Publication Design with Incentives in Mind*

Ravi Jagadeesan† Davide Viviano‡

October 22, 2024
(Preliminary, comments welcome)

Abstract

Publication decisions shape the process of scientific communication but also influence what and how research is conducted through their impact on the researcher's private incentives. We introduce a framework to study optimal publication decisions when researchers can choose (i) how the study is conducted and (ii) whether to manipulate the research findings (e.g., via selective reporting). When manipulation is not possible, but research entails substantial private costs for the researchers, some unsurprising findings must be published to compensate for such costs. As a result, it is optimal to incentivize cheaper experiments with lower accuracy as publication constraints become more binding. When manipulation occurs, it is optimal to allow for biased studies and publish some unsurprising results. Even if it is possible to mandate signals to deter manipulation (e.g., via pre-analysis plans), this is suboptimal when the signals entail high (perceived) private costs for researchers.

*We thank Isaiah Andrews, Arun Chandrasekhar, Kevin Chen, Paul Niehaus, Jesse Shapiro, Jann Spiess, Elie Tamer, 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.

1 Introduction

Publication decisions can 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 publish the most informative results (e.g. Abadie, 2020; Frankel and Kasy, 2022). It is typically considered most informative to publish surprising results.

However, implicit in the goal of publishing is an equally important task: which findings get published determines the incentives for researchers on 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 decisions about each of these aspects of a study broadly as a *research design*. As researchers choose their research designs, their private incentives may influence many aspects of it. On the other hand, "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 question of how researchers' incentives should impact the design of scientific communication.

This paper studies optimal publication decisions when the researcher chooses the research design based on private costs and benefits. We frame this question as a mechanism design problem: a social planner (principal) chooses optimal publication rules, taking into account the incentives of the researcher (agent). The social planner aims to publish results that are most informative to the public, net of the cost of (or constraints for) publishing, as in Frankel and Kasy (2022)'s single-agent model. We focus on how publication decisions here affect which research designs are chosen in the first place: given the publication rule, the researcher chooses the design that maximizes her value for publication 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). However, medical researchers may respond in how they conduct their study through the size, length, cost of the study, composition of the control group (Thorlund et al., 2020), and—in some cases—in which specific findings to report (e.g., Riveros et al., 2013; Shinohara et al., 2015).

In Section 2, we introduce a framework that allows us to study how researchers' incentives affect (i) how the study is conducted and (ii) which findings are reported (e.g., occurring with data manipulation or selective reporting). We focus on Bayes' optimal decision rules under Gaussian priors—whose variance captures the audience's uncertainty about the effect of the study—and a quadratic loss that measures accuracy in the research findings. Our main results disentangle the contributions of (i) and (ii) to the optimal publication rules.

In Section 3, we focus on (i) and abstract from data manipulation. Researchers choose between (quasi)experiments with different variances and *costs* before the experiment result is realized—as with pre-analysis plans—and truthfully report an unbiased estimate of the study's effect. Thus, the research design is common knowledge. The planner decides whether to publish the study as a function of the study's result and design, and upon publication, the public updates its beliefs. The variance captures sampling variation, as well as ex-ante variation due to a study random effect (which may arise with a study ex-post bias drawn from a known distribution, see e.g. Rhys Bernard et al., 2024). That is, the variance defines, in reduced form, the overall ex-ante mean-squared error.

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). Whether a social planner should incentivize one or the other (or combinations of these) is becoming a pressing issue for medical studies, given its relevance for cost reduction (Jahanshahi et al., 2021; Food and Drug Administration, 2023). 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).

Since the research design is common knowledge, the planner can enforce the design she prefers provided that it is individually rational for a researcher to implement the design. In the absence of this individual rationality constraint, the first best policy would force the researcher to conduct the study with the lowest mean-squared error. When the feasible designs are two experiments that are inexpensive for the researcher to conduct, this policy can be feasibly implemented by the planner; the individual rationality constraint does not bind. This aligns with the lay intuition that (with a quadratic loss) we should prefer studies with the lowest mean-squared error irrespective of the researcher's private costs.

However, when the research cost of one of the designs is above a "tipping point," the publication decision rule can substantially affect whether it is individually rational for the researcher to implement the first-best design. As a result, the planner's Bayes optimal publication implements an inexpensive experiment with non-negligible mean-squared error instead of an experiment with *no* mean-squared error but sufficiently high cost. This deviation from the first best is exacerbated when the planner faces more stringent constraints for publication: with a higher cost of publication, the planner's preference shifts towards *less* expensive experiments. Intuitively, for a costly experiment to be incentive-compatible, the planner must increase its publication chances—irrespective of whether its findings are surprising—to make researchers willing to run the experiment. Findings with little effect on the audience's belief get published at a possibly large (attention) cost since otherwise the researcher would not conduct an expensive study. This analysis suggests that medical studies with observational control groups are preferred over sufficiently expensive placebo groups, even when the two have the same mean-squared error.

We next turn in Section 4 to scenarios where the researcher, *after* observing the study's result, can report a biased statistic from such a study. The bias is chosen by the researcher and unobserved by the planner. The audience is unaware of possible manipulations of published findings. We think of this model as a stylized description of settings with potential selective reporting or data manipulation. The cost incurred by the researcher is increasing in the bias (e.g., capturing reputational costs).

Each publication rule generates different degrees of data manipulation. Due to the absence of direct transfers and the inability to promise a publication probability above one, the planner's mechanism design problem is effectively one with limited transfers. As a result, standard methods as in Mirrlees (1971) and Myerson (1981) do not apply. We characterize the solution of the mechanism design problem and show that the optimal publication rule always satisfies three properties: (a) it increases the cutoff for which findings always get published compared to settings without data manipulation; (b) just below this cutoff, it randomizes publication chances, making the researcher indifferent between manipulating the data or not; (c) in the planner's preferred equilibrium, some data manipulation *does* occur.

To gain further insight, suppose first that the planner wrongly chooses the publication rule that would be optimal without manipulation. This corresponds to publishing only if the reported statistic is above a cutoff (function of the prior distribution, variance and cost of publication). Then bunching would occur at the cutoff: researchers whose results are around the cutoff would engage in manipulation (as documented in settings with p-hacking, e.g. Elliott et al., 2022). The planner would therefore incur a loss due to possibly large bias.

As a second step, the planner publishes results below the cutoff with some positive probability to make the researcher indifferent between manipulating and not the data. Manipulation (and bunching) does not occur anymore, but some uninteresting results do get published. This is undesirable in the presence of publication costs or constraints.

As the last step, the planner increases the cutoff just enough to compensate for the loss of publishing uninteresting results. Because the cutoff is more stringent, in the planner's preferred equilibrium, some researchers just below the new cutoff engage in data manipulation to increase their publication chances. The loss from publishing some biased studies is second order relative to not publishing interesting findings.

In summary, publication decisions require slightly bigger effects to be certainly published, publish *some* uninteresting results, and tolerate some bias at the margin. This is in stark contrast with scenarios without manipulation, where only interesting results get published.

As a final exercise in Section 5, we combine these two models and ask whether the planner should mandate the researcher to send a costly signal prohibiting data manipulation (e.g., enforcing a pre-analysis plan) or allow for data manipulation without such a signal. The (perceived) cost of the pre-analysis plan is private and only incurred by the researcher. These scenarios correspond to our first and second models, respectively. In the absence of any researcher's cost associated with the signal (or study), the planner always prefers enforcing unbiased estimates (see for example Spiess, 2024; Kasy and Spiess, 2023). Our core focus here instead is on the cost associated with the design: taking into account the researcher's incentives, mandating a costly signal may decrease the supply of research.

For example, when the pre-analysis plan entails a substantial burden on the researcher and increases her perceived cost of conducting the study (or decreases her perceived benefit), the planner prefers *no* pre-analysis plan and enforces the publication rule with a more stringent cutoff. This preference is exacerbated as the publication constraints become *more* binding. Intuitively, with a higher perceived cost of the pre-analysis plan, researchers may give up on conducting the study unless the planner publishes findings that are less interesting to the public. This is undesirable for the planner when there are publication constraints. In this case, enforcing the optimal publication rule under (unobserved) data manipulation increases the planner's objective and does not restrict the supply of research. This final result, therefore, complements a large empirical literature on this topic (Olken, 2015; Banerjee et al., 2020; Miguel, 2021) by illustrating *trade-offs* in the use of a pre-analysis plan which, we show here, are tightly linked with its perceived cost on the researcher.

Related literature This paper connects to a growing literature that studies economic models to analyze statistical protocols. In the context of scientific communication, Abadie (2020); Andrews and Shapiro (2021); Andrews and Kasy (2019); Kitagawa and Vu (2023), and most closely related Frankel and Kasy (2022) have analyzed how research findings are or should be reported in research studies to inform the public. Different from this paper, none of these references account for the researcher's incentives in the design of the optimal communication protocol. Our goal here is to formally analyze the effect of different communication strategies on the incentives of the researcher. This allows us to show that optimal publication should not only depend on the relevance of the findings, but also on the research costs and incentives associated with these.

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 incentives of researchers. Other work in this line include Chassang et al. (2012); Tetenov (2016); Banerjee et al. (2017); Spiess (2024); Henry and Ottaviani (2019); Banerjee et al. (2020); Di Tillio et al. (2021); Viviano et al. (2024); Williams (2021); McCloskey and Michaillat (2022); Yoder (2022); Bates et al. (2022); Libgober (2022); Bates et al. (2023); Kasy and Spiess (2023), with a particular focus here on the publication process. Different from these references, we *both* analyze settings where the researcher may choose the research design absent private information, or choose the design and *manipulate* reported findings with private information. This allows us to formalize trade-offs between different publication strategies, and formally study ideas when/whether unsurprising results should be published, and whether data manipulation should occur in equilibrium (Glaeser, 2006).

In particular, an important distinction from some of these models studying approval decisions, such as Tetenov (2016); Viviano et al. (2024); Bates et al. (2022, 2023) is that these papers assume that researchers thruthfully report the statistics sampled from their study (abstracting from questions about data manipulation or design choice studied in this paper). Spiess (2024) 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.

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) reader, and a social planner. The reader and social planner are interested in learning a given parameter $\theta_0 \in \Theta$, such as the average treatment effect of a given intervention. A researcher conducts a study to inform the reader about θ_0. A study is summarized by (X, Δ), where Δ denotes the design and

$$X(\Delta)|\theta_0 \sim \mathcal{N}\Big(\theta_0 + \beta_\Delta, S_\Delta^2\Big), \quad \theta_0 \sim \mathcal{N}(0, \eta_0^2) \tag{1}$$

where β_Δ and S_Δ^2 denote the bias and variance of a design Δ and η_0^2 the prior variance.

If a study arrives, it will be evaluated according to a decision rule $p(X, \Delta) \in [0, 1]$, interpreted as the decision to publish the findings of the study, where $p \in \mathcal{P}$ lies in a pre-specified space $\mathcal{P}$ which is subset of all Borel measurable functions. Constraints encoded in $\mathcal{P}$ may capture constraints on the information about Δ available to the social planner. The reader takes an action

$$a_p^\star(X(\Delta), \Delta) = \begin{cases} \frac{X\eta_0^2}{S_\Delta^2 + \eta_0^2} & \text{if } p(X, \Delta) = 1 \\ 0 & \text{otherwise.} \end{cases}$$

That is, the audience forms posterior beliefs about θ_0 assuming $\beta_\Delta = 0$, and ignores publication decisions when findings are not published. As $\eta_0^2 \to \infty$ or $S_\Delta^2 \to 0$, this is equivalent to stating that the action of the audience is to use X as a prediction for the parameter θ_0 when the finding is published, and use the prior mean otherwise.

Given (X, Δ, θ_0), the planner incurs a loss

$$\mathcal{L}_p(X, \Delta, \theta_0) = \Big(\theta_0 - a_p^\star(X, \Delta)\Big)^2 - c_p p(X, \Delta). \tag{2}$$

Here, we are interested in the distance between the audience's action and θ_0; c_p denotes the publication (opportunity) cost, capturing publication or attention constraints.

As standard practice, whenever the researcher is indifferent between two designs, we implicitly assume she chooses the design that minimizes the planner's expected loss.

Conditional Δ, publication decision p, and statistics $X(\Delta)$, the researcher's (expected) payoff for given (possibly randomized) publication decision $p(\cdot)$ equals

$$v_p(X(\Delta), \Delta) = p(X(\Delta), \Delta) - C_\Delta, \quad C_\Delta \leq 1 \tag{3}$$

where C_Δ is the (expected) cost to publication over its benefit associated with a given design Δ. Note that C_Δ may denote direct, indirect or even perceived costs associated with the design Δ.[1] Note that there are no economic transfers between the planner and the researcher: the researcher can only be incentivized by increasing its publication probability.

We will consider separately two scenarios: in Section 3, we assume that the researcher

[1] We assume $C_\Delta \leq 1$, since otherwise the design Δ is trivially never chosen by the researcher.

chooses Δ before observing $X(\Delta)$, and Δ is of common knowledge. In Section 4, the researcher first observes each statistic associated with Δ and then privately chooses Δ.

3 Publication rules with observed design choice

This section studies a simple model where each design presents no bias $\beta_\Delta = 0$, but different designs may be associated with different variances S_Δ^2 and costs C_Δ. The researcher chooses the design Δ *before* observing the statistics $X(\Delta)$ and Δ becomes common knowledge.

Assumption 3.1. Let $\Delta \in \{E, O\}$, i.e., researchers can choose between two different designs (E, O), with $\beta_E = \beta_O = 0$. Each design is associated with variances $S_E^2 \leq S_O^2 < \kappa$, and costs $C_E \geq C_O \geq 0$, for arbitrary positive $\kappa < \infty$. For given p, the researcher chooses Δ to maximize $u_p(\Delta) = \mathbb{E}[v_p(X(\Delta), \Delta)]$ where the expectation is with respect to $(X(\Delta), \theta_0)$ distributed as in Equation (1). Once the researcher chooses Δ, this becomes of common knowledge. The social planner minimizes $\mathbb{E}[\mathcal{L}_p(X, \Delta, \theta_0)]$ where the expectation is over (X, θ_0) as in Equation (1) with $\beta_\Delta = 0$.

The design Δ and its corresponding variance S_Δ^2 are observed by the social planner. We will refer to $\Delta = E$ as an experiment and $\Delta = O$ as an observational study, using the choice between an experiment or observational study as a leading example throughout this section.[2]

Returning to our example in a clinical study, $\Delta = E$ is a medical study with a treatment and placebo control group and $\Delta = O$ is a medical study with a pre-specified synthetic control group. The experiment has a positive cost, e.g., associated with raising funds to also recruit patients in the placebo group. The observational study has a smaller cost since researchers only have to recruit the treatment but not control group. Similarly, we assume $\beta_O = 0$, and think of S_O^2 as the overall ex-ante mean-squared error of the study. When a bias occurs in the observational study, we think of this as a random effect whose contribution enters directly into S_O^2 as we discuss further in Example 3.1.[3]

Example 3.1 (Experiment and Observational study)**.** 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 estimation 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

[2] As shown below Equation (7), it is possible to directly extend our results to settings where $\Delta \in \{\emptyset, O, E\}$, where $\Delta = \emptyset$ implies that no study is conducted. This is omitted here for brevity.

[3] As we further discuss just below Equation (7) it is also possible to consider "no study", equivalent to $S_O^2 = \infty$.

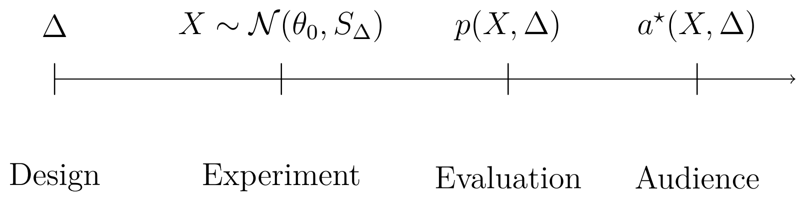

Figure 1: Illustration of the variables in the model. First, researchers pre-specify the population of interest. Second, they run an experiment and draw a statistic X. Finally, a social planner evaluates the experiment based on a decision rule $p(X, \Delta) \in [0, 1]$. Finally, the reader forms a posterior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

the distribution of b_O through a meta-analysis, where σ_B^2 captures the impact of the random bias in statistics' distribution.[4] We can write $X(O) = \theta_0 + \varepsilon'_O, \varepsilon'_O \sim \mathcal{N}(0, S_O^2), S_O^2 = \sigma_B^2 + \sigma_O^2$, where σ_B^2 is the variance of an irreducible error arising the variance of the bias. □

Example 3.2 (Cheap and costly experiment)**.** Suppose that E corresponds to an experiment with costly implementation C_E, where O is an experiment with possibly different variance, but smaller cost of implementation. In this case $\mathbb{E}[X(E)|\theta_0] = \mathbb{E}[X(O)|\theta_0] = \theta_0$ while the two experiments may have different costs (and variances). □

3.1 Publication of experiments

As a first step, we characterize optimal publication when the researcher *must* run an experiment ($\Delta = E$). Before doing so, we need to introduce the following definitions.

Definition 3.1 (Constrained publication rule)**.** Define

$$p_\Delta^\star \in \arg\min_{p \in \mathcal{P}} \mathbb{E}\left[\mathcal{L}_p(X(\Delta), \Delta, \theta_0)\right], \quad \text{such that} \quad \mathbb{E}[v_p(X(\Delta), \Delta)] \geq 0, \tag{4}$$

the best planner's action for a given design Δ under the individual rationality constraint, where $\mathcal{P}$ is the set of all Borel measurable functions.

Definition 3.1 defines a publication rule that minimizes the loss of the social planner for a *given* design Δ, forcing the utility of the researcher to be weakly positive.

Definition 3.2 (Cheap and expensive design)**.** Design Δ is a *cheap* if $C_\Delta \leq \mathbb{P}\left(|X(\Delta)| \geq \gamma_\Delta^*\right)$, where $\gamma_\Delta^* = \frac{S_\Delta^2 + \eta_0^2}{\eta_0^2}\sqrt{c_p}$, and *expensive* otherwise.

[4]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). It is also possible to extend the model to allow the bias to be a random effect, and the audience *not* to incorporate information about the variance of the bias in its updating rule, see Remark 4.

Definition 3.2 classifies experiments into two categories depending on whether the cost of the experiment is sufficiently smaller than the cost of publication and its variance. The choice of the threshold depends on three main components: the prior variance of the parameter of interest η_0^2, the variance of the experiment S_E^2, and the cost of publication c_p. Whenever c_p increases, C_E must become smaller for an experiment to be classified as cheap.

Lemma 3.1 (Publication rules for experiments)**.** The following holds:

$$p^\star_\Delta(X) = \begin{cases} 1\left\{|X| \geq \left|\Phi^{-1}(C_\Delta/2)\right| \sqrt{S_\Delta^2 + \eta_0^2}\right\} & \text{if } \Delta \text{ is an expensive design,} \\ 1\left\{|X| \geq \gamma_\Delta^*\right\}, & \text{otherwise.} \end{cases} \tag{5}$$

Proof of Lemma 3.1. See Appendix A.2. □

Lemma 3.1 characterizes optimal decision rules for experiments under the constraint that the researcher runs an experiment. Here, $p_E^\star$ depends on the variance of the statistic $X(E)$, the prior variance η_0^2, the cost of publication c_p and the *cost* of the experiment C_E. Whenever the experiment is a cheap study, the publication rule is as in Frankel and Kasy (2022); when instead it is an expensive study, the publication rule is similar in spirit to maximin approval decisions in Tetenov (2016) who study hypothesis testing without endogenous choice of the design. Therefore, the dependence of $p_E^\star$ with the cost of the experiment C_E differs from publication rules in Frankel and Kasy (2022), who do not consider the researcher's incentives in the publication process. Similarly, the dependence of the publication rule with the cost of attention c_p differs from approval decisions in Tetenov (2016) and Viviano et al. (2024).

Also, different from previous references, Lemma 3.1 distinguishes two scenarios. For cheap experiments, the threshold does *not* depend on the cost of the experiment. For sufficiently costly experiments, the threshold is increasing in the cost of the experiment.

Remark 1 (Prior mean different from zero)**.** Our framework directly extends to $\theta_0 \sim \mathcal{N}(\mu, \eta_0^2)$ for prior mean μ. In general, μ defines audience's action in the absence of publication. In this case, the optimal publication rule takes the form of $|X - \mu| \geq t$ for a threshold t as in Lemma 3.1. Therefore, as in Frankel and Kasy (2022), we should interpret surprising results as those that move the audience away from its default action μ. □

3.2 Choosing the design

Given Lemma 3.1, we study optimal choice between an observational study or experiment. Let $\mathcal{L}_\Delta^* = \mathbb{E}\left[\mathcal{L}_{p_\Delta^\star}(X(\Delta), \Delta, \theta_0)\right]$ denote the optimal expected loss conditional on implementing design Δ. Because Δ is observed by the social planner, the social planner can incentivize the loss-minimizing design *ex ante* by setting

$$p^\star(X, \Delta) = 1\{\Delta^{\text{planner}} = \Delta\} p^\star_\Delta(X), \quad \Delta^{\text{planner}} \in \arg\min_{\Delta \in \{E,O\}} \mathcal{L}^*_\Delta. \tag{6}$$

It is immediate that $p^\star(X, \Delta)$ maximizes the planner's utility, and it is implementable.

Motivated by this observation, we compare the loss function of the social planner when the researcher implements the experiment or observational study, and the planner chooses the optimal publication rule for each design $p^\star_\Delta$.

As a first step, note that from Lemma A.6 we can write for an expensive experiment

$$\frac{\mathcal{L}^*_E}{\eta_0^2} = 1 + \frac{C_E c_p}{\eta_0^2} - \frac{\eta_0^2}{S_E^2 + \eta_0^2} + r' \tag{7}$$

for a small reminder $r' = \mathcal{O}((1 - C_E)^3)$. Here the denominator η_0^2 corresponds to the social planner's loss when *no study* is implemented. Therefore, whenever $C_e \frac{c_p}{\eta_0^2} > \frac{\eta_0^2}{S_E^2 + \eta_0^2}$ the planner prefers *no* study as opposed to the expensive experiment (even if the research cost is paid privately by the researcher).

This result follows from "scarcity of attention" in our model: for a higher cost of publication, we want to publish fewer results. However, experiments with larger implementation cost C_E must be published with higher frequency, to incentivize the researcher to run the study. Therefore, when publication constraints become more binding, the preference of the social planner shifts towards *less* expensive experiments. We formalize this intuition below.

Proposition 3.1 (Experiment vs. observational study)**.** Let Assumption 3.1 hold.

(i) *Two expensive designs*: Suppose E and O are both expensive designs, with $C_E \geq C_O > \delta > 0$ for a positive constant $\delta > 0$. Then

$$\frac{\mathcal{L}^*_E - \mathcal{L}^*_O}{\eta_0^2} = \frac{c_p}{\eta_0^2}(C_E - C_O) + \frac{\eta_0^2}{(S_E^2 + \eta_0^2)(S_O^2 + \eta_0^2)}(S_E^2 - S_O^2) + r \tag{8}$$

for a remainder r with $|r| \leq K_\delta \Big\{(1 - C_E)^3 + (1 - C_O)^3\Big\}$, for finite constant $K_\delta < \infty$.

(ii) *Expensive vs. cheap design*: Suppose E is an expensive design with $C_E \geq \delta > 0$ and O is a cheap design. Then Equation (8) holds with $\lambda^*_O = 1 - \frac{1}{3\sqrt{2\pi}} \frac{c_p^{1/2}}{\eta_0} \sqrt{\frac{S_O^2}{\eta_0^2} + 1}$ in lieu of C_O and r' in lieu of r, $|r'| \leq K_{\kappa,\delta} \left\{(1 - C_E)^3 + \frac{c_p^2}{\eta_0^4}\right\}$ for a finite constant $K_{\kappa,\delta} < \infty$.

(iii) *Two cheap designs*: If E and O are both cheap designs, $\mathcal{L}^*_E - \mathcal{L}^*_O > 0$ if and only if $S_E^2 > S_O^2$.

Proof of Proposition 3.1. See Appendix A.3. □

Proposition 3.1 provides explicit comparisons between an experiment and an observational study. Here (i) compares two expensive designs, (ii) an expensive experiment and a cheap observational study and (iii) a cheap experiment and observational study.

Clearly, if the observational study has a much larger variance than the experiment ($S_O - S_E \gg 0$), publishing the observational study is sub-optimal. In statistical terms, if a cheap study only "weakly identifies" θ_0, this is expected to be dominated by an expensive study. Similarly, if both the experiment and observational study are cheap, the private research cost should not affect the planner's decision between the two.

Consider instead the more interesting scenario where research costs are binding. For instance, we may be in scenario (i), where the observational study is expensive (but possibly not "too costly"), or scenario (ii), where the experiment is expensive, and the observational study is cheap.[5] Returning to our medical study example, we may expect that the observational study is expensive when it requires recruiting the treatment group, and it is cheap if recruiting the treatment group is costless. In these cases, experiments with *lower* variance than observational studies may not be preferred over observational studies. This is because as the cost becomes binding the planner must publish some uninteresting results at a (possibly large) publication costs. 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. As the cost of publication becomes more binding, the planner's preference shifts towards less costly studies.

Table 1 summarizes the implication of our findings as the cost of the experiment C_E and the cost of attention c_p vary. Similarly, Figure 2 shows how the planner's preference shifts towards noisy observational studies even when the experiment has *no* variance and the observational study is a cheap study.

In summary, these results are suggestive that attention costs can make the planner's utility decreasing in research costs, even when these are private. This introduces trade-offs in the optimal accuracy of an experiment as we further discuss in Remark 2.

Remark 2 (Choice of the sample size)**.** Our framework allows us to study choices of optimal sample size taking into account the corresponding cost, once we let $C_E(S_E^2)$ be an explicit function of the variance S_E^2. Let $C_E(S_E^2) = c_f + \frac{c_v}{S_E^2}$ as a function of a fixed cost c_f and variable cost c_v. For a cheap experiment, we simply choose the smallest variance that guarantees that the researcher runs the experiment or not. For an expensive experiment, we

[5]For (ii) we typically expect $C_E > \lambda_O^*$ as λ_O^* approximates the publication probability for a cheap study. Here, to simplify exposition, here we provide comparisons up to reminders r, r', where r' is small when $c_p/\eta_0^2 < 1$, i.e., when it is typically better to publish some study. The Appendix contains a complete characterization of such reminders.

Table 1: Practical implications of Proposition 3.1 when O is a cheap design. The table summarizes the intuition that for cheap experiments, we should prefer experiments over observational studies if their mean squared error is smaller than the one of an observational. For expensive experiments, comparisons also depend on the cost of the experiment.

Cost of experiment	Cost of publication	MSE comparison	Choosen design
C_E small	c_p any	$S_E^2 - S_O^2 < 0$	Experiment
C_E any	c_p any	$S_E^2 - S_O^2 > 0$	Obs studies
C_E any	c_p small	$S_E^2 - S_O^2 < 0$	Experiment
C_E large	c_p any	$S_E^2 - S_O^2 \ll 0$	Experiment
C_E large	c_p large	$S_E^2 - S_O^2 \not\ll 0$	Obs studies

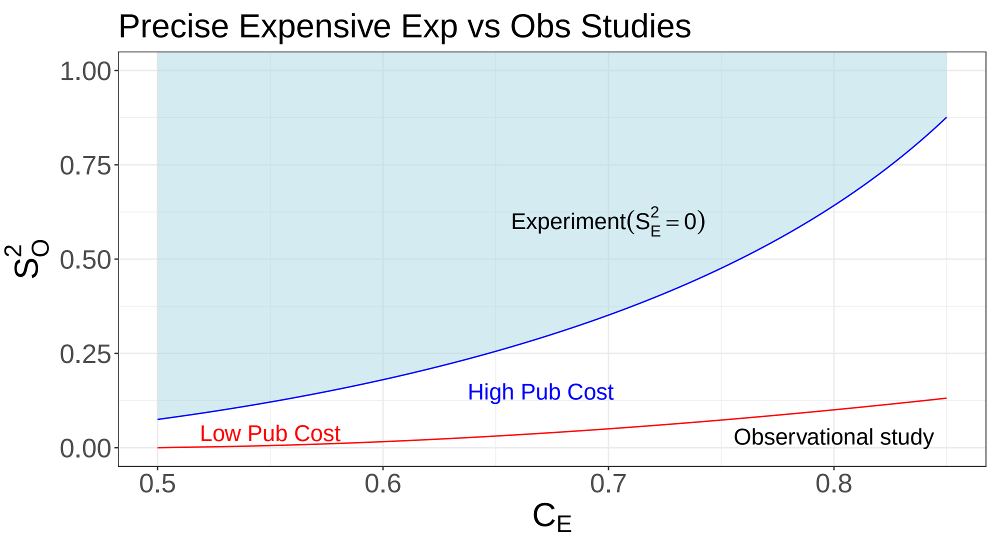

Figure 2: The figure reports on the x-axis the cost of a non-cheap experiment with no variance ($S_E^2 = 0$), and on the y-axis the rescaled squared error S_O^2/η_0^2 of a cheapobservational study using the exact expression obtained from Lemmas A.2 and A.3 The red line denotes the frontier of values for cost of publication $c_p/\eta_0^2 = 0.5$, and the blue lines for the higher cost of publication $c_p/\eta_0^2 = 1$. 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 an observational study may be preferred over an experiment, even if its mean squared error is larger (but not too large), whenever the experiment is sufficiently costly. As the cost of publication increases, the planner prefers more the observational study.

can invoke Lemma A.5 and obtain (approximate) first order conditions $c_p c_v = \frac{S_E^4}{(S_E^2/\eta_0^2+1)^2}$. It follows that whenever $\eta_0 \gg S_E$, at the optimum

$$S_E^2 \approx \sqrt{c_p c_v}.$$

That is, for an expensive experiment, we choose the variance proportional to the square root of the cost of publication and variable cost. This choice is dictated by the researcher's constraints. As the variable cost increases, the researcher's cost increases, which is compensated by an increase in the variance. Similarly, as the publication constraint becomes more binding, we prefer less costly (and hence less accurate) experiments. This motivates a variance S_E^2 *increasing* in the cost of publication c_p. □

Remark 3 (Multiple publications for costly experiments)**.** To incentivize costly experiments, it is possible to allow researchers to publish two papers using the same data collected in the experiment. This form of incentive implicitly increases the rewards to researchers for large experiments or equivalently halves the cost C_E of the experiment paid by the researcher and doubles the cost of publication to $2c_p$. Publishing two papers with the same experimental data may be optimal for some costly experiments. □

Remark 4 (Harmful observational studies)**.** Here, a cheap design (whether it is observational or experimental) always outperforms conducting no study. One could consider scenarios where observational studies are harmful to the audience even in the absence of data manipulation. This occurs when the audience fails to incorporate the bias of a study when updating their beliefs. Returning to Example 3.1, this corresponds to assuming that the audience's posterior belief *overweight* the statistic from an observational study, forming an incorrect belief $X(O)w$ as opposed to $X(O)w'$, where $w = \frac{\eta_0^2}{\sigma_O^2+\eta_0^2} \geq w' = \frac{\eta_0^2}{\sigma_O^2+\eta_0^2+\sigma_B^2}$.

In these contexts, the planner's preference shift towards experiments. For example, the planner may prefer a cheap experiment over a cheap observational study even when the experiment has a larger mean squared error than the observational study. □

4 Publication rules with p-hacking

In this section, we turn to 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 4.1. Consider a class of designs $\Delta \in \mathbb{R} \cup \emptyset$, with $\beta_\Delta = \Delta$ known to the researcher but not to the social planner, and $X(\Delta) = \theta_0 + \beta_\Delta + \varepsilon, \varepsilon|\theta_0 \sim \mathcal{N}(0, S_E^2)$, for $\Delta \neq \emptyset$. The researcher observes $\theta_0 + \varepsilon$ and chooses Δ (and so β_Δ) to maximizes her realized payoff $v_p(X(\Delta), \Delta)$ in Equation (3), after observing $X(\Delta)$. Let $C_\Delta = c_d|\beta_\Delta| + c_e$ for $\Delta \neq \emptyset$ where $c_d < \infty, c_e < 1$, and $C_\emptyset = 0$. The social planner chooses $p(X)$ as a function of X only (and not Δ).

Figure 3 illustrates the model: the researcher chooses deterministically the bias of the reported statistic. She, however, pays 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. In our stylized description, this is approximated by defining X as the sum of $\theta_0 + \varepsilon$ plus a bias arising from manipulation. The cost

C_Δ captures reputational or computational costs associated with the manipulation, assumed to be increasing and linear in the bias β_Δ. The researcher observes $\theta_0 + \varepsilon$, and hence, she maximizes her *realized* utility conditional on the observed statistics when choosing Δ. As we discuss in Section 2, the audience (but not the planner) is unaware that published findings may have some data manipulation.

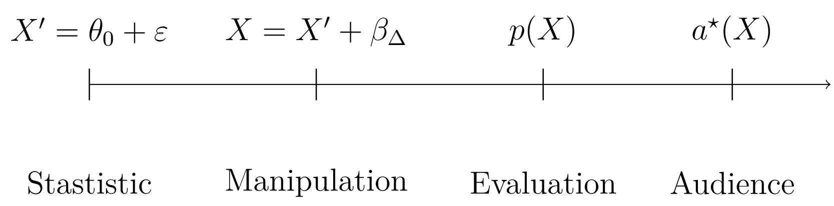

Figure 3: Illustration of the variables in the model. First, researchers observe the vector of statistics. They then manipulate the design by introducing a bias into the statistics and maximize their private utility. The social planner does not observe the bias, and evaluate the study based on a publication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

Finally, note that 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 observational studies. We interpret this assumption as stating that the variance of the statistic is observable to the planner and audience and truthfully reported by the researcher, focusing on mechanisms that introduce unobserved bias in the reported statistic.

4.1 Optimal publication rule under manipulation

The planner does not observe β_Δ, knows S_E^2, and minimizes her expected loss over (θ_0, ε), keeping into account the best reaction of the researcher. That is, we define an optimal publication rule as

$$p^\star \in \arg\max_{p \in \mathcal{P}} \mathbb{E}_{X,\theta_0}\left[\mathcal{L}_p(X(\Delta_p^\star), \Delta_p^\star, \theta_0)\right], \quad \Delta_p^\star = \arg\max_{\Delta} v_p(X(\Delta), \Delta) \tag{9}$$

where $\mathcal{P}$ is the set of all Borel measurable functions $p(X, d)$ constant in d (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 4.1 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 1. A *linearly smoothed cutoff rule* with cutoff X^* and slope m is defined by

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

A linearly smoothed cutoff rule considered here is a deterministic publication rule above and below thresholds $(X^* - \frac{1}{m}, X^*)$ for given m and it randomizes the publication chances between these two values, with publication probability increasing in the value of the reported statistic $|X|$. To gain further intuition, consider the threshold $\gamma_E^* = \frac{S_E^2 + \eta_0^2}{\eta_0^2}\sqrt{c_p}$. Then $X^* = \gamma_E^*$ and $m = \infty$ corresponds to a publication rule in Frankel and Kasy (2022), i.e., a publication rule for cheap experiments. When $m < \infty$ and $X^* > \gamma_E^*$, the linearly smoothed cutoff rule publishes with probability one more surprising results than rules in the absence of manipulation, while randomizing around the threshold. In the following theorem we characterize the optimal publication probability in contexts with manipulation.

Theorem 1 (Optimal publication probability)**.** *Under the model in Assumption 4.1:*

(a) There is a cutoff $X^* \in \left(\gamma_E^*, \gamma_E^* + \frac{1-c_e}{c_d}\right)$ such that p_{X^*,c_d} is optimal.

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

Proof. See Appendix A.4. □

Theorem 1 characterizes the optimal publication rule under manipulation. The rule is a smoothed cutoff rule that: (i) does not publish results below a certain cutoff $X^* - \frac{1}{c_d}$; (ii) randomizes whenever X is just above the cutoff $X^* - \frac{1}{c_d}$ but below X^* and it publishes for results above X^*.

To gain further intuition, it is instructive to compare this with the optimal publication rule for a *cheap* experiment in Lemma 3.1. For simplicity, suppose $c_e = 0$ so that we abstract from fixed costs. Consider first a scenario in which the social planner ignores manipulation. Then we would observe *bunching* around the cutoff to publish γ_E^*, as researchers with $\theta_0 + \varepsilon \in (\gamma_E^* - \frac{1}{c_d}, \gamma_E^*)$ would introduce a bias to publish. Researchers with $\theta_0 + \varepsilon < \gamma_E^* - \frac{1}{c_d}$ would find it nonprofitable to introduce any bias (as the cost would not compensate the benefits) and therefore would not publish.

Next, suppose that the planner introduces randomization in the publication rule whenever $X \in (\gamma_E^* - \frac{1}{c_d}, \gamma_E^*)$ as described in Theorem 1. It follows that the randomization device in

Theorem 1 makes the researcher indifferent between manipulating and not manipulating the data, at the cost of publishing some unsurprising results. However, this choice is sub-optimal as the planner may publish too many unsurprising results.

The last step is to increase the threshold X^* to compensate the loss from publishing unsurprising results. In summary, optimal publication rules with manipulation (a) increase the cutoff for publication; (b) allow for randomization at the margin below the cutoff.

Table 2 illustrates main features associated with each action. The first action is equivalent to ignoring manipulation, as in settings considered in Frankel and Kasy (2022). Manipulation in this scenario has testable implications since it introduces large bunching around the cutoff, see e.g., the discussion in Elliott et al. (2022) among others. Through sufficient randomization below the cutoff, we can guarantee no bunching (and no manipulation). The last step is then to increase the cutoff. We should then observe just some bunching above the old cutoff and below the new one (see Proposition 4.1). This corresponds to the planner's preferred publication rule.

Table 2: Sequence of actions to decrease the loss function. We start from a simple cutoff rule for cheap experiment wrongly assuming no data manipulation. Under this rule, we observe bunching of X around the cutoff. We then randomize below the cutoff to make researcher indifferent between manipulation and not. Finally, we increase the cutoff to publish less unsurprising results.

Action	Testable observation	Published findings	Manipulation
Select optimal cutoff rule without manipulation	Large bunching	Only surprising findings are published	Large
Add randomization below cutoff	No bunching	Many unsurprising findings are published	None
Increase the cutoff + randomize below	Some bunching	Some unsurprising findings are published	Some

4.2 Some implications

In this section, we explore some of the implications of Theorem 1. The first implication we study is whether, in the *planner's preferred equilibrium*, some researcher would manipulate their findings.

Proposition 4.1 (Manipulation in equilibrium)**.** Under the model in Assumption 4.1, under the publication rule in Theorem 1 and the planner's preferred equilibrium for $|\theta_0 + \varepsilon| \in (\gamma_E^*, X^*)$, we have $|\beta_{\Delta^\star_{p^\star}}| > 0$.

Proof. See Appendix A.5. □

Proposition 4.1 states that some manipulation should occur just below the threshold. Intuitively, if researchers could not manipulate their findings, then at the optimum, we would have published findings above the threshold γ_E^*. However, because manipulation can occur, the social planner increases the threshold from γ_E^* to X^*. It follows that manipulation is beneficial in the interval between the old threshold γ_E^* and the "new" threshold X^*. It guarantees that findings that should have been published in the absence of manipulation do get published, while the cost of manipulation is second order. This result suggests that in equilibrium, we should not force manipulation to be exactly zero. A practical implication of Proposition 4.1 is that we observe a discontinuous distribution of X below X^* and above γ_E^* *at the optimum.*

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

Proposition 4.2 (Implementation costs)**.** Under the model in Assumption 4.1, for $c_e \geq 1 - c_d\gamma_E^*$, the set of optimal cutoffs X^* in Theorem 1 is decreasing in c_e in the strong set order.

Proof. See Appendix A.6. □

Proposition 4.2 shows that for expensive 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 Proposition 3.1 in the absence of manipulation: in the absence of manipulation, the planner may force the researcher to run cheap over expensive studies. However, with manipulation, fixed costs imply a lower chance of manipulation, motivating increasing the chance of publication for such studies.

Next, we states that some findings below the optimal cutoff γ_E^* *in the absence* of data manipulation, do get published in the planner's preferred equilibrium under manipulation.

Proposition 4.3 (Some unsurprising results are published)**.** Under the model in Assumption 4.1, under the publication rule in Theorem 1 and the planner's preferred equilibrium, for some values of $|\theta_0 + \varepsilon| < \gamma_E^*$, we have $p^\star(X) > 0$ but $\beta_{\Delta_{p^*}^*} = 0$.

Proof. See Appendix A.7. □

Proposition 4.3 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 optimal publication rule without manipulation as in Frankel and Kasy (2022).

5 Implications for alternative publication mechanisms

In this section we study the implications of our results for alternative publication mechanisms, such requiring pre-analysis plans or introducing a registry of non-surprising results.

5.1 Pre-analysis plans

We first study the implications of Theorem 1 for pre-analysis plans.

Definition 5.1 (Pre-analysis vs possible manipulation)**.** Consider the following:

(A) Pre-analysis plan: Researchers cannot manipulate their findings ($c_d = \infty$), thrutfully report $X = \theta_0 + \varepsilon$, and pay a fixed cost $c_e > \delta$, for arbitrary constant $\delta > 0$; this scenario corresponds to the model in Assumption 3.1, with $\Delta = E$ and cost c_e. The social planner chooses a publication rule $p_E^\star$ as in Definition 3.1 and incurs a loss $\mathcal{L}_E^* = \mathbb{E}[\mathcal{L}_{p_E^\star}(X, E, \theta_0)]$ where X, θ_0 are distributed as in Equation (1) with $\beta_E = 0$.

(B) Possible manipulation: Assumption 4.1 holds, with $\Delta \in \mathbb{R} \cup \emptyset$ with $c_d < \infty, c_e = 0$, corresponding to researcher able to manipulate their findings after observing $\theta_0 + \varepsilon$. The social planner chooses a publication rule $p^\star$ as in Equation (9) and incurs its corresponding expected loss $\mathcal{L}_M^* = \mathbb{E}_{X,\theta_0}\left[\mathcal{L}_{p^\star}(X(\Delta_{p^\star}^\star), \Delta_{p^\star}^\star, \theta_0)\right]$ 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.[6] The cost c_e may capture either costs of conducting a new experiment, or perceived (psychological) costs of the pre-analysis plan.

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). For simplicity here, we implicitly assume the two studies have the same variance equal to S_E^2, although one could extend these results to different variances.

Proposition 5.1 (Value of pre-analysis plan)**.** Under (A) and (B) in Definition 5.1:

(i) If $\Delta = E$ is a cheap design, then $\mathcal{L}_E^* - \mathcal{L}_M^* \leq 0$;

[6] For simplicity, we consider here settings where, once the researcher commits to (A), the planner chooses the optimal publication probability such that the researcher is better off to (ex-ante) choose to conduct the experiment instead of doing no study. This implies that the benefits from such a study under the optimal publication rule outweigh the benefits from no study.

(ii) If $\Delta = E$ is an expensive design, then

$$\frac{1}{\eta_0^2}\Big(\mathcal{L}_E^* - \mathcal{L}_M^*\Big) \geq \frac{c_p}{\eta_0^2}(c_e - \lambda_E^*) - \frac{1}{\eta_0^2 c_d^2} + r'$$

where $\lambda_E^* = 1 - \frac{1}{\sqrt{2\pi}3}\frac{c_p^{1/2}}{\eta_0}\sqrt{\frac{S_O^2}{\eta_0^2}+1}, |r'| \leq K_{\kappa,\delta}\Big((c_e-1)^3 + \frac{c_p^2}{\eta_0^4}\Big)$ for a finite constant $K_{\kappa,\delta} < \infty$.

Proof. See Appendix A.8. □

In Proposition 5.1, we show that a pre-analysis plan is preferred over no pre-analysis plan if the cost of the pre-analysis plan is sufficiently low.

When the (perceived) cost of a pre-analysis plan is high, the trade-off depends on (i) the perceived cost of a pre-analysis plan c_e and (ii) the cost of data manipulation c_d. Clearly, if $c_d = 0$, then 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, *no* pre-analysis plan is preferred for sufficiently high researcher's perceived cost. This is despite the cost of the pre-analysis plan c_e being private and paid by the researcher and not by the planner.

Different from (and complementary to) models where either no experimentation or no reputational costs occur (e.g. Kasy and Spiess, 2023; Spiess, 2024), these results illustrate *trade-offs* in the choice of the pre-analysis plans. A larger perceived cost of pre-analysis may decrease the supply of research, making the planner prefer possible manipulation.

5.2 Registry of unsurprising results

A core theme is that if attention is costly, publication decisions must take into account the *private* research cost. It is natural to ask whether we can think of publication mechanisms that reward studies without paying such a cost of attention.

For instance, consider a scenario where the researcher conducts a (possibly) expensive study with cost C_E in the model in Section 3 – i.e., abstracting from manipulation for simplicity. Because attention is costly, we only want to publish sufficiently surprising results. Suppose that we implement the first-best policy, ignoring the researcher's costs. Following Lemma 3.1, this takes the form $p_E^\star = 1\{|X| \geq \gamma_E^*\}$ of a threshold crossing rule.

In our model in Section 3, if the experiment is costly, however, the researcher may choose the outside option. To guarantee implementability of $p_E^\star$ suppose now we can reward the researcher whose results are not published in the journal by publishing the results in a repository of unsurprising results. This repository may have no cost c_p and guarantee a reward R to the researcher. The researcher's objective function upon conducting the experiment reads as

$$\mathbb{E}[v_{p^*_E}(X(E), E)] = \mathbb{P}(|X(E)| \geq \gamma^*_E) + R\mathbb{P}(|X(E)| \leq \gamma^*_E) - C_E.$$

When choosing $R \geq \frac{C_E+1}{\mathbb{P}(|X(E)| \geq \gamma^*_E)} - 1$, the researcher is always indifferent between conducting or not the research study. Rewarding results through a "repository of unsurprising results", with a reward increasing in the cost of the experiment guarantees that the first-best policy is implementable at no attention cost for the audience.

6 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 (quasi)experiments with lower accuracy (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) knowingly 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 that would be optimal if the planner naively assumed researchers do not manipulate. 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 large perceived burden on the researcher.

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 5, studying more complex action spaces can shed light on alternative mechanisms relevant to scientific communication.

Future research should also 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 possible reputational costs associated with it, such parameters may be learned using information from meta-analyses through the distribution of the submitted findings (different form, but in the spirit of Andrews and Kasy, 2019). In this case, experimentation or different norms across different journals can help to identify 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., 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.

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.

McCloskey, A. and P. Michaillat (2022). Incentive-compatible critical values. Working paper, National Bureau of Economic Research.

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

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*, 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.

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 Omitted Proofs

We will define $\Phi(\cdot)$ the cumulative density function of a Standard Normal Distribution and $\phi(\cdot)$ its probabiliy density function. We will use the $y = \mathcal{O}(x)$ notation to indicate that $y \leq Kx$ for a finite constant K.

A.1 Preliminary Lemmata

The following lemma provides an approximation to a function of the cost that will be used in the characterization of the planner's loss function for expensive designs.

Lemma A.1. Let $\Lambda(C_E) = \Phi^{-1}(1 - C_E/2)\phi(\Phi^{-1}(1 - C_E/2)), C_E > \delta > 0$ for a constant $\delta > 0$. Then $\Lambda(C_E) = \frac{1}{2}(1 - C_E) + r_l$ where $|r_l| \leq K_\delta\left((1 - C_E)^3\right)$, for a finite constant $K_\delta < \infty$.

Proof of Lemma A.1. The proof follows directly from a second order Taylor approximation to $\Phi^{-1}(1 - C_E/2)$ around $C_E = 1$ (where the second order derivative equals zero) and a first order Taylor approximation to $\phi(\Phi^{-1}(1 - C_E/2))$ around $C_E = 1$ (where the derivative equal zero), and using the fact that $\Phi^{-1}(1/2) = 0$. □

The following lemma provides a simple decomposition of the social planner's loss function conditional on the realized statistic X.

Lemma A.2 (Loss function)**.** Suppose that Assumption 3.1 holds. Then

$$\mathbb{E}\Big[\mathcal{L}_p(X(\Delta), \Delta, \theta_0)\Big|X(\Delta)\Big] = \Big[c_p - \frac{\eta_0^4}{(S_\Delta^2 + \eta_0^2)^2}X(\Delta)^2\Big]p(X, \Delta) + \mathbb{E}[\theta_0^2|X(\Delta)]$$

Proof of Lemma A.2. Recall that under Assumption 3.1, $\beta_\Delta = 0$. Using the fact that $\mathbb{E}[\theta_0|X(\Delta)] = X(\Delta)\frac{\eta_0^2}{S_\Delta^2+\eta_0^2}$, and $\mathbb{E}\Big[\theta_0|X(\Delta)\Big] = \frac{\eta_0^2}{S_\Delta^2+\eta_0^2}X(\Delta)$. We can write

$$\begin{aligned}
\mathbb{E}\Big[\mathcal{L}_p(X(\Delta), \Delta, \theta_0)\Big|X, \beta_\Delta\Big] &= \Big[\Big(\frac{\eta_0^2}{S_\Delta^2 + \eta_0^2}\Big)^2\beta_\Delta^2 + c_p - \Big(\frac{\eta_0^2}{S_\Delta^2 + \eta_0^2}(X(\Delta) - \beta_\Delta)\Big)^2\Big]p(X, \Delta) + \mathbb{E}[\theta^2|X] \\
&\quad - 2\frac{\eta_0^2}{S_\Delta^2 + \eta_0^2}\beta_\Delta\mathbb{E}[\theta_0|X(\Delta)]p(X, \Delta) + 2\frac{\eta_0^2(X(\Delta) - \beta_\Delta)}{S_\Delta^2 + \eta_0^2}\frac{\eta_0^2}{S_\Delta^2 + \eta_0^2}\beta_\Delta p(X, \Delta) \\
&= \Big[c_p - \Big(\frac{\eta_0^2}{S_\Delta^2 + \eta_0^2}X(\Delta)\Big)^2\Big]p(X, \Delta) + \mathbb{E}[\theta_0^2|X(\Delta)].
\end{aligned}$$

□

The following lemma provides an exact characterization of the social planner's loss function for a publication rule corresponding to a threshold rule. This will then be used to characterize the loss function for cheap designs.

Lemma A.3 (Loss for threshold rule)**.** Suppose that Assumption 3.1 holds. Then for any rule $p_\Delta = 1\{\frac{|X(\Delta)|}{\sqrt{S_\Delta^2+\eta_0^2}} \geq t_\Delta\}$, for given threshold t_Δ

$$\mathbb{E}\Big[\mathcal{L}_{p_\Delta}(X(\Delta), \Delta, \theta_0)\Big] = \eta_0^2 + 2\Phi(-t_\Delta)\left[c_p - \frac{\eta_0^4}{S_\Delta^2+\eta_0^2}\Big[1 + \frac{t_\Delta \phi(t_\Delta)}{(1-\Phi(t_\Delta))}\Big]\right].$$

Proof of Lemma A.3. First, we write

$$\mathbb{E}\left[\mathcal{L}_{p^\star_\Delta}(X(\Delta), \Delta, \theta_0)|X(\Delta)\right] = \mathbb{E}[\theta_0^2|X] - \left(\Big(\frac{\eta_0^2}{S_\Delta^2+\eta_0^2}\Big)^2 X^2 - c_p\right)1\Big\{\frac{1}{S_\Delta^2+\eta_0^2}X^2 \geq t_\Delta^2\Big\}$$

We can write

$$\begin{aligned}&\mathbb{E}\left[\Big(\frac{\eta_0^2}{S_\Delta^2+\eta_0^2}\Big)^2 X^2 1\Big\{\frac{1}{S_\Delta^2+\eta_0^2}X^2 \geq t_\Delta^2\Big\}\right]\\ &= \underbrace{\mathbb{E}\left[\Big(\frac{\eta_0^2}{S_\Delta^2+\eta_0^2}\Big)^2 X^2 \Big| \frac{|X|}{\sqrt{S_\Delta^2+\eta_0^2}} \geq t_\Delta\Big\}\right]}_{(I)} \times \underbrace{\mathbb{P}\left(\frac{|X|}{\sqrt{S_\Delta^2+\eta_0^2}} \geq t_\Delta\right)}_{(II)}.\end{aligned}$$

Recall that $X \sim \mathcal{N}(0, \eta_0^2 + S_\Delta^2)$. It is easy to show that by symmetry of the Gaussian distribution

$$(I) = \mathbb{E}\left[\Big(\frac{\eta_0^2}{S_\Delta^2+\eta_0^2}\Big)^2 X^2 \Big| \frac{X}{\sqrt{S_\Delta^2+\eta_0^2}} \geq t_\Delta\right].$$

Using properties of the (truncated) Gaussian distribution, we can write

$$(I) = \frac{\eta_0^4}{S_\Delta^2+\eta_0^2}\Big[1 + \frac{t_\Delta \phi(t_\Delta)}{1-\Phi(t_\Delta)}\Big], \quad (II) = 2\Phi(-t_\Delta),$$

and the lemma follows. □

The following lemma is a simple definition of $p^\star_\Delta$ in Lemma 3.1.

Lemma A.4. For $p^\star_\Delta$ in Equation (5), it follows that $p^\star_\Delta = 1\{\frac{|X|}{\sqrt{S_\Delta^2+\eta_0^2}} \geq t_\Delta\}$ where $t_\Delta = \Big(\frac{\sqrt{S_\Delta^2+\eta_0^2}}{\eta_0^2}\Big)\sqrt{c_p - \max\{\lambda(C_\Delta), 0\}}$, and $\lambda(C_\Delta) = c_p - \Big(\frac{\eta_0^2}{\sqrt{S_\Delta+\eta_0^2}}\Phi^{-1}(1 - C_\Delta/2)\Big)^2$.

Proof. The proof is immediate from rearrangement. □

The following lemma provides an exact characterization of the loss planner's loss function in the presence of an expensive design.

Lemma A.5 (Loss for large costs)**.** Suppose that Assumption 3.1 holds. Suppose that $\lambda(C_\Delta) > 0$ as defined in Lemma A.4. Then

$$\mathbb{E}[\mathcal{L}_{p^\star_\Delta}(X(\Delta), S_\Delta, \theta_0)] =$$
$$\eta_0^2 + C_\Delta\Big[c_p - \frac{\eta_0^4}{S_\Delta^2 + \eta_0^2}\Big] - \frac{2\eta_0^4}{S_\Delta^2 + \eta_0^2}\Phi^{-1}(1 - C_\Delta/2)\phi(\Phi^{-1}(1 - C_\Delta/2)).$$

Proof of Lemma A.5. The proof follows directly from Lemma A.3, noting that when $\lambda(C_\Delta) > 0$, the constraint is binding, and $C_\Delta/2 = \Phi(-t_\Delta)$. □

A.2 Proof of Lemma 3.1

Using Lemma A.2 and the lagrangian formulation with multiplier λ, the objective reads as

$$\int \Big[c_p - \lambda - \Big(\frac{\eta_0^2}{S_E^2 + \eta_0^2}x\Big)^2\Big]p(x)dF_X(x) + \lambda C_E$$

where $F_X = \mathcal{N}(0, S_E^2 + \eta_0^2)$, and $\lambda \geq 0$ is the lagrangian multiplier. We can solve and obtain $p_\lambda(x) = 1\{(\frac{\eta_0^2}{S^2+\eta_0^2}x)^2 \geq c_p - \lambda\}$ as the minimizer of the above expression. From complementary slackness, whenever the constraint is binding, we can write

$$\lambda : \int p_\lambda(x)dF_X(x) = C_E,$$

so that (since $X \sim \mathcal{N}(0, S_E^2 + \eta_0^2)$),

$$2\Phi\Big(-\sqrt{c_p - \lambda}\frac{\sqrt{S_E^2 + \eta_0^2}}{\eta_0^2}\Big) = C_E \Rightarrow \lambda = c_p - \Big(\frac{\eta_0^2}{\sqrt{S_E + \eta_0^2}}\Phi^{-1}(1 - C_E/2)\Big)^2.$$

When the constraint is not binding, $\lambda = 0$ from complementary slackness. Under Definition 3.2, an experiment is expensive when $\lambda \geq 0$ and cheap otherwise. Our result directly follows from simple rearrangement of p_λ under these two scenarios using Lemma A.4.

A.3 Proof of Proposition 3.1

We prove (i), (ii), (iii) separately.

Proof of (i)**.** Consider first the loss function of the expensive experiment. We approximate this using the following lemma. The following lemma provides an approximation of the planner's loss in the presence of an expensive design.

Lemma A.6 (Approximate loss for large costs)**.** Suppose that Assumption 3.1 holds. Suppose that $\lambda(C_\Delta) > 0$, where $C_\Delta > \delta$ for arbitrary constant $\delta > 0$. Then

$$\mathbb{E}[\mathcal{L}_{p^\star_\Delta}(X(\Delta), S_\Delta, \theta_0)] = \eta_0^2 + C_\Delta\Big[c_p - \frac{\eta_0^4}{S_\Delta^2 + \eta_0^2}\Big] - \frac{\eta_0^4}{S_\Delta^2 + \eta_0^2}(1 - C_\Delta) + \mathcal{O}(\eta_0^2(1 - C_\Delta)^3).$$

Proof of Lemma A.6. The lemma follows directly from Lemma A.5 and Lemma A.1. □

Then by Lemmas A.4, A.6, we can write

$$\mathbb{E}[\mathcal{L}_{p^\star_E}(X(E), S_E, \theta_0)] = \eta_0^2 + C_E\Big[c_p - \frac{\eta_0^4}{(S_E^2 + \eta_0^2)}\Big] - \frac{\eta_0^4}{(S_E^2 + \eta_0^2)}(1 - C_E) + \mathcal{O}(\eta_0^2(1 - C_E)^3).$$

Similarly it follows for O being an expensive design. The proof completes from rearrangement.

Proof of (ii)**.** Before stating the result, we need to introduce the following lemma. The following lemma approximates the planner's loss function in the presence of a cheap design.

Lemma A.7 (Taylor approximation)**.** Suppose that Assumption 3.1 holds and consider $p_\Delta = 1\{\frac{|X|}{\sqrt{S_\Delta^2+\eta_0^2}} \geq t_\Delta\}$, where $t_\Delta = \frac{\sqrt{S_\Delta^2+\eta_0^2}}{\eta_0^2}\sqrt{c_p}$. Then

$$\mathbb{E}[\mathcal{L}_{p_\Delta}(X(\Delta), S_\Delta, \theta_0)] = \eta_0^2 + c_p - \frac{\phi(0)}{3}c_p^{3/2}\frac{\sqrt{S_\Delta^2 + \eta_0^2}}{\eta_0^2} - \frac{\eta_0^4}{S_\Delta^2 + \eta_0^2} + \mathcal{O}(t_\Delta^3 c_p + \eta_0^2 t_\Delta^4).$$

Proof of Lemma A.7. From Lemma A.3, we can write

$$\mathbb{E}[\mathcal{L}_{p_\Delta}(X(\Delta), S_\Delta, \theta_0)] = \eta_0^2 + \underbrace{2\Phi(-t_\Delta)c_p}_{(I)} - \underbrace{2\Phi(-t_\Delta)\frac{\eta_0^4}{S_\Delta^2 + \eta_0^2}}_{(II)} - \underbrace{2\frac{\eta_0^4}{S_\Delta^2 + \eta_0^2}t_\Delta\phi(t_\Delta)}_{(III)}.$$

Consider now a second order Taylor expansion to $\Phi(-t_\Delta) = \Phi(0) - t_\Delta\phi(0) + \frac{1}{2}\phi'(0)t_\Delta^2 + \mathcal{O}(t_\Delta^3)$. Because $\phi'(0) = 0$, this gives us

$$(I) = c_p - c_p t_\Delta\phi(0) + \mathcal{O}(t_\Delta^3 c_p).$$

Consider now a third order Taylor expansion to $\Phi(-t_\Delta) = \Phi(0) - t_\Delta\phi(0) - \frac{1}{6}\phi''(0)t_\Delta^3 + \mathcal{O}(t_\Delta^4)$. Using the fact that $\phi'(x) = -x\phi(x)$, we have $\phi''(0) = -\phi(0)$, giving us $\Phi(-t_\Delta) = \Phi(0) - t_\Delta\phi(0) + \frac{1}{6}\phi(0)t_\Delta^3 + \mathcal{O}(t_\Delta^4)$. Therefore, we can write

$$(II) = 2\Big[\frac{1}{2} - t_\Delta\phi(0) + \frac{1}{6}\phi(0)t_\Delta^3\Big]\frac{\eta_0^4}{S_\Delta^2 + \eta_0^2} + \mathcal{O}(t_\Delta^4\eta_0^2).$$

Finally, consider a second order Taylor approximation to $\phi(t_\Delta) = \phi(0) - \frac{1}{2}t_\Delta^2\phi(0) + \mathcal{O}(t_\Delta^3)$. We can write

$$(III) = 2\frac{\eta_0^4}{S_\Delta^2 + \eta_0^2}t_\Delta\Big[\phi(0) - \frac{1}{2}t_\Delta^2\phi(0)\Big] + \mathcal{O}(t_\Delta^4\eta_0^2).$$

Rearranging terms we obtain

$$(I) - (II) - (III) = c_p - c_p t_\Delta\phi(0) - \frac{\eta_0^4}{S_\Delta^2 + \eta_0^2} + \frac{2\phi(0)}{3}t_\Delta^3\frac{\eta_0^4}{S_\Delta^2 + \eta_0^2} + \mathcal{O}(t_\Delta^3 c_p + t_\Delta^4\eta_0^2)$$

We now use the fact that $t_\Delta = \frac{\sqrt{S_\Delta^2+\eta_0^2}}{\eta_0^2}\sqrt{c_p}$ under the assumptions stated. We obtain

$$c_p t_\Delta\phi(0) - \frac{2\phi(0)}{3}t_\Delta^3\frac{\eta_0^4}{S_\Delta^2 + \eta_0^2} = \frac{1}{3}\phi(0)c_p^{3/2}\frac{\sqrt{S_\Delta^2 + \eta_0^2}}{\eta_0^2}$$

completing the proof. □

For the observational study, because it is a cheap design, we invoke Lemma A.7, and write (where $t_O = \frac{\sqrt{S_O^2+\eta_0^2}}{\eta_0^2}\sqrt{c_p}$ from Lemma A.4)

$$\mathbb{E}[\mathcal{L}_{p_O^\star}(X(O), S_O, \theta_0)] = \eta_0^2 + c_p - \frac{\phi(0)}{3}c_p^{3/2}\frac{\sqrt{S_O^2 + \eta_0^2}}{\eta_0^2} - \frac{\eta_0^4}{S_O^2 + \eta_0^2} + \mathcal{O}(t_O^3 c_p + \eta_0^2 t_O^4).$$

By dividing each expression by η_0^2 and using the fact that S_O^2 is bounded by a finite constant completes the proof.

Proof of (*iii*)**.** Claim (iii) follows directly from Lemma A.3 since the loss function is decreasing in the variance S_Δ^2 and for both the cost is not binding, so that $t_\Delta = \frac{\sqrt{S_\Delta^2+\eta_0^2}}{\eta_0^2}\sqrt{c_p}$ in Lemma A.3. To show this, from Lemma A.4 and following the same steps in the proof of Lemma A.3, it is easy to show that

$$\mathbb{E}[\mathcal{L}_{p_E^\star}(X(O), O, \theta_0)] = \eta_0^2 + 2\Phi(-t_E)\Big[c_p - \frac{\eta_0^2}{S_O^2 + \eta_0^2}(1 + \frac{t_E\phi(t_E)}{(1 - \Phi(t_E))})\Big].$$

That is, the loss for $p_E^\star$ evaluated at $X(O)|\theta_0 \sim \mathcal{N}(\theta_0, S_O^2)$ is increasing in S_O^2. It follows that

$$\mathbb{E}[\mathcal{L}_{p_E^\star}(X(O), O, \theta_0)] - \mathbb{E}[\mathcal{L}_{p_E^\star}(X(E), E, \theta_0)] < 0$$

if and only if $S_O^2 < S_E^2$. Because $\mathbb{E}[\mathcal{L}_{p_E^\star}(X(O), O, \theta_0)] \geq \mathbb{E}[\mathcal{L}_{p_O^\star}(X(O), O, \theta_0)]$ by definition of $p_O^\star$, it follows that if $S_O^2 < S_E^2$, $\mathbb{E}[\mathcal{L}_{p_O^\star}(X(O), O, \theta_0)] - \mathbb{E}[\mathcal{L}_{p_E^\star}(X(E), E, \theta_0)] < 0$. Using the analog argument with $S_E^2 < S_O^2$ we obtain the opposite result, completing the claim.

A.4 Proof of Theorem 1

Let $X_0 = X(0)$ denote the type of the researcher.

Note that $\theta_0 | X_0 \sim \mathcal{N}\left(\frac{\eta_0^2}{\eta_0^2+S_E^2} X_0, \frac{S_E^2 \eta_0^2}{S_E^2+\eta_0^2}\right)$. Hence, writing $\omega = \frac{\eta_0^2}{\eta_0^2+S_E^2}$ the planner's expected loss conditional on X_0 if the researcher chooses a nontrivial design Δ that has publication probability p is

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

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 $0 \leq v \leq 1$, define

$$\mathcal{L}^*(X_0, v) = \min_{p, \beta_\Delta \in [0,1] \times \mathbb{R} \,\middle|\, p - c_d |\beta_\Delta| = v} \left\{ \left(\omega^2 (\beta_\Delta^2 - X_0^2) + c_p\right) p \right\} \tag{10}$$

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

Lemma A.8. (a) If $|X_0| \neq \gamma$, then there is a unique optimizer $(\tilde{p}(X_0, v), \beta_\Delta(X_0, v))$ in (10) given by

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

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

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

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

Proof. Writing $|\beta_\Delta| = \frac{p - u - C_e}{c_d}$, note that (10) can be written equivalently as

$$\mathcal{L}^*(X_0, v) = \min_{p \in [0,1] | p \geq v} \left\{ \left[\omega^2 \left(\frac{p - v}{c_d} \right)^2 - \omega^2 X_0^2 + c_p \right] p \right\}. \tag{11}$$

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

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

$$\mathcal{L}^*(X_0, v) = \min_{p\in[0,1]} \left\{ \left[\omega^2 \left(\frac{p-v}{c_d} \right)^2 - \omega^2 X_0^2 + c_p \right] p \right\}, \tag{12}$$

and moreover that all optimizers $\tilde{p}(X_0, v)$ satisfy $\tilde{p}(X_0, v) > v$. Indeed, taking $p = v$ in (12), we can see that the right hand side is negative. It follows that in optimum in (12), 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 $\mathcal{L}^*(X_0; v) < 0$. It also follows from first-order condition for the optimality of p that

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

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

$$\frac{\partial \mathcal{L}^*(X_0, v)}{\partial v} = \frac{2\omega^2(p^*(v) - v)p^*(v)}{c_d^2} < 0$$

$$\frac{\partial \mathcal{L}^*(X_0; v)}{\partial X_0} = -2\omega^2 X_0 p^*(v) < 0 \qquad \text{for } X_0 > 0.$$

- Case 2: $|X_0| \leq \gamma^*$. In this case, the objective in (11) 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 $\mathcal{L}^*(X_0, v) = [c_p - \omega^2 X_0^2]\, v$. For $|X_0| = \gamma^*$, this function is zero. For $|X_0| < \gamma^*$, 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 $X_0 = \gamma^*$ an indirect utility of $u^* \in [0, 1 - C_e]$. The linearly smoothed cutoff rule $p^*_{\gamma^* + \frac{1-C_e-u^*}{c_d}, c_d}$ lies within this class, and under the planner's preferred equilibrium, delivers expected loss conditional on X_0 of this publication rule is

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

Lemma A.9. Within the class of publication rules that provide type $X_0 = \gamma^*$ an indirect utility of $u^* \in [0, 1 - C_e]$, for all types $X_0 > 0$:

(a) the linearly smoothed cutoff rule $p_{\gamma^* + \frac{1-C_e-u^*}{c_d}, c_d}$ 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 $p_{\gamma^* + \frac{1-C_e-u^*}{c_d}, c_d}$.

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

- Case 1: $X_0 > \gamma^*$. If type X_0 obtains utility u, then type γ^* could obtain utility at least $u - c_d(X_0 - \gamma^*)$ by choosing a design with the same mean as the design chosen by X_0. Hence, we must have that $u - c_d(X_0 - \gamma^*) \leq u^*$. Obviously, we must have that $u \leq 1 - C_e$. By Lemma A.8(b) and the definition of $\mathcal{L}^*$, the expected loss conditional on X_0 must be at least

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

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

- Case 2: $X_0 = \gamma^*$. By Lemma A.8(b) and the definition of $\mathcal{L}^*$, the expected loss conditional on X_0 must be at least $\omega^2(X_0^2 + S_E^2) = \mathcal{L}^*(X_0, u^* + C_e) + \omega^2(X_0^2 + S_E^2)$.

- Case 3: $\gamma^* - \frac{u^*}{c_d} \leq X_0 < \gamma^*$. Type X_0 could obtain utility at least $u^* + c_d(X_0 - \gamma^*) > 0$ by choosing a design with the same mean as the design chosen by γ^*. By Lemma A.8(b) and the definition of $\mathcal{L}^*$, the expected loss conditional on X_0 must be at least

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

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

- Case 4: $X_0 < \gamma^* - \frac{u^*}{c_d}$. By Lemma A.8(b) and the definition of $\mathcal{L}^*$, 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. □

To prove the first part of the theorem, note that Lemma A.9(a) implies that there exists a utility level $u^* \in [0, 1 - C_e]$ for type γ^* such that $p_{\gamma^* + \frac{1-C_e-u^*}{c_d}, c_d}$ (under the planner's preferred equilibrium) is optimal. Writing $v^* = u^* + C_e$, the expected loss of $p_{\gamma^*+\frac{1-v^*}{c_d},c_d}$ (under the planner's preferred equilibrium) is

$$\mathcal{E}(v^*, C_e) = \mathbb{E}_{X_0}\left[\mathcal{L}^*(X_0, \min\{v^* + c_d(X_0 - \gamma^*), 1\})\, 1\left\{|X_0| \geq \gamma^* - \frac{v^* - C_e}{c_d}\right\}\right] + \eta_0^2.$$

Differentiating under the integral sign using Lemma A.8(b), we see that $\frac{\partial \mathcal{E}}{\partial v^*}\Big|_{v^*=C_e} < 0$ and that $\frac{\partial \mathcal{E}}{\partial v^*}\Big|_{v^*=1} > 0$. So for $p_{\gamma^*+\frac{1-C_e-u^*}{c_d},c_d}$ to be optimal, we must have $0 < u^* < 1 - C_e$, hence

$$0 < \frac{1 - C_e - u^*}{c_d} < \frac{1 - C_e}{c_d}.$$

To prove the second part of the theorem, consider any optimal publication rule p^* that delivers utility level u^* to type $X_0 = \gamma^*$. By Lemma A.8(a), the publication rule p^* (under the planner's preferred equilibrium) must deliver expected loss conditional on X_0 equal to that of $p_{\gamma^*+\frac{1-C_e-u^*}{c_d},c_d}$ for almost all type X_0, and $p_{\gamma^*+\frac{1-C_e-u^*}{c_d},c_d}$ must be optimal. In particular, we must have that $0 < u^* < 1 - C_e$. Using these consequences of optimality, we prove the two assertions in this part separately.

- Suppose for sake of deriving a contradiction that $p^*(X) \neq p_{\gamma^*+\frac{1-C_e-u^*}{c_d},c_d}(X)$ for a positive measure of $X > 0$ with $p_{\gamma^*+\frac{1-C_e-u^*}{c_d},c_d}(X) > C_e$. Then, at least one of the following must occur.
  - Case 1: $p^*(X) \neq 1$ for a positive measure of $X > \gamma^* + \frac{1-C_e-u^*}{c_d}$. Then types $X_0 > \gamma^* + \frac{1-C_e-u^*}{c_d}$ with $p^*(X_0) < 1$ must obtain indirect utility less than $1 - C_e$, which, by Lemma A.9(b), must lead to expected loss conditional on X_0 strictly greater than that of $p_{\gamma^*+\frac{1-C_e-u^*}{c_d},c_d}$—a contradiction.
  - Case 2: $p^*(X) \neq u^* + C_e + c_d(X - \gamma^*)$ for a positive measure of results $X \in \left(\gamma^* - \frac{u^*}{c_d}, \gamma^* + \frac{1-C_e-u^*}{c_d}\right)$. Letting $\tilde{p}(X_0, v)$ be as defined in Lemma A.8(a), by continuity, a positive measure of types $X_0 \in \left(\gamma^* - \frac{u^*}{c_d}, \gamma^* + \frac{1-C_e-u^*}{c_d}\right)$ must satisfy

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

    Hence, if such types are to obtain indrect utility $u^* + c_d(X_0 - \gamma^*)$, they must have publication probability different from $\tilde{p}(X, u^* + C_e + c_d(X_0 - \gamma^*))$, which, by Lemma A.8(a), would also lead to expected loss conditional on X_0 strictly greater than that of $p_{\gamma^*+\frac{1-C_e-u^*}{c_d},c_d}$. But Lemma A.9(b) implies for $X_0 \in \left(\gamma^* - \frac{u^*}{c_d}, \gamma^* + \frac{1-C_e-u^*}{c_d}\right)$, to obtain expected loss conditional on X_0 equal to that of $p_{\gamma^*+\frac{1-C_e-u^*}{c_d},c_d}$, type X_0 must obtain indirect utility $u^* + c_d(X_0 - \gamma^*)$—a contradiction.

- Suppose for sake of deriving a contradiction that $p^*(X) > C_e$ for a positive measure of $X > 0$ with $p_{\gamma^*+\frac{1-C_e-u^*}{c_d},c_d}(X) \le C_e$. Then types $X_0 \in (0, \gamma^* - \frac{1-u^*}{c_d})$ with $p^*(X_0) > C_e$ must obtain positive indirect utility, which, by Lemma A.9(b), must lead to expected loss conditional on X_0 strictly greater than that of $p_{\gamma^*+\frac{1-C_e-u^*}{c_d},c_d}$—a contradiction.

A.5 Proof of Proposition 4.1

In the notation of the proof of Theorem 1, let $X_0 \in (\gamma^*, X^*)$. The expected loss conditional on X_0 if type X_0 chooses a nontrivial design Δ is

$$\left(\omega^2(\beta_\Delta^2 - X_0^2) + c_p\right) p,$$

where $p = c_d(|\beta_\Delta| - (X^* - X_0))$. Lemma A.8(a) implies that there is a unique minimizer, which satisfies $\beta_\Delta > 0$.

A.6 Proof of Proposition 4.2

In the notation of the proof of Theorem 1, consider the function $\mathcal{E}(v^*, C_e)$ on the domain $\{(v^*, C_e) \mid 1 \ge v^* \ge C_e \ge 1 - c_d\gamma^*\}$. Differentiating $\mathcal{E}^*$ using the fundamental theorem of calculus implies that

$$\frac{\partial \mathcal{E}}{\partial C_e} = -\frac{2}{c_d}\mathcal{L}^*\left(\gamma^* - \frac{v^* - C_e}{c_d}, C_e\right)\frac{\exp\left(-\frac{\left(\gamma^* - \frac{v^*-C_e}{c_d}\right)^2}{2(S_E^2+\eta_0^2)}\right)}{\sqrt{2\pi(S_E^2+\eta_0^2)}}.$$

Lemma A.8(b) implies that $\mathcal{L}^*\left(\gamma^* - \frac{v^*-C_e}{c_d}, C_e\right) \ge 0$, and it then follows from Lemma A.8(c) that $\frac{\partial^2 \mathcal{E}}{\partial v^* \partial C_e} \le 0$. Hence, $\mathcal{E}$ is a submodular on $\{(v^*, C_e) \mid 1 \ge v^* \ge C_e \ge 1 - c_d\gamma^*\}$, which is a lattice. Topkis's Theorem implies that

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

is increasing in the strong set order in C_e over $[1 - c_d\gamma^*, 1]$. Since $\mathcal{E}(v^*, C_e)$ is the expected loss of $p_{\gamma^*+\frac{1-v^*}{c_d},c_d}$ (under the planner's preferred equilibrium), the proposition follows.

A.7 Proof of Proposition 4.3

In the notation of the proof of Theorem 1, take $|X_0| \in (X_0^* - \frac{1-c_e}{c_d}, \gamma_E^*)$. The expected loss conditional on X_0 if type X_0 chooses a nontrivial design Δ is

$$\left(\omega^2(\beta_\Delta^2 - X_0^2) + c_p\right) p,$$

where $p = c_d(|\beta_\Delta| - (X^* - X_0))$. Lemma A.8(a) implies that there is a unique minimizer, which satisfies $\beta_\Delta = 0$. Since $|X_0| > \frac{1-c_e}{c_d}$, we have $p > c_e$, so type X_0 will choose a nontrivial design and obtain a nonzero publication probability.

A.8 Proof of Proposition 5.1

The first claim follows directly from the fact that if $\Delta = E$ is a cheap design, $p_E^\star$ is the minimizer of the loss function in the absence of incentive compatible constraints and where the individual rationality constraint is not binding at the optimum (i.e., at the optimum $p_E^\star$ is the same if $c_e = 0$ from Lemma 3.1). On the other hand, $\mathcal{L}_M^*$ further imposes additional incentive compatibility constraints. It follows that $p_E^\star$ is the minimizer of the social planner's expected loss function allowing for a less restrictive function class of publication rule.

For the second claim, from Theorem 1, we can write

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

where p' is a sub-optimal publication rule. It follows that under p' all results for which $|\theta_0+\varepsilon| \in [\gamma_E^*, \gamma_E^*+\frac{1}{c_d}]$ are manipulated with $|\beta_{\Delta_{p'}^\star}| = \gamma_E^*+\frac{1}{c_d}-|\theta_0+\varepsilon|$. For all $|\theta_0+\varepsilon| < \gamma_E^*+\frac{1}{c_d}$, $|\beta_{\Delta_{p'}^\star}| = 0$. As a result, we can write, after expanding the quadratic expression in the loss function

$$\mathcal{L}_M^* \leq \underbrace{\mathbb{E}\Big[\mathcal{L}_{p''}(\theta_0+\varepsilon, 0, \theta_0)\Big]}_{(I)} + \underbrace{\mathbb{E}[\beta_{\Delta_{p'}^\star}^2 p''(\theta_0+\varepsilon)] + 2\mathbb{E}[\beta_{\Delta_{p'}^\star}\varepsilon p''(\theta_0+\varepsilon)]}_{(II)}, \quad p''(X) = 1\left\{|X| \geq \gamma_E^*\right\}.$$

The expression for (I) is obtained directly from Lemma A.7. For (II), we first note that we can write with simple rearrangement, after simplifying the main terms, and using the expression for $\beta_{\Delta_{p'}^\star} = \text{sign}(\theta_0+\varepsilon)\Big[\gamma_E^* + \frac{1}{c_d} - |\theta_0+\varepsilon|\Big]$

$$\mathbb{E}[\beta_{\Delta_{p'}^\star}\varepsilon] = c_0\mathbb{E}[\text{sign}(\theta_0+\varepsilon)\varepsilon||\theta_0+\varepsilon||]p''(|\theta_0+\varepsilon|)$$

for a finite constant c_0. It is easy to show that because θ_0 is centered around zero and independent of ε, $\mathbb{E}[\text{sign}(\theta_0 + \varepsilon)\varepsilon||\theta_0 + \varepsilon|] = 0$. For the first term in (II), we can write since the absolute bias is bounded from above by $1/c_d^2$

$$(II) \leq \frac{1}{c_d^2}.$$

The final conclusion follows directly from the Taylor approximation to $\mathcal{L}_E^*$ from Lemma A.6 and rearrangement.