^2 + C_\Delta c_p - \text{PostVarRed}(\Delta) + O(\eta^2(1-C_\Delta)^3).$$

Proof. The lemma follows from Lemma 3 and Lemma 4(ii) as $\text{PostVarRed}(\Delta) \leq \eta^2$. □

Lemma 6 implies that

$$\mathcal{L}^\star_E - \mathcal{L}^\star_O = (C_E - C_O)c_p - \big(\text{PostVarRed}(E) - \text{PostVarRed}(O)\big) + \big(O(\eta^2(1-C_E)^3) - O(\eta^2(1-C_O)^3)\big).$$

The proposition follows by simple rearrangement.

B.7 Proof of Theorem 1

We first consider the case in which O is expensive, and apply that case to deduce the result for the case in which O is cheap. When O is expensive, by Lemma 3, we can write

$$\mathcal{L}^\star_E - \mathcal{L}^\star_O = C_E c_p - \text{PostVarRed}(E)\Upsilon(C_E) - C_O c_p + \text{PostVarRed}(O)\Upsilon(C_O). \tag{10}$$

Proof of (a) By Lemma 4(i), since $C_E \geq C_O$ (by Assumption 1), it follows from (10) that

$$\mathcal{L}^\star_E - \mathcal{L}^\star_O \leq (C_E - C_O)c_p - \big(\text{PostVarRed}(E) - \text{PostVarRed}(O)\big)\Upsilon(C_E).$$

Note that as $\frac{\pi}{6} < 1$ and $C_E \in [0, 1]$, Lemma 4(ii) implies that

$$\Upsilon(C_E) \geq 1 - \frac{\pi}{6}(1 - C_E)^3 \geq 1 - (1 - C_E) = C_E.$$

Since PostVarRed$(E) \geq$ PostVarRed(O) (by Assumption 1), it follows that

$$\mathcal{L}_E^\star - \mathcal{L}_O^\star \leq (C_E - C_O)c_p - \Big(\text{PostVarRed}(E) - \text{PostVarRed}(O)\Big)C_E \leq 0.$$

Proof of (b) By Parts (ii) and (iii) of Lemma 4, it follows from (10) that

$$\begin{aligned}\mathcal{L}_E^\star - \mathcal{L}_O^\star \geq (C_E - C_O)c_p &- \text{PostVarRed}(E) \\ &+ \text{PostVarRed}(O) - \frac{1}{3}\,\text{PostVarRed}(O)(1 - C_O)\Phi^{-1}\Big(1 - \frac{C_O}{2}\Big)^2.\end{aligned}$$

As O is expensive and $X(O) \sim \mathcal{N}(0, \eta^2 + S_O^2)$, we have

$$C_0 \geq \mathbb{P}(|X(O)| \geq \gamma_O^\star) = 2\bigg(1 - \Phi\Big(\sqrt{\frac{c_p}{\text{PostVarRed}(O)}}\Big)\bigg).$$

Rearranging terms and applying Φ^{-1} yields that

$$\Phi^{-1}\Big(1 - \frac{C_O}{2}\Big)^2 \leq \frac{c_p}{\text{PostVarRed}(O)}$$

Hence, we have that

$$\begin{aligned}\mathcal{L}_E^\star - \mathcal{L}_O^\star &\geq (C_E - C_O)c_p - \text{PostVarRed}(E) + \text{PostVarRed}(O) - \frac{(1 - C_O)c_p}{3} \\ &= \Big(C_E - \frac{2C_O + 1}{3}\Big)c_p - \Big(\text{PostVarRed}(E) - \text{PostVarRed}(O)\Big) \geq 0.\end{aligned}$$

Proof for the case in which O is cheap Consider a design O' that is unbiased on average with $S_{O'} = S_O$ and $C_{O'} = \mathbb{P}(|X(O)| \geq \gamma_O^\star))$. By Lemma 3, we have that $\mathcal{L}_{O'}^\star = \mathcal{L}_{O'}^\star$. Moreover, O' is expensive. Hence, we can conclude the assertions for the comparison between E and O by applying (a) and (b) to compare E to the expensive design O'.

C Proof omitted from Section 4

C.1 Proof of Theorem 2

We will refer to a trivial design as a design where the researcher does not conduct the study (and incur zero utility). We let to $\gamma^\star = \frac{S_E^2+\eta^2}{\eta^2}\sqrt{c_p}$.

Let $X_0 = X(0) = \theta_0 + \varepsilon$ denote the type of the researcher. Writing $\omega = \frac{\eta^2}{\eta^2+S_E^2}$, note that $\theta_0 | X_0 \sim \mathcal{N}\left(\frac{\eta^2}{\eta^2+S_E^2}X_0, \frac{S_E^2\eta^2}{S_E^2+\eta^2}\right)$, Hence, the planner's expected loss conditional on X_0 if the researcher chooses a nontrivial design Δ that has publication probability p is

$$\left(\omega^2\beta_\Delta^2 + c_p\right)p + \omega^2 X_0^2(1-p) + \omega^2 S_E^2 = \left(\omega^2(\beta_\Delta^2 - X_0^2) + c_p\right)p + \omega^2(X_0^2 + S_E^2). \tag{11}$$

Moreover, if the researcher chooses the trivial design, then the planner's expected loss conditional on X_0 is $\omega^2(X_0^2 + S_E^2)$. For $0 \le v \le 1$, define

$$\mathcal{L}^\star(X_0, v) = \min_{p,\beta_\Delta \in [0,1]\times\mathbb{R} \big| p - c_d|\beta_\Delta| = v} \left\{\left(\omega^2(\beta_\Delta^2 - X_0^2) + c_p\right)p\right\} \tag{12}$$

denote the minimum expected loss conditional on X_0 generated by a nontrivial design and publication probability that delivers utility exactly $v - C_0$ to the researcher.

We next characterize several properties of $\mathcal{L}^\star(X_0, v)$ and the optimizer in (12), which we use in the proof of this theorem and other results from Section 4.

Lemma 7. *Under Assumption 2:*

(a) If $|X_0| \neq \gamma^\star$, then there is a unique optimizer $(\tilde{p}(X_0, v), \beta_\Delta(X_0, v))$ in (12) given by

$$\tilde{p}(X_0, v) = \begin{cases} v & \text{if } |X_0| < \gamma^\star \\ \min\left\{1, \frac{2v + \sqrt{v^2 + 3c_d^2(X_0^2 - (\gamma^\star)^2)}}{3}\right\} & \text{if } |X_0| > \gamma^\star \end{cases}$$

and $\beta_\Delta(X_0, v) = \frac{\tilde{p}(X_0,v) - v}{c_d}$.

(b) If $|X_0| > \gamma^\star$ (resp. $|X_0| = \gamma^\star$, $|X_0| < \gamma^\star$), then $\mathcal{L}^\star(X_0, v)$ is negative (resp. zero, positive) for $v > 0$, and has negative (resp., zero, positive) derivative in v over $(0, 1)$.

(c) $\mathcal{L}^\star(X_0, v)$ has negative derivative in X_0 over $(0, \gamma^\star) \cup (\gamma^\star, \infty)$.

Proof. Writing $|\beta_\Delta| = \frac{p - u - C_0}{c_d}$, note that (12) can be written equivalently as

$$\mathcal{L}^\star(X_0, v) = \min_{p \in [0,1] | p \ge v} \left\{\left[\omega^2\left(\frac{p - v}{c_d}\right)^2 - \omega^2 X_0^2 + c_p\right]p\right\}. \tag{13}$$

Noting that $\gamma^\star = \frac{1}{\omega}\sqrt{c_p}$, we divide into cases based on the value of $|X_0|$ to complete the proof.

- Case 1: $|X_0| > \gamma^\star$. In this case, we claim that for $0 < v < 1$, the quantity $\mathcal{L}^\star(X_0; v)$ is the optimum of the relaxed problem

$$\mathcal{L}^\star(X_0, v) = \min_{p\in[0,1]} \left\{ \left[\omega^2 \left(\frac{p-v}{c_d} \right)^2 - \omega^2 X_0^2 + c_p \right] p \right\}, \tag{14}$$

  and moreover that all optimizers $\tilde{p}(X_0, v)$ satisfy $\tilde{p}(X_0, v) > v$. Indeed, taking $p = v$ in (14), we can see that the right hand side is negative. It follows that in optimum in (14), we must have that $\omega^2 \left(\frac{p-v}{c_d}\right)^2 - \omega^2 X_0^2 + c_p < 0$. The first-order condition for the optimality of p then entails that $\tilde{p}(X_0, v) > v$, as desired.

  In particular, we then have that $\mathcal{L}^\star(X_0; v) < 0$. It also follows from first-order condition for the optimality of p that

$$\tilde{p}(X_0, v) = \min \left\{ 1, \frac{2v + \sqrt{v^2 + 3c_d^2(X_0^2 - (\gamma^\star)^2)}}{3} \right\}.$$

  The Envelope Theorem (Milgrom and Segal, 2002, Theorem 3) guarantees that $\mathcal{L}^\star(X_0; v)$ is partially differentiable in v and X_0, and that

$$\begin{aligned} \frac{\partial \mathcal{L}^\star(X_0, v)}{\partial v} &= \frac{2\omega^2(p^\star(v) - v)p^\star(v)}{c_d^2} < 0 \\ \frac{\partial \mathcal{L}^\star(X_0; v)}{\partial X_0} &= -2\omega^2 X_0 p^\star(v) < 0 \qquad \text{for } X_0 > 0. \end{aligned}$$

- Case 2: $|X_0| \leq \gamma^\star$. In this case, the objective in (13) is increasing in p on the interval $[v, 1]$. The optimum is therefore achieved at $\tilde{p}(X_0, v) = v$ (uniquely for $|X_0| < \gamma$), so we have that $\mathcal{L}^\star(X_0, v) = [c_p - \omega^2 X_0^2]\, v$. For $|X_0| = \gamma^\star$, this function is zero. For $|X_0| < \gamma^\star$, this function is positive for $v > 0$, has positive derivative in v, and has negative derivative in X_0 for $X_0 > 0$.

The cases exhaust all possibilities, which completes the proof. □

Let us now consider the constrained problem in which the planner must choose a publication rule that provides type $X_0 = \gamma^\star$ an indirect utility of $u^\star \in [0, 1 - C_0]$. The linearly smoothed cutoff rule $p^\star_{\gamma^\star + \frac{1 - C_0 - u^\star}{c_d}, c_d}$ lies within this class, and under the planner's preferred equilibrium, delivers expected loss conditional on X_0 of this publication rule is

$$\begin{cases} \omega^2(X_0^2 + S_E^2) & \text{if } |X_0| \leq \gamma^\star - \frac{u^\star}{c_d} \\ \mathcal{L}^\star(X_0, u^\star + c_d(X_0 - \gamma^\star) + C_0) + \omega^2(X_0^2 + S_E^2) & \text{if } \gamma^\star - \frac{u^\star}{c_d} < |X_0| < \gamma^\star + \frac{1-C_0-u^\star}{c_d} \\ \mathcal{L}^\star(X_0, 1) + \omega^2(X_0^2 + S_E^2) & \text{if } |X_0| \geq \gamma^\star + \frac{1-C_0-u^\star}{c_d} \end{cases}.$$

Lemma 8. *Within the class of publication rules that provide type $X_0 = \gamma^\star$ an indirect utility of $u^\star \in [0, 1 - C_0]$, for all types $X_0 > 0$:*

(a) the linearly smoothed cutoff rule $p_{\gamma^\star + \frac{1-C_0-u^\star}{c_d}, c_d}$ minimizes the expected loss conditional on X_0 (under the planner's preferred equilibrium), and

(b) any other publication rule within this class minimizes the expected loss conditional on X_0 must provide the same indirect utility to type X_0 as $p_{\gamma^\star + \frac{1-C_0-u^\star}{c_d}, c_d}$.

Proof. We divide into cases based on the value of X_0.

- Case 1: $X_0 > \gamma^\star$. If type X_0 obtains utility u, then type $\gamma^\star$ could obtain utility at least $u - c_d(X_0 - \gamma^\star)$ by choosing a design with the same mean as the design chosen by X_0. Hence, we must have that $u - c_d(X_0 - \gamma^\star) \leq u^\star$. Obviously, we must have that $u \leq 1 - C_0$. By Lemma 7(b) and the definition of $\mathcal{L}^\star$, the expected loss conditional on X_0 must be at least

$$\mathcal{L}^\star\big(X_0, \min\big\{u^\star + c_d(X_0 - \gamma^\star) + C_0, 1\big\}\big) + \omega^2(X_0^2 + S_E^2)$$

  with equality only if the indirect utility to X_0 is $\min\big\{u^\star + c_d(X_0 - \gamma^\star) + C_0, 1\big\}$.

- Case 2: $X_0 = \gamma^\star$. By Lemma 7(b) and the definition of $\mathcal{L}^\star$, the expected loss conditional on X_0 must be at least $\omega^2(X_0^2 + S_E^2) = \mathcal{L}^\star(X_0, u^\star + C_0) + \omega^2(X_0^2 + S_E^2)$.

- Case 3: $\gamma^\star - \frac{u^\star}{c_d} \leq X_0 < \gamma^\star$. Type X_0 could obtain utility at least $u^\star + c_d(X_0 - \gamma^\star) > 0$ by choosing a design with the same mean as the design chosen by $\gamma^\star$. By Lemma 7(b) and the definition of $\mathcal{L}^\star$, the expected loss conditional on X_0 must be at least

$$\mathcal{L}^\star(X_0, u^\star + c_d(X_0 - \gamma^\star) + C_0) + \omega^2(X_0^2 + S_E^2),$$

  with equality only if the indirect utility to type X_0 is $u^\star + c_d(X_0 - \gamma^\star)$.

- Case 4: $X_0 < \gamma^\star - \frac{u^\star}{c_d}$. By Lemma 7(b) and the definition of $\mathcal{L}^\star$, the expected loss conditional on X_0 must be at least $\omega^2(X_0^2 + S_E^2)$, with equality only if the indirect utility to type X_0 is 0.

The cases exhaust all possibilities, which completes the proof. $\square$

We are now ready to complete the proof of the theorem.

Proof of Part (a) Lemma 8(a) implies there exists a utility level $u^\star \in [0, 1 - C_0]$ for type $\gamma^\star$ such that $p_{\gamma^\star + \frac{1-C_0-u^\star}{c_d}, c_d}$ (under the planner's preferred equilibrium) is optimal. Writing $v^\star = u^\star + C_0$, the expected loss of $p_{\gamma^\star + \frac{1-v^\star}{c_d}, c_d}$ (under the planner's preferred equilibrium) is

$$\mathcal{E}(v^\star, C_0) = \mathbb{E}_{X_0}\left[\mathcal{L}^\star(X_0, \min\{v^\star + c_d(X_0 - \gamma^\star), 1\})\, 1\left\{|X_0| \geq \gamma^\star - \frac{v^\star - C_0}{c_d}\right\}\right] + \eta^2. \quad (15)$$

Differentiating under the integral sign using Lemma 7(b), we see that $\frac{\partial \mathcal{E}}{\partial v^\star}\big|_{v^\star = C_0} < 0$ and that $\frac{\partial \mathcal{E}}{\partial v^\star}\big|_{v^\star = 1} > 0$. So for $p_{\gamma^\star + \frac{1-C_0-u^\star}{c_d}, c_d}$ to be optimal, we must have $0 < u^\star < 1 - C_0$, hence

$$0 < \frac{1 - C_0 - u^\star}{c_d} < \frac{1 - C_0}{c_d}.$$

Proof of Part (b) Consider an optimal publication rule $p^\star$ that delivers utility $u^\star$ to type $X_0 = \gamma^\star$. By Lemma 7(a), the publication rule $p^\star$ (under the planner's preferred equilibrium) must deliver expected loss conditional on X_0 equal to that of $p_{\gamma^\star + \frac{1-C_0-u^\star}{c_d}, c_d}$ for almost all types X_0, and $p_{\gamma^\star + \frac{1-C_0-u^\star}{c_d}, c_d}$ must be optimal. In particular, we have $0 < u^\star < 1 - C_0$. Using these consequences of optimality, we prove the two assertions in this part separately.

- Suppose for sake of deriving a contradiction that $p^\star(X) \neq p_{\gamma^\star + \frac{1-C_0-u^\star}{c_d}, c_d}(X)$ for a positive measure of $X > 0$ with $p_{\gamma^\star + \frac{1-C_0-u^\star}{c_d}, c_d}(X) > C_0$. Then, at least one of the following must occur.
  - Case 1: $p^\star(X) \neq 1$ for a positive measure of $X > \gamma^\star + \frac{1-C_0-u^\star}{c_d}$. Then types $X_0 > \gamma^\star + \frac{1-C_0-u^\star}{c_d}$ with $p^\star(X_0) < 1$ must obtain indirect utility less than $1 - C_0$, which, by Lemma 8(b), must lead to expected loss conditional on X_0 strictly greater than that of $p_{\gamma^\star + \frac{1-C_0-u^\star}{c_d}, c_d}$—a contradiction.
  - Case 2: $p^\star(X) \neq u^\star + C_0 + c_d(X - \gamma^\star)$ for a positive measure of results $X \in \left(\gamma^\star - \frac{u^\star}{c_d}, \gamma^\star + \frac{1-C_0-u^\star}{c_d}\right)$. Letting $\tilde{p}(X_0, v)$ be as defined in Lemma 7(a), by continuity, a positive measure of types $X_0 \in \left(\gamma^\star - \frac{u^\star}{c_d}, \gamma^\star + \frac{1-C_0-u^\star}{c_d}\right)$ must satisfy

$$p^\star\left(X_0 + \frac{\tilde{p}(X, u^\star + C_0 + c_d(X_0 - \gamma^\star)) - u^\star - C_0}{c_d}\right) \neq \tilde{p}(X, u^\star + C_0 + c_d(X_0 - \gamma^\star)).$$

Hence, if such types are to obtain indrect utility $u^\star + c_d(X_0 - \gamma^\star)$, they must have publication probability different from $\tilde{p}(X, u^\star + C_0 + c_d(X_0 - \gamma^\star))$, which, by Lemma 7(a), would also lead to expected loss conditional on X_0 strictly greater than that of $p_{\gamma^\star + \frac{1-C_0-u^\star}{c_d}, c_d}$. But Lemma 8(b) implies for $X_0 \in \left(\gamma^\star - \frac{u^\star}{c_d}, \gamma^\star + \frac{1-C_0-u^\star}{c_d}\right)$, to obtain expected loss conditional on X_0 equal to that of $p_{\gamma^\star + \frac{1-C_0-u^\star}{c_d}, c_d}$, type X_0 must obtain indirect utility $u^\star + c_d(X_0 - \gamma^\star)$—a contradiction.

- Suppose for sake of deriving a contradiction that $p^\star(X) > C_0$ for a positive measure of $X > 0$ with $p_{\gamma^\star + \frac{1-C_0-u^\star}{c_d}, c_d}(X) \le C_0$. Then types $X_0 \in (0, \gamma^\star - \frac{1-u^\star}{c_d})$ with $p^\star(X_0) > C_0$ must obtain positive indirect utility, which, by Lemma 8(b), must lead to expected loss conditional on X_0 strictly greater than that of $p_{\gamma^\star + \frac{1-C_0-u^\star}{c_d}, c_d}$—a contradiction.

C.2 Proof of Propositions 6 and 7

Let $X^\star$ be as in Theorem 2(b). We use the notation of the proof of Theorem 2.

Let $X_0 \in (X^\star - \frac{1-C_0}{c_d}, X^\star)$. Under publication rule $p^\star$, type X_0 is indifferent between all designs Δ with $0 \le \beta_\Delta \le X^\star - X_0$, and strictly prefers such designs to all other designs. These designs yield utility $v - C_0$, where $v = 1 - c_d(X^\star - X_0) > C_0$. Hence, optimality implies that almost surely, the lowest loss design Δ with $0 \le \beta_\Delta \le X^\star - X_0$ will be chosen. (11) implies that the expected loss conditional on X_0 if type X_0 chooses such a design Δ is

$$\left(\omega^2(\beta_\Delta^2 - X_0^2) + c_p\right) p + \omega^2(X_0^2 + S_E^2) \quad \text{where} \quad p = 1 - c_d(X^\star - X_0 - \beta_\Delta).$$

Hence, defining $\beta_\Delta(X_0, v)$ as in Lemma 7(a), the loss-minimizing design has bias $\beta_\Delta(X_0, v)$, and the publication probability will be $p^\star(X) = \tilde{p}(X_0, v)$.

To prove Proposition 6(a), suppose that $X_0 \in (X_0^\star - \frac{1-C_0}{c_d}, \gamma^\star)$. Lemma 7(a) implies that $\tilde{p}(X_0, v) = v$ and $\beta_\Delta(X_0, v) = 0$. Hence, almost surely, we have $p^\star(X) > C_0$ and $\beta_{\Delta^\star_{p^\star}} = 0$.

To prove Proposition 6(b), suppose that $X_0 > \gamma^\star$ is such that $\tilde{p}(X_0, v) < 1$; there is a positive measure of such X_0 as it follows from Lemma 7(a) that $\tilde{p}(X_0, 1 - c_d(X^\star - X_0))$ is continuous in X_0 and satisfies $\tilde{p}(\gamma^\star, 1 - c_d(X^\star - \gamma^\star)) = 1 - c_d(X^\star - \gamma^\star) < 1$. Almost surely, we then have that $p^\star(X) = \tilde{p}(X_0, v) < 1$.

To prove Proposition 7(a), suppose that $X_0 \in (\gamma^\star, X^\star)$. If $\tilde{p}(X_0, v) < 1$, then by Lemma 7(a), we have

$$c_d\beta_\Delta(X_0, v) = \tilde{p}(X_0, v) - v = \frac{2v + \sqrt{v^2 + 3c_d^2(X_0^2 - (\gamma^\star)^2)}}{3} - v > v - v = 0.$$

On the other hand, if $\tilde{p}(X_0, v) = 1$, then then by Lemma 7(a), we have $\beta_\Delta(X_0, v) = 1 - v > 0$.

Hence, almost surely, we have that $|\beta_{\Delta^\star_{p^\star}}| > 0$. The case of $X_0 \in (-X^\star, -\gamma^\star)$ is analogous.

To prove Proposition 7(b), note that the bound $\sqrt{v^2 + x} \leq v + \frac{x}{2v}$ yields that

$$\frac{2\big(1 - c_d(X^\star - X^\star)\big) + \sqrt{\big(1 - c_d(X^\star - X^\star)\big)^2 + 3c_d^2\big((X^\star)^2 - (\gamma^\star)^2\big)}}{3} > 1.$$

Hence, by continuity, there exists $\zeta > 0$ such that for all $X_0 \in (X^\star - \zeta, X^\star)$, we have

$$\frac{2v + \sqrt{v^2 + 3c_d^2(X_0^2 - (\gamma^\star)^2)}}{3} > 1.$$

For such X_0, by Lemma 7(a), we have $\tilde{p}(X_0, v) = 1$ and $X_0 + \beta_\Delta(X_0, v) = X_0 + \frac{1-v}{c_d} = X^\star$. Hence, almost surely, we have that $p^\star(X) = 1$ and that $|X_0 + \beta_{\Delta^\star_{p^\star}}| = X^\star$. The case of $X_0 \in (-X^\star, -X^\star + \zeta)$ is analogous.

C.3 Proof of Proposition 8

Consider the function $\mathcal{E}(v^\star, C_0)$ defined in (15) in the proof of Theorem 2. Differentiating $\mathcal{E}^\star$ using the Fundamental Theorem of Calculus implies that

$$\frac{\partial \mathcal{E}}{\partial C_0} = -\frac{2}{c_d}\mathcal{L}^\star\left(\gamma^\star - \frac{v^\star - C_0}{c_d}, C_0\right)\frac{\exp\left(-\frac{\left(\gamma^\star - \frac{v^\star - C_0}{c_d}\right)^2}{2(S_E^2 + \eta^2)}\right)}{\sqrt{2\pi(S_E^2 + \eta^2)}}.$$

Lemma 7(b) implies that $\mathcal{L}^\star\left(\gamma^\star - \frac{v^\star - C_0}{c_d}, C_0\right) > 0$ for $0 < C_0 < v^* < 1$, and it then follows from Lemma 7(c) that $\frac{\partial^2 \mathcal{E}}{\partial v^\star \partial C_0} < 0$. As the set $\mathcal{D} = \{(v^\star, C_0) \mid 1 \geq v^\star \geq C_0 \geq 1 - c_d\gamma^\star\}$ is a lattice, $\mathcal{E}$ is a submodular function on $\mathcal{D}$. Topkis's Theorem implies that the optimizer set

$$\arg\min_{v^\star \in [C_0, 1]} \mathcal{E}(v^\star, C_0)$$

is increasing in C_0 in the strong set order over the interval $C_0 \in [1 - c_d\gamma^\star, 1]$. As $\mathcal{E}(v^\star, C_0)$ is the expected loss of $p_{\gamma^\star + \frac{1-v^\star}{c_d}, c_d}$ (under the planner's preferred equilibrium), the set of optimal cutoffs $X^\star$ in Theorem 2 is

$$\gamma^\star + \frac{1}{c_d} - \frac{1}{c_d}\arg\min_{v^\star \in [C_0, 1]} \mathcal{E}(v^\star, C_0),$$

which is decreasing in the strong set order over the interval $C_0 \in [1 - c_d\gamma^\star, 1]$.

D Proof of Proposition 9

We prove each claim separately.

Proof of (i) For the first claim it suffices to note that

$$\mathcal{L}_M^\star \geq \min_{\Delta,p} \mathbb{E}\Big[\mathcal{L}_p(X(\Delta),\Delta,\theta_0)\Big] \tag{16}$$

where in the right hand side we minimize both over the design Δ and the publication rule p. Note that for any publication rule p, minimizer over Δ on the right-hand side of Equation (16) sets $\beta_\Delta = 0$ since it is easy to show that $\mathbb{E}\Big[\mathcal{L}_p(X(\Delta),\Delta,\theta_0)\Big]$ is increasing in the bias β_Δ. In addition, because we have a cheap design, for $\beta_\Delta = 0$, the minimizer over the right hand side expression equals $p^\star = 1\{|X| \geq \gamma_E^\star\}$ from Proposition 1. Therefore, for a cheap design E, it follows that

$$\min_{\Delta,p} \mathbb{E}\Big[\mathcal{L}_p(X(\Delta),\Delta,\theta_0)\Big] = \mathcal{L}_E^\star$$

completing the claim for (i).

Proof of (ii) For the second claim, from Theorem 2, we can write

$$\mathcal{L}_M^\star \leq \mathbb{E}\Big[\mathcal{L}_{p'}(X(\Delta_{p'}^\star),\Delta_{p'}^\star,\theta_0)\Big], \quad p'(X) = 1\left\{|X| \geq \gamma_E^\star + \frac{1}{c_d}\right\} \tag{17}$$

since p' is a sub-optimal publication rule. Therefore, by expanding the expression on the right-hand side of Equation (17)

$$\begin{aligned}
\mathcal{L}_M^\star &\leq \mathbb{E}\left[\left(\varepsilon + \beta_{\Delta_{p'}^\star}\right)^2 p'(X) + \eta^2(1-p'(X)) + c_p p'(X)\right] \\
&= \underbrace{\mathbb{E}\left[\varepsilon^2 p'(X) + \eta^2(1-p'(X)) + c_p p'(X)\right]}_{(A)} + \underbrace{\mathbb{E}[\beta^2_{\Delta_{p'}^\star} p'(X)] + 2\mathbb{E}[\beta_{\Delta_{p'}^\star}\varepsilon p'(X)]}_{(B)}
\end{aligned}$$

It follows that under p' for all results for which $|\theta_0+\varepsilon| \in [\gamma_E^\star, \gamma_E^\star + \frac{1}{c_d}]$, the researcher chooses $|\beta_{\Delta_{p'}^\star}| = \gamma_E^\star + \frac{1}{c_d} - |\theta_0+\varepsilon|$. For all $|\theta_0+\varepsilon| < \gamma_E^\star$, the researcher chooses $|\beta_{\Delta_{p'}^\star}| = 0$ (and similarly $|\beta_{\Delta_{p'}^\star}| = 0$ if $|\theta_0+\varepsilon| > \gamma_E^\star + \frac{1}{c_d}$). Therefore, we can write in equilibrium $p'(X(\Delta_{p'}^\star)) = 1\{|\theta_0+\varepsilon| \geq \gamma_E^\star\}$. As a result, we can write as we define $p''(x) = 1\{|x| \geq \gamma_E^\star\}$

$$\begin{aligned}
(A) &= \mathbb{E}\left[\varepsilon^2 p'(X) + \eta^2(1-p'(X)) + c_p p'(X)\right] = \mathbb{E}\Big[\mathcal{L}_{p''}(\theta_0+\varepsilon, 0, \theta_0)\Big] \\
(B) &= \mathbb{E}[\beta^2_{\Delta_{p'}^\star} p''(\theta_0+\varepsilon)] + 2\mathbb{E}[\beta_{\Delta_{p'}^\star}\varepsilon p''(\theta_0+\varepsilon)].
\end{aligned}$$

That is (A) equals the loss function as if there was no manipulation, under a publication rule p''. (B) captures the manipulation component. We next study (B).

We consider each term in (B). First note that for the first term, we can write

$$\beta^2_{\Delta^\star_{p'}} \leq \frac{1}{c_d^2} \Rightarrow \mathbb{E}[\beta^2_{\Delta^\star_{p'}} p''(\theta_0+\varepsilon)] \leq \frac{1}{c_d^2}$$

since the researcher will never choose a bias larger than $1/c_d$ under $p'(x)$. For the second term, we can write

$$\mathbb{E}[\beta_{\Delta^\star_{p'}} \varepsilon p''(\theta_0+\varepsilon)] \leq \sqrt{\mathbb{E}[\beta^2_{\Delta^\star_{p'}}] S_E^2} \leq \frac{S_E}{c_d}$$

where the first inequality follows from from Cauchy-Schwarz inequality, and the second inequality by the bound on the bias in equalibrium.

Collecting the terms, we can write

$$\mathcal{L}_E^\star - \mathcal{L}_M^\star \geq \mathcal{L}_E^\star - \mathbb{E}\Big[\mathcal{L}_{p''}(\theta_0+\varepsilon, 0, \theta_0)\Big] - \frac{1+2S_E c_d}{c_d^2}.$$

Here, $\mathcal{L}_E^\star$ is the loss for an expensive design with no manipulation with cost C_E. On the other hand, $\mathbb{E}\Big[\mathcal{L}_{p''}(\theta_0+\varepsilon, 0, \theta_0)\Big] = \mathcal{L}_O^\star$ for a design O with cost $C_O = 0$ and variance $S_O^2 = S_E^2$. It follows that $\mathcal{L}_E^\star - \mathbb{E}\Big[\mathcal{L}_{p''}(\theta_0+\varepsilon, 0, \theta_0)\Big] = \mathrm{IC}(E)$, completing the proof.