Online Appendix for: Which findings should be published?

Alexander Frankel[*] Maximilian Kasy[†]

December 5, 2020

C Extensions

C.1 Alternative objective: learning

Separate from any decision problem, the public might value more precise knowledge of the state of the world for its own sake. One natural way of measuring the precision of beliefs is by looking at the variance. We formalize a *learning* objective by supposing that the public seeks a publication rule that minimizes the expected variance of the

[*]University of Chicago Booth School of Business. Address: 5807 South Woodlawn Avenue Chicago, IL 60637. E-Mail: afrankel@chicagobooth.edu.

[†]Department of Economics, Oxford University. Address: Manor Road Building, Manor Road, Oxford OX1 3UQ, United Kingdom. E-Mail: maximilian.kasy@economics.ox.ac.uk.

posterior beliefs π_1. Formally, under the learning objective we replace the earlier "relevance" welfare function $W(D, a, \theta)$ from Equation (1) with

$$W(D, \pi_1) = -\operatorname{Var}_{\theta \sim \pi_1}[\theta] - Dc, \tag{19}$$

where $c > 0$ continues to represent the social opportunity cost of publication. The *learning-optimal* publication rule p is the one which maximizes the ex-ante expectation of (19).

Suppose that the public uses Bayesian updating. There is then a clear connection between learning and relevance. The posterior expectation of the relevance utility under a quadratic loss utility function $-(a-\theta)^2$ – with the public choosing an action equal to its expectation of the state – is minus the posterior variance. That is exactly the learning welfare. So the learning-optimal policy is identical to the policy that maximizes the quadratic loss relevance objective under Bayesian updating, regardless of assumptions about signals or priors. In order to maximize learning and minimize uncertainty over the state of the world, then, it remains optimal to publish only those studies which induce extreme posteriors. This gives an alternative interpretation of some previous results that were motivated by decision problems.

C.2 Alternative objective: accuracy

Under an *accuracy* objective, a journal seeks to publish point estimates X that are as close as possible to the true state of the world θ. These estimates can be thought of as the ones that would be the most "replicable" by future studies. Letting $\Theta = \mathcal{X} = \mathbb{R}$, we formalize our accuracy objective by replacing welfare from (1) with

$$W(D, \theta, X) = D \cdot (-(X-\theta)^2 + b), \tag{20}$$

where $b > 0$ indicates the shadow benefit of publication; if no study arrives, welfare is normalized to zero. For simplicity, we now assume a quadratic loss from publishing values of X further from θ. (We consider a generalized loss function below.)

If the goal is to publish accurate results, a non-selective rule will do better than one that publishes only extreme findings. But, as we show, a different kind of selective rule can do even better. Let the *accuracy-optimal* publication rule be the one maximizing the ex-ante expectation of this welfare function.

Under the accuracy objective, publication depends only on the belief $\pi_1^{(X,S)}$. The accuracy-optimal rule publishes a study $(X, S) = (x, s)$ if the interim expected welfare from (20) is greater than 0, i.e., if

$$\mathbb{E}_{\theta \sim \pi_1^{(x,s)}}[(x-\theta)^2] \leq b. \tag{21}$$

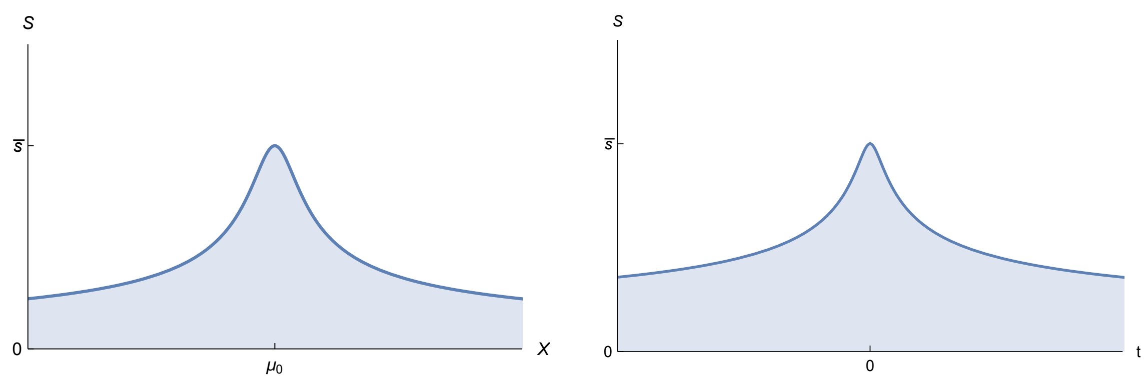
We can explicitly solve for this rule when there are normal priors: publish if $(X - \mu_0)^2 \leq \left(1 + \frac{\sigma_0^2}{S^2}\right)\left(b + b\frac{\sigma_0^2}{S^2} - \sigma_0^2\right)$.[1] At any standard error S, it is accuracy-

[1] As a first step to deriving this expression, rewrite (21) as $\operatorname{Var}_{\theta \sim \pi_1^{(x,s)}}[\theta] + (x - \mathbb{E}_{\theta \sim \pi_1^{(x,s)}}[\theta])^2 \leq b$. Then plug in the variance and expectation from (4) to derive the publication rule above.

optimal to publish studies with the point estimate X in a symmetric interval about μ_0; see Figure 1. (At sufficiently high standard errors, it may be the case that no studies are published.)

In other words, the accuracy-optimal publication rule has the opposite form as the publication rule maximizing quadratic loss relevance: at a given standard error, it publishes moderate findings and does not publish extreme ones. By the same token, publishing only extreme findings at a given standard error would minimize accuracy. This is because point estimates closer to the prior mean are thought (under the interim belief) to be closer to the true state. For intuition, recall that the distance of the point estimate from the interim mean, $X - \mu_1^{(X,S)}$, is linear in the distance of the point estimate from the prior mean, $X - \mu_0$. Of course, the accuracy-optimal publication rule is still partially aligned with the earlier (relevance-)optimal rules in that it publishes a larger range of point estimates when standard errors are smaller.

Figure 1: Accuracy-optimal publication region (shaded) for quadratic distance, normal prior.



(a) In terms of the point estimate X and the standard error S.

(b) In terms of the t-statistic $t = (X - \mu_0)/S$ and the standard error S.

If $b < \sigma_0^2$, as pictured, then no studies are published for $S > \bar{s}$, with $\bar{s} = \frac{\sigma_0\sqrt{b}}{\sqrt{\sigma_0^2 - b}}$. If instead $b \geq \sigma_0^2$, then an interval of X containing $[\mu_0 - (b - \sigma_0^2), \mu_0 + (b - \sigma_0^2)]$ would be published for any S.

Just as the relevance-optimal rule is bad for accuracy, so too is the accuracy-optimal rule bad for relevance. For a fixed standard error and a fixed share of studies to be published, the rule of publishing only moderate point estimates would actually minimize quadratic loss utility – and would therefore also be the worst for the learning objective.[2] A non-selective publication rule would be intermediate on both quadratic loss relevance and on accuracy.

[2] As described in Appendix A, we solved for the rule that maximized quadratic loss utility for Bayesian updating by first showing that the problem was equivalent to $\max_p \max_{a^0} EW(p, a^0)$; rearranging the order of maximization let us conclude that the globally optimal p was also interim-optimal given a^0. To solve for the policy that minimizes quadratic loss utility (at a fixed and

Without giving an explicit characterization, the same qualitative result of publishing moderate results to maximize accuracy would hold if we were to generalize the accuracy objective (20) beyond a quadratic cost of distance. Consider a generalized accuracy objective of

$$W(D, \theta, X) = D \cdot (-\delta((X - \theta)^2) + b), \tag{20'}$$

for a strictly increasing function $\delta(\cdot)$. (An arbitrary increasing function of $(X - \theta)^2$ is equivalent to an arbitrary increasing function of $|X - \theta|$.) One can establish that under normal priors, the generalized accuracy-optimal policy maximizing (20′) takes the same qualitative form as that maximizing (20): at a given standard error, either point estimates in a symmetric interval around μ_0 are published, or no point estimates are published.

Proposition 1. *Let there be normal priors. The publication rule maximizing the generalized accuracy objective* (20′) *takes the following form: at* $S = s$*, either no studies* (X, s) *are published, or there exists* k *such that a study* (X, s) *is published if and only if* $(X - \mu_0)^2 \leq k$*.*

C.3 A model with researcher incentives

Thus far, we have taken submissions to the journal to be exogenous. In reality submissions come about from a sequence of decisions by researchers: which topics to work on, what designs S to choose, and which findings X to actually write up and submit. In solving for an optimal journal publication rule, one ought to take into account the researchers' endogenous response to the incentives provided. To illustrate, this section presents a stylized *model with incentives* that explores a publication-motivated researcher's choices of whether to conduct a study and how to design that study.

Our analysis here complements some other recent theoretical investigations of how researcher or experimenter design choices may respond to incentives. In our example, the researcher's type will be commonly known and the design of a (submitted) study will be publicly observable, as in Henry and Ottaviani (2017) or the main analysis of McClellan (2017). Tetenov (2016) and Yoder (2018) study how a principal can screen across heterogeneous experimenters with privately known types. Libgober (2015) considers a setting in which study findings are observable, but the study design that led to a finding may be obscured.

Set-up. There is a single researcher who takes a research topic as given. There is a common prior $\theta \sim \pi_0$ shared by all parties: the researcher, the journal, and the public.

commonly known standard error), one solves $\min_p \max_{a^0} EW(p, a^0)$. By a minimax theorem, one can rearrange the order of minimization and maximization and conclude that the globally pessimal p is also interim-pessimal given a^0, and the interim-pessimal policy is to publish moderate results.

The timing of the game is as follows. First, the journal publicly commits to a publication rule p for studies on this topic. Then the researcher chooses whether to conduct a study and, if so, what study design S to use; the researcher will submit the results of any study to a journal. Then the game proceeds as in Section 2. If a study (X, S) is submitted it is published with probability $p(X, S)$, and finally the public updates its belief and takes a policy action. The key distinction from the original model is that the study submission probability q and the distribution of study designs F_S are now endogenous to the publication rule p.

To keep the analysis simple, we will restrict attention to naive updating. We will also continue to focus on a normal signal structure, with $S \in \mathbb{R}_{++}$ and $X|\theta, S \sim \mathcal{N}(\theta, S^2)$.

The researcher's problem. The researcher observes the publication rule p and then decides whether to conduct a study. If she does conduct a study then she chooses its standard error $S \in (0, \infty)$.

Normalize the researcher's outside option payoff from not conducting a study to 0. If a study is conducted, the researcher values its publication, but pays a cost that depends on the precision of the study. Specifically, the researcher gets a benefit of 1 for getting a study published, independently of the study's results. The researcher pays a cost $\kappa(S)$ for conducting a study with standard error S, with $\kappa : (0, \infty) \to \mathbb{R}_+$. (Assumptions such as $\kappa'(S) < 0$ would be natural – the researcher pays more for an experiment with a larger sample size, say – but we do not actually need to impose any conditions on the cost function for the results that follow.) So the researcher's ultimate payoff if she conducts a study with standard error S and publication outcome D is

$$D - \kappa(S).$$

Denote the researcher's expected payoff from conducting a study with standard error $S = s$, given journal publication rule p, by $V(s, p)$:

$$V(s, p) = \mathbb{E}_{\theta \sim \pi_0, X \sim \mathcal{N}(\theta, s^2)}[p(X, s)] - \kappa(s).$$

The researcher's participation constraint for being willing to conduct a study is

$$\max_{s \in (0,\infty)} V(s, p) \geq 0, \tag{P}$$

where we assume that the maximum is attained. Conditional on conducting a study, the researcher's choice of standard error S is determined by the incentive compatibility condition

$$S \in \arg \max_{s \in (0,\infty)} V(s, p). \tag{IC}$$

As before, we will assume that an argmax exists for any relevant p, without giving explicit conditions on primitives to guarantee that this will be the case.

The journal's problem. Let the journal maximize the expectation of welfare W given by the policy payoff minus any cost of publication:

$$W = U(a, \theta) - Dc.$$

That is, we suppose that the journal does not place any weight on the researcher's utility. Furthermore, assume that the public updates naively, so that the public's default action is fixed at $a^0 = a^*(\pi_0)$.

The journal's objective function takes the same form as in the original model, with the key distinction that the arrival of studies is no longer exogenous to the publication rule p. First, the study submission probability q depends on p: $q = 1$ if the participation constraint (P) is satisfied, and $q = 0$ otherwise. Second, conditional on participation, the standard error S depends on p through the incentive compatibility condition (IC). As is standard, assume that the researcher resolves indifferences in favor of the journal's preferences. The journal's problem is to choose an *incentive-optimal* publication rule p that maximizes expected welfare subject to these endogenous responses.

Observe that, conditional on the arrival of a study, the journal's gross interim benefit of publication is unchanged from its earlier definition in (6). A study that induces a journal interim belief of $\pi_1^{(X,S)}$ when the public's default action is $a^0 = a^*(\pi_0)$ yields gross interim benefit of $\Delta(\pi_1^{(X,S)}, a^*(\pi_0))$.

In the original model with exogenous study submission, the journal's optimal policy was given by the interim-optimal publication rule in which a study is published if and only if $\Delta(\pi_1^{(X,S)}, a^*(\pi_0)) \geq c$; indicate this interim-optimal publication rule by $p^{I(a^*(\pi_0))}$. Let us impose the assumption that the researcher would in fact be willing to participate if the journal were to use the publication rule $p^{I(a^*(\pi_0))}$ and would submit a study with $S = s^{\text{int}}$. This assumption will simplify both the solution and the exposition of our results.

Assumption 1. *The participation constraint* (P) *is satisfied under the interim-optimal publication rule* $p = p^{I(a^*(\pi_0))}$. *Let* $s^{\text{int}} \in \arg\max_s V\left(s; p^{I(a^*(\pi_0))}\right)$ *be the researcher's choice of study design in response to the interim-optimal publication rule.*

Characterizing the optimal publication rule.

Proposition 2. *Consider the model with incentives under naive updating, and suppose that Assumption 1 holds. Then there exist* $\overline{s} \leq s^{\text{int}}$, $\lambda \geq 0$, *and* $\rho \in [0, 1]$ *such that the following rule* p *is incentive-optimal:*

$$p(X,S) = \begin{cases} 1 & \textit{if } S = \overline{s} \textit{ and } \Delta(\pi_1^{(X,S)}, a^*(\pi_0)) > c - \lambda, \\ & \quad \textit{or if } S < \overline{s} \textit{ and } \Delta(\pi_1^{(X,S)}, a^*(\pi_0)) \geq c \\ \rho & \textit{if } S = \overline{s} \textit{ and } \Delta(\pi_1^{(X,S)}, a^*(\pi_0)) = c - \lambda \\ 0 & \textit{otherwise} \end{cases}.$$

Given this rule, the researcher chooses to conduct a study with $S = \overline{s}$.

The form of the optimal rule – at least at the chosen study design $S = \overline{s}$ – is very similar to the interim-optimal rule that was used in the model without incentives. A study is published if the gross interim benefit is sufficiently high.

However, the journal distorts publication from the interim-optimal rule in two ways. First, the journal does not publish any studies with standard error $S > \overline{s}$. The researcher is therefore induced to invest additional resources into the precision of studies and to reduce S from s^{int} to $\overline{s}$. Second, at $S = \overline{s}$ the journal relaxes the interim benefit threshold for publication from c to $c - \lambda$ in order to encourage researcher participation. Without that relaxation, a researcher might decide that a study at $S = \overline{s}$ would be too costly to conduct given its low likelihood of being published. (While in equilibrium the researcher never chooses $S < \overline{s}$, the journal has no reason to distort the publication rule at those more precise designs.)

In the original model without incentives, a journal which internalized all costs and benefits of publication would not need commitment power: ex-ante payoffs were maximized by publishing according to what was interim-optimal after receiving a study. Having added researcher incentives, the two distortions now require two forms of journal commitment. The journal commits not to publish imprecise studies, even if such a study was conducted and turned out to have extremely striking results. This commitment is never actually tested on the equilibrium path, though – imprecise studies are not conducted. The journal also commits to publish studies with weak findings when they have the appropriate precision. This second form of commitment is tested, as these studies are submitted (and published) in equilibrium.

One key simplification of this model of incentives is the assumption that there is no heterogeneity across researchers. This fact guarantees that researchers would always choose to conduct a study with a single standard error, known in advance. In a richer model, we would expect publication rules to reward more precise studies with higher publication probabilities in a more continuous manner than what we found here.

C.4 Imperfectly observed study designs

In determining whether to publish a study, a journal cares about the study's true information content. It may not be enough to treat the *reported* standard error as the variable S in our model of normal signals. As previously discussed, one concern is external validity: the parameter being estimated in the study may only be a proxy for the policy parameter of interest. Another concern is that the study may be internally flawed: a study with a misspecified model or an unconvincing identification strategy may report a very small standard error without actually being close to the truth.

When the study design is imperfectly observed, the point estimate can itself be informative as to the study's precision. To be concrete, assume that there are normal priors with mean normalized to 0 and there are normal signals, so that $\theta \sim \mathcal{N}(0, \sigma_0^2)$ and $X \sim \mathcal{N}(\theta, S^2)$. But now assume that the realization of $S \sim F_S$ is *unobserved* by the journal and the public. As noted in Subramanyam (1996), observing a point estimate with a larger magnitude $|X|$ leads to higher beliefs on the unobserved noise S. In our application, a small point estimate would suggest that the study design was precise, while a large point estimate would be suggestive of some hidden noise. The extreme realization might be attributed to a violation of the identifying assumptions, to a coding error, or to some other unseen flaw.

Continuing the example with X but not S observed by the journal and public, and with $\theta \sim \mathcal{N}(0, \sigma_0^2)$ and $X \sim \mathcal{N}(\theta, S^2)$, suppose further that there is quadratic loss utility. The journal makes a publication decision based on the posterior mean of θ, now conditional on X but not S:

$$\mu_1^{(X)} = \mathbb{E}[\theta|X] = \mathbb{E}[\mathbb{E}[\theta|X,S]|X] = \mathbb{E}\left[\frac{\sigma_0^2}{S^2+\sigma_0^2}|X\right] \cdot X.$$

The journal wants to publish if the interim benefit $(\mu_1^{(X)} - \mu_1^0)^2$ exceeds the publication cost c. A higher belief on S due to a larger point estimate $|X|$ translates into a lower weight $\mathbb{E}\left[\frac{\sigma_0^2}{S^2+\sigma_0^2}|X\right]$ on the point estimate. Indeed, when the prior on S is sufficiently dispersed, $\mathbb{E}\left[\frac{\sigma_0^2}{S^2+\sigma_0^2}|X\right]$ can decrease fast enough that $\mathbb{E}[\theta|X]$ is nonmonotonic and falls to 0 as X goes to infinity. (In addition to Subramanyam (1996), see discussion of this issue in Dawid (1973), O'Hagan (1979), and Harbaugh et al. (2016).) An intermediate point estimate would therefore move an observer's mean belief more than a very large, "implausible," point estimate would. Let us restate that our results in Section 3 support publishing "extreme results" in the sense of results that *lead to extreme beliefs*. If extreme signal realizations are written off as implausible, then they would not lead to extreme beliefs and thus should not be published.

A related possibility is that the study design S, capturing the true informational content of the study's findings, is better observed by the journal than by the public. After all, the journal editor and referees are experts who are charged with carefully evaluating the quality of a paper; a policymaker reading the study might not have this expertise. Consider a model where the journal observes (X, S) when making a publication decision, while if a paper is published the public sees only X. In such a model, the public can make an inference on the quality of the study design from the fact that the study was published. Publication implies that the journal had chosen to certify the study as clearing the bar of peer review. Suppose additionally that even unpublished studies are publicly available as working papers or preprints. In this case the only role of "publication" by a journal is certification or signaling value. A formal analysis of optimal publication rules in such an environment is an interesting topic for future research.

D Additional Results

D.1 Additional comparative statics

Proposition 3 presents additional comparative statics on the binary action publication rule solved for in Proposition 2. Recall that this result assumed a normal prior, and allowed for either Bayesian or naive updating.

Proposition 3. *Under the hypotheses and publication rule of Proposition 2,*

1. *The publication cutoff* $\left(1+\frac{S^2}{\sigma_0^2}\right)c-\frac{S^2}{\sigma_0^2}\mu_0$ *in terms of the point estimate is independent of the study arrival probability* q*. It is decreasing in the mean* μ_0*. It is larger when the standard error* S *is larger, the prior variance* σ_0^2 *is smaller, or the cost of publication* c *is larger.*
2. *The publication cutoff* $\left(\frac{1}{S}+\frac{S}{\sigma_0^2}\right)(c-\mu_0)$ *in terms of the t-statistic is nonmonotonic and convex in the standard error* S*: it has minimum at* $S=\sigma_0$ *and goes to infinity as* $S\to 0$ *or* $S\to\infty$*.*

Next we present additional comparative statics for the gross interim benefit of publishing null results in two-period model. Here we maintain the assumptions of Proposition 4, looking at naive updating, normal priors, and quadratic loss utility.

Proposition 4. *Under the hypotheses of Proposition 4, the gross interim benefit of publishing a result* (X_1,S_1) *with* $X_1=\mu_0$*, given by*

$$(1-\alpha)\frac{\sigma_0^8 s_2^4}{(\sigma_0^2+S_1^2)(\sigma_0^2+s_2^2)^2(\sigma_0^2 S_1^2+\sigma_0^2 s_2^2+S_1^2 s_2^2)},$$

is:

1. *decreasing in* α*, going to 0 as* $\alpha\to 1$*;*
2. *increasing in* σ_0*, going to 0 as* $\sigma_0\to 0$*;*
3. *decreasing in* S_1*, going to 0 as* $S_1\to\infty$*;*
4. *nonmonotonic and quasiconcave in* s_2*, approaching 0 as* $s_2\to 0$ *or* $s_2\to\infty$*.*

D.2 Size control for selective publication rules

Fix a z-score $z>0$. In this subsection we show how to construct selective publication rules for which the coverage probability of the confidence interval $[X-zS, X+zS]$ is equal to $\Phi[z]-\Phi[-z]$ for all θ. (Of course, as established by Theorem 2 part 4, such publication rules can not take the form of those in Theorem 1.) This exercise demonstrates that while non-selectivity is sufficient for confidence intervals to control size, it is not necessary.

Case of S=1: Normalizing $S = 1$, let the distribution of the finding X be given by $X \sim \mathcal{N}(\theta, 1)$ and the publication probability be given by $p(X)$. Then the coverage probability of a confidence interval of the form $[X - z, X + z]$ is given by

$$P(\theta \in [X - z, X + z]) = \frac{\int p(\theta + \epsilon)\mathbf{1}(\epsilon \in [-z, z])\varphi(\epsilon)d\epsilon}{\int p(\theta + \epsilon)\varphi(\epsilon)d\epsilon}.$$

This coverage probability is equal to its nominal level, $\Phi(z) - \Phi(-z)$, for all θ, if and only if

$$\int p(\theta + \epsilon) \left[\mathbf{1}(\epsilon \in [-z, z]) - (\Phi(z) - \Phi(-z))\right] \varphi(\epsilon)d\epsilon = 0 \text{ for all } \theta.$$

Taking the Fourier transform $\mathscr{F}$ of this expression, and recalling that the Fourier transform maps convolutions into products, the above expression is equivalent to the condition

$$\mathscr{F}\left(p(\cdot)\right) \cdot \mathscr{F}\left(\left[\mathbf{1}(\cdot \in [-z, z]) - (\Phi(z) - \Phi(-z))\right] \varphi(\cdot)\right) \equiv 0.$$

If the coverage probability is equal to its nominal level, we thus get that $\mathscr{F}(p(\cdot))$ has to equal zero everywhere except possibly at points where $\mathscr{F}\left(\left[\mathbf{1}(\cdot \in [-z, z]) - (\Phi(z) - \Phi(-z))\right] \varphi(\cdot)\right) = 0$. Reversely, by the Fourier inversion theorem,[3] this condition is also sufficient for the coverage probability to be equal to its nominal level.

The Fourier transform $\mathscr{F}\left(\left[\mathbf{1}(\cdot \in [-z, z]) - (\Phi(z) - \Phi(-z))\right] \varphi(\cdot)\right)$ is real-valued, even, and continuous. Let t^* be any zero of this Fourier transform. Then for any publication rule of the form $p(x) = r_0 + r_1 \cdot \sin(t^* \cdot x) + r_2 \cdot \cos(t^* \cdot x)$ we get that nominal size control is satisfied. (Of course, one must ensure that the publication probability is bounded between 0 and 1.) We can also take linear combinations of these functions over different roots t^*. These are the only publication rules with nominal size control.

While we cannot obtain analytic solutions, at any z we can numerically solve for such roots. For instance, for $z = 1.96$, solutions include $t^* \simeq 2.11045$, 3.49544, etc. So under either of the publication rules $p(x) = .5 + .5\cos(2.11045x)$ or $p(x) = .5 + .5\cos(3.49544x)$, for example, the probability of $\theta \in [X - 1.96, X + 1.96]$ conditional on publication would be 95% at all θ.

General case: Fixing z, suppose that $p(x)$ is some publication rule that satisfies nominal coverage for $S = 1$. Then $p(x, s) = p(x/s)$ achieves nominal coverage for $S = s$.

D.3 Two-period model with binary actions

Consider the two-period model with normal priors and naive updating. Proposition 4 in Section 5 presented the gross interim benefit of publication – and therefore the

[3] `https://en.wikipedia.org/wiki/Fourier_inversion_theorem`

optimal publication rule – for that setting under quadratic loss utility. Here, we will illustrate how some conclusions can change under binary action utility. We focus on characterizing how the interim benefit of publication varies as a function of the point estimate of the first-period study, X_1.

First, recall the quadratic loss analysis. With quadratic loss utility, the benefit of publishing towards the $t = 1$ action payoff – that is, the expected increase in $\alpha U(a_1, \theta)$ – is quadratic in $(X_1 - \mu_0)$, giving a symmetric benefit of publishing more extreme results in either direction. The benefit of publishing towards the $t = 2$ action payoff – the expected increase in $(1 - \alpha)U(a_2, \theta)$ – has one term that is quadratic in $(X_1 - \mu_0)$ and another term that is positive and constant in X_1. There is a benefit of publishing any result, including a null result with $X_1 = \mu_0$, and an additional benefit of publishing more extreme results. These disaggregated benefits are illustrated in panel (a) of Figure 2.

Now consider the model with binary action utility. The public's optimal action is $a = 0$ when its posterior mean is negative and $a = 1$ when its posterior mean is positive. Assume that $\mu_0 < 0$, and recall that we consider the case of naive updating, so the default action at $t = 1$ under nonpublication is $a = 0$. In that case the benefit towards the $t = 1$ payoff is $\alpha\mu_1^{(X_1,S_1)}$ if $\mu_1^{(X_1,S_1)} > 0$ and is 0 otherwise.[4] Since $\mu_1^{(X_1,S_1)}$ increases linearly with X_1, the benefit is zero at every X_1 from minus infinity through some positive number, and it increases linearly for larger X_1. See the blue curve in panel (b) of Figure 2.

Conditional on (X_1, S_1) and on X_2, the realized benefit of publication towards the $t = 2$ payoff is $(1 - \alpha)|\mu_2^{(X_1,S_1),(X_2)}|$ if $\mu_2^{(X_1,S_1),(X_2)}$ and $\mu_2^{0,(X_2)}$ are of different signs, and is zero otherwise. The publication decision is made at $t = 1$, and so the benefit is evaluated by taking expectation over X_2 (under the $t = 1$ interim beliefs $\pi_1^{(X_1,S_1)}$). See the orange curve in panel (b) of Figure 2 for an illustration of this expected $t = 2$ benefit. As we see, this $t = 2$ benefit is somewhat subtle.

The first thing to note is that the expected $t = 2$ payoff is strictly positive everywhere except $X_1 = 0$. The $t = 2$ benefit of publishing a result with $X_1 = 0$ is zero (as is the $t = 1$ benefit) because a study reporting $X_1 = 0$ never changes the period 2 action. The action depends on the sign of the mean, and a study with $X_1 = 0$ moves the posterior mean closer to zero without changing the sign.

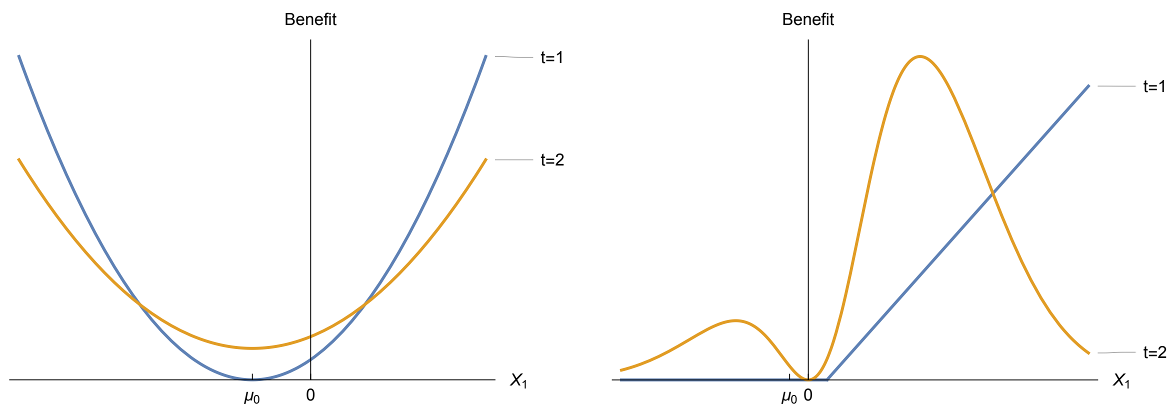
Moving away from $X_1 = 0$, there is a positive $t = 2$ benefit of publishing a result X_1 with an intermediate positive or negative value. Publishing a positive finding avoids the public's mistake of taking the action $a = 0$, in accord with its priors, when the unpublished period-1 study would actually indicate that the state is positive. Publishing a negative finding avoids the public's mistake of taking $a = 1$ after a positive finding in the second period, when the period-1 study would have indicated

[4] We follow the notational convention of the proof of Proposition 4 here, in which $\mu_1^{(X_1,S_1)}$ is the period 1 mean belief conditional on observing the period 1 study; $\mu_2^{(X_1,S_1),(X_2)}$ is the period 2 mean conditional on observing both studies; and $\mu_2^{0,(X_2)}$ is the period 2 mean after observing the second study if the first was not published.

a negative state. Figure 2 shows that these costs are asymmetric (a conclusion we see in other numerical examples): there is a larger cost of failing to publish a study with a positive result, one that goes against the public's prior.

Finally, as X_1 gets more extreme in either direction, the $t = 2$ payoff benefit approaches zero. This is because "$t = 2$" is defined as the time after some additional information has arrived.[5] And an extreme X_1 is suggestive of an extreme state, meaning that the period 2 signal is very likely to reveal whether the state is positive or negative. For instance, if X_1 has a very large positive value, then we expect X_2 to have a very large positive value as well. So publishing this study would give a $t = 1$ benefit by moving the first period action from $a_1 = 0$ to $a_1 = 1$. But the public will take $a_2 = 1$ in the second period regardless of whether the first period study is published.

Figure 2: Dynamic interim payoffs



(a) Quadratic loss utility

(b) Binary action utility

For both examples, we set $S_1 = 2$, $\sigma_0 = 2$, $\mu_0 = -1$, and $s_2 = 2$. The relative weight coefficient on the first period, α, is chosen to make the curves of similar scale as graphed; increasing α scales up the $t = 1$ benefit relative to that at $t = 2$. For quadratic loss utility, we have chosen $\alpha = .3$, with X_1 ranging from -5 to 3. For binary action utility, we have chosen $\alpha = .05$, with X_1 ranging from -10 to 15.

E Proofs for Appendix results

E.1 Proofs for Appendix A

Proof of Lemma 2. Follows from arguments in the text of Appendix A. □

[5]If there is a longer expected wait before new studies arrive and actions are updated, that corresponds in our model to a larger weight α on the $t = 1$ payoff.

Proof of Lemma 3. Follows from arguments in the text of Appendix A. $\square$

Proof of Proposition 5. By Lemma 3 part 1, it suffices to show that $a = a^*(\pi_1^{0,p^{I(a)}})$ is uniquely solved by $a = \mu_0$ – in other words, that $a^0 = \mu_0$ is the unique fixed point when we map default actions to interim optimal publication rules, and then map publication rules back to default actions.

Conditional on a study (X, S) arriving when the default action is a^0, the journal will not publish if $(\mu_1^{(X,S)} - a^0)^2 < c$, i.e., if $\mu_1^{(X,S)}$ lies in the interval $(a^0 - \sqrt{c}, a^0 + \sqrt{c})$. Let $\bar{\mu}(a^0)$ indicate $\mathbb{E}[\theta|\mu_1^{(X,S)} \in (a^0 - \sqrt{c}, a^0 + \sqrt{c})]$, the expected state conditional on a study arriving and not being published. If this expectation is undefined due to the event $\mu_1^{(X,S)} \in (a^0 - \sqrt{c}, a^0 + \sqrt{c})$ occurring with zero probability, let $\bar{\mu}(a^0) = \mu_0$.

The mean of the default belief – and therefore the implied default action – conditional on nonpublication will be a convex combination of $\bar{\mu}(a^0)$ (with weight q) and μ_0 (weight $1 - q$). Therefore, to show that $a^0 = \mu_0$ is the unique fixed point, it is sufficient to show the following three items: (i) for $a^0 = \mu_0$, it holds that $\bar{\mu}(a^0) = a^0$; (ii) for any $a^0 < \mu_0$, it holds that $\bar{\mu}(a^0) > a^0$; and (iii) for any $a^0 > \mu_0$, it holds that $\bar{\mu}(a^0) < a^0$. (If we had assumed $q < 1$ then it would be sufficient to show (ii) and (iii) with weak inequalities.)

Item (i) follows from the fact that $\mu_1^{(X,S)}$ is symmetric about μ_0, and therefore it remains symmetric when this random variable is truncated outside of the interval $(\mu_0 - \sqrt{c}, \mu_0 + \sqrt{c})$. The proofs of items (ii) and (iii) will be identical to each other, up to the direction of inequalities, so let us focus on proving (ii). Fix $a^0 < \mu_0$. First, if there is a zero probability that $\mu_1^{(X,S)} \in (a^0 - \sqrt{c}, a^0 + \sqrt{c})$, then $\bar{\mu}(a^0) = \mu_0 > a^0$ and we are done. Otherwise, notice that symmetry about μ_0 combined with single-peakedness means that the pdf of $\mu_1^{(X,S)}$ is larger at $a^0 + k$ than at $a^0 - k$ for any $k > 0$, with the inequality being strict for any ϵ such that either pdf value is nonzero. Hence the mean of $\mu_1^{(X,S)}$ conditional on being in the interval $(a^0 - \sqrt{c}, a^0 + \sqrt{c})$ is strictly above the midpoint a^0. That completes the proof of item (ii). $\square$

Proof of Proposition 6. By Lemma 3 part 2, it suffices to show that the payoff under default action $a^0 = 0$ is higher than under default action $a^0 = 1$, i.e., that $EW\left(p^{I(0)}, 0\right) \geq EW\left(p^{I(1)}, 1\right)$. The interim optimal publication rule $p^{I(a^0)}$ is as follows: for $a^0 = 0$ publish if $\mu_1^{(X,S)} \geq c$, and for $a^0 = 1$ publish if $\mu_1^{(X,S)} \leq -c$. Expanding out $EW(p, a^0)$ from (5) for each possible value of a^0,

$$EW(p^{I(0)}, 0) = q\mathbb{E}\left[\begin{cases} \mu_1^{(X,S)} - c & \text{if } \mu_1^{(X,S)} \geq c \\ 0 & \text{if } \mu_1^{(X,S)} < c \end{cases}\right]$$

$$EW(p^{I(1)}, 1) = q\mathbb{E}\left[\begin{cases} \mu_1^{(X,S)} & \text{if } \mu_1^{(X,S)} > -c \\ -c & \text{if } \mu_1^{(X,S)} \leq -c \end{cases}\right] + (1 - q)\mu_0.$$

Taking the difference,

$$EW(p^{I(0)},0) - EW(p^{I(1)},1) = q\mathbb{E}\left[\begin{cases} -c & \text{if } \mu_1^{(X,S)} \geq c \\ -\mu_1^{(X,S)} & \text{if } \mu_1^{(X,S)} \in (-c,c) \\ c & \text{if } \mu_1^{(X,S)} \leq -c \end{cases}\right] - (1-q)\mu_0. \tag{22}$$

We seek to show that this difference is nonnegative. Since $\mu_0 \leq 0$ by assumption, it is sufficient to show that the expectation term is nonnegative.

To show that the expectation term is nonnegative, first define a weakly increasing function $l : \mathbb{R} \to \mathbb{R}_+$ as follows:

$$l(k) = \begin{cases} 0 & \text{if } k \leq 0 \\ k & \text{if } k \in (0,c) \\ c & \text{if } k > c \end{cases}.$$

The expectation term in (22) can be rewritten as $\mathbb{E}[l(-\mu_1^{(X,S)})] - \mathbb{E}[l(\mu_1^{(X,S)})]$, and so it is sufficient to show that this difference is nonnegative.

Next, observe that the distribution of $-\mu_1^{(X,S)}$ first order stochastically dominates that of $\mu_1^{(X,S)}$:

$$\begin{aligned} P(-\mu_1^{(X,S)} \leq k) &= 1 - P(\mu_1^{(X,S)} \leq -k) \\ &\leq 1 - P(\mu_1^{(X,S)} \geq k) \\ &= P(\mu_1^{(X,S)} \leq k), \end{aligned}$$

where the inequality comes from the assumption of $P(\mu_1^{(X,S)} \leq -k) \geq P(\mu_1^{(X,S)} \geq k)$. By FOSD, then, the expectation of $l(-\mu_1^{(X,S)})$ is weakly larger than the expectation of $l(\mu_1^{(X,S)})$, completing the proof. $\square$

E.2 Proofs for Appendix B

Proof of Lemma 4. Let $a' = a^*(\pi')$, $a'' = a^*(\pi'')$, and $a''' = a^*(\pi''')$. Moreover, recall that for any actions $\underline{a} \leq \overline{a}$ and any distributions $\underline{\pi} \leq_{FOSD} \overline{\pi}$, supermodularity implies that

$$\mathbb{E}_{\theta\sim\overline{\pi}}[U(\underline{a},\theta)] + \mathbb{E}_{\theta\sim\underline{\pi}}[U(\overline{a},\theta)] \leq \mathbb{E}_{\theta\sim\underline{\pi}}[U(\underline{a},\theta)] + \mathbb{E}_{\theta\sim\overline{\pi}}[U(\overline{a},\theta)]. \tag{23}$$

Now consider the two exhaustive cases of $a^0 \leq a''$ and $a^0 \geq a''$.

If $a^0 \leq a''$, then

$$\begin{aligned} \mathbb{E}_{\theta\sim\pi'''}[U(a^0,\theta)] + \mathbb{E}_{\theta\sim\pi''}[U(a'',\theta)] &\leq \mathbb{E}_{\theta\sim\pi''}[U(a^0,\theta)] + \mathbb{E}_{\theta\sim\pi'''}[U(a'',\theta)] \\ &\leq \mathbb{E}_{\theta\sim\pi''}[U(a^0,\theta)] + \mathbb{E}_{\theta\sim\pi'''}[U(a''',\theta)] \\ \Rightarrow\ \mathbb{E}_{\theta\sim\pi''}[U(a'',\theta)] - \mathbb{E}_{\theta\sim\pi''}[U(a^0,\theta)] &\leq \mathbb{E}_{\theta\sim\pi'''}[U(a''',\theta)] - \mathbb{E}_{\theta\sim\pi'''}[U(a^0,\theta)] \\ \Rightarrow \Delta(\pi'',a^0) &\leq \Delta(\pi''',a^0), \end{aligned}$$

where, on the first line, the first inequality follows from (23) and the second inequality follows from the fact that $a''' = a^*(\pi''')$. The second line then rearranges terms from the left-hand side and the right-hand side of the first line.

Alternatively, if $a^0 \geq a''$, then by a similar argument

$$\begin{aligned}
&\mathbb{E}_{\theta\sim\pi'}[U(a^0,\theta)] + \mathbb{E}_{\theta\sim\pi''}[U(a'',\theta)] \leq \mathbb{E}_{\theta\sim\pi''}[U(a^0,\theta)] + \mathbb{E}_{\theta\sim\pi'}[U(a'',\theta)] \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \leq \mathbb{E}_{\theta\sim\pi''}[U(a^0,\theta)] + \mathbb{E}_{\theta\sim\pi'}[U(a',\theta)] \\
\Rightarrow\ &\mathbb{E}_{\theta\sim\pi''}[U(a'',\theta)] - \mathbb{E}_{\theta\sim\pi''}[U(a^0,\theta)] \leq \mathbb{E}_{\theta\sim\pi'}[U(a',\theta)] - \mathbb{E}_{\theta\sim\pi'}[U(a^0,\theta)] \\
&\qquad\qquad\qquad\qquad\qquad \Rightarrow \Delta(\pi'',a^0) \leq \Delta(\pi',a^0). \qquad \square
\end{aligned}$$

Proof of Lemma 5. As stated, when publication is non-selective, the distribution of $X|\theta, S = s, D = 1$ is identical to the distribution $X|\theta, S = s$ for every s. Parts 1 and 4 follow immediately from that observation. Part 2 follows from the definition of non-selective publication: $p(x, s)$ constant in x implies that $\mathbb{E}[p(X, S)|\theta, S = s]$ is equal to that same constant. To show part 3, recall that the independence of S and θ implies that if $\mathbb{E}[p(X, S)|\theta, S = s]$ is constant for each s, then it is constant in expectation across S, and so $\mathbb{E}[p(X, S)|\theta]$ is constant as well. The result then follows from (3). $\square$

Proof of Lemma 6. First observe that

$$\begin{aligned}
&\mathbb{E}_{\theta\sim\pi}[(X_2^{(s_2)} - \theta)^2] = s_2^2 \\
\Rightarrow\ &\lim_{s_2\to 0} \mathbb{E}_{\theta\sim\pi}[(X_2^{(s_2)} - \theta)^2] = 0.
\end{aligned} \tag{24}$$

Next recall that for any s_2 and any realization $X_2^{(s_2)} = x$, the posterior mean of the updated belief, $m(x; \pi_1^I, s_2)$, minimizes the expected square distance to θ:

$$\begin{aligned}
&m(x;\pi,s_2) \in \arg\min_{g_{s_2}:\mathbb{R}\to\mathbb{R}} \mathbb{E}_{\theta\sim\pi}[(g_{s_2}(x) - \theta)^2 | X_2^{(s_2)} = x] \\
\Rightarrow\ &\mathbb{E}_{\theta\sim\pi}[(m(x;\pi,s_2) - \theta)^2 | X_2^{(s_2)} = x] \\
&\quad \leq \mathbb{E}_{\theta\sim\pi}[(g_{s_2}(x) - \theta)^2 | X_2^{(s_2)} = x]\ \forall g_{s_2}.
\end{aligned}$$

Since this inequality holds for each realization $X_2^{(s_2)} = x$, it also holds in expectation:

$$\mathbb{E}_{\theta\sim\pi}[(m(X_2^{(s_2)};\pi) - \theta)^2] \leq \mathbb{E}_{\theta\sim\pi}[(g_{s_2}(X_2^{(s_2)}) - \theta)^2]\ \forall g_{s_2}.$$

Plugging in $g_{s_2}(x)$ equal to the identity function x,

$$0 \leq \mathbb{E}_{\theta\sim\pi}[(m(X_2^{(s_2)};\pi,s_2) - \theta)^2] \leq \mathbb{E}_{\theta\sim\pi}[(X_2^{(s_2)} - \theta)^2].$$

Taking the limit as $s_2 \to 0$ as in (24), the right-hand side of the above expression converges to 0, and hence

$$\lim_{s_2\to 0} \mathbb{E}_{\theta\sim\pi}[(m(X_2^{(s_2)};\pi,s_2) - \theta)^2] \to 0. \tag{25}$$

So we see that $m(X_2^{(s_2)};\pi,s_2)$ and $X_2^{(s_2)}$ both converge to θ in mean-square as $s_2 \to 0$. We can conclude that $m(X_2^{(s_2)};\pi,s_2)$ converges to $X_2^{(s_2)}$ in mean-square, and hence we have proven our result, if $m(X_2^{(s_2)};\pi,s_2)$, $X_2^{(s_2)}$, and θ are all square-integrable under $\theta \sim \pi$. In turn it suffices to show that these random variables all have a finite mean and a variance. By assumption, the mean and variance of θ under π are finite. Then $X_2^{(s_2)}$ and $m(X_2^{(s_2)};\pi,s_2)$ also share the mean of θ under π for all s_2. The variance of X_2 is given by $\mathrm{Var}_{\theta\sim\pi}(\theta)+s_2^2$. Finally, the variance of $m(X_2;\pi,s_2)$ is bounded above by $\mathrm{Var}_{\theta\sim\pi}(\theta)$ by the Law of Total Variance: the variance of the posterior mean given some signal is bounded above by the variance of the prior. $\square$

Proof of Lemma 7. Applying a transformation with $\lambda = 1/s_2$, let $\hat{X}_2^{(\lambda)} = \lambda X_2^{(1/\lambda)}$; $X_2^{(\lambda)}$ is equal to the t-statistic $X_2^{(s_2)}/s_2$. That is, $\hat{X}_2^{(\lambda)}|\theta \sim \mathcal{N}(\lambda\theta, 1)$, where $\hat{X}_2^{(\lambda)}|\theta$ has pdf at $\hat{x}$ of $\varphi(\hat{x}-\lambda\theta)$. Correspondingly, let

$$\hat{m}(\hat{x};\pi,\lambda) = \mathbb{E}_{\theta\sim\pi}[\theta|\hat{X}_2^{(\lambda)} = \hat{x}]$$

be the public's period-2 expectation of θ given period-1 belief π followed by period-2 observation $\hat{X}_2^{(\lambda)} = \hat{x}$, i.e., given $X_2^{(1/\lambda)} = \hat{x}/\lambda$. This transformation will be convenient because as $s_2 \to \infty$ and $\lambda = 1/s_2 \to 0$, the variable $\hat{X}^{(\lambda)}|\theta$ approaches a standard normal, whereas $X_2^{(s_2)}|\theta$ approaches an improper distribution with infinite variance.

We seek to show that for any π with mean μ_1 that is bounded by Pareto tails with finite variance, it holds that

$$\lim_{\lambda\to 0} \mathbb{E}_{\theta\sim\pi}[(\hat{m}(\hat{X}^{(\lambda)};\pi,\lambda)-\mu_1)^2] = 0. \tag{26}$$

Writing the expectation from (26) out in integral form,

$$\mathbb{E}_{\theta\sim\pi}[(\hat{m}(\hat{X}^{(\lambda)};\pi,\lambda)-\mu_1)^2] = \int\int (\hat{m}(\hat{x};\pi,\lambda)-\mu_1)^2\varphi(\hat{x}-\lambda\theta)d\pi(\theta)d\hat{x}.$$

By Lebesgue's dominated convergence theorem, to show (26), it suffices to show (i) for all $\hat{x}$, $\lim_{\lambda\to 0}\int(\hat{m}(\hat{x};\pi,\lambda)-\mu_1)^2\varphi(\hat{x}-\lambda\theta)d\pi(\theta) = 0$; and (ii) there exists a "dominating" function $g:\mathbb{R}_+ \to \mathbb{R}_+$ that is Lebesgue-integrable, i.e., $\int g(\hat{x})d\hat{x}$ is finite, such that for λ sufficiently small, $\int(\hat{m}(\hat{x};\pi,\lambda)-\mu_1)^2\varphi(\hat{x}-\lambda\theta)d\pi(\theta) \le g(\hat{x})$ for all $\hat{x}$.

Step 1: *Show that for all $\hat{x}$, $\lim_{\lambda\to 0}\int(\hat{m}(\hat{x};\pi,\lambda)-\mu_1)^2\varphi(\hat{x}-\lambda\theta)d\pi(\theta) = 0$.*

It holds that

$$\begin{aligned}\int(\hat{m}(\hat{x};\pi,\lambda)-\mu_1)^2\varphi(\hat{x}-\lambda\theta)d\pi(\theta) &= (\hat{m}(\hat{x};\pi,\lambda)-\mu_1)^2\int\varphi(\hat{x}-\lambda\theta)d\pi(\theta)\\ &\le (\hat{m}(\hat{x};\pi,\lambda)-\mu_1)^2\int\varphi(0)d\pi(\theta)\\ &= (\hat{m}(\hat{x};\pi,\lambda)-\mu_1)^2\varphi(0).\end{aligned}$$

So to show the desired result that $\int(\hat{m}(\hat{x};\pi,\lambda)-\mu_1)^2\varphi(\hat{x}-\lambda\theta)d\pi(\theta)$ converges to 0 for all $\hat{x}$, it suffices to show that $(\hat{m}(\hat{x};\pi,\lambda)-\mu_1)^2$ converges to 0 for all $\hat{x}$. In turn, it suffices to show that $\hat{m}(\hat{x};\pi,\lambda)$ converges to μ_1 for any fixed $\hat{x}$. Writing $\hat{m}(\hat{x};\pi,\lambda)$ in integral form,

$$\hat{m}(\hat{x};\pi,\lambda)=\frac{\int\theta\varphi(\hat{x}-\lambda\theta)d\pi(\theta)}{\int\varphi(\hat{x}-\lambda\theta)d\pi(\theta)} \tag{27}$$

In the denominator of (27), for all θ, $\varphi(\hat{x}-\lambda\theta)\to\varphi(\hat{x})$ as $\lambda\to 0$. Moreover, $\varphi(\hat{x}-\lambda\theta)\le\varphi(0)$ for all θ and λ, and $\int\varphi(0)d\pi(\theta)=\varphi(0)<\infty$. So $\varphi(0)$ is a dominating function for $\varphi(\hat{x}-\lambda\theta)$ that is integrable with respect to π_0, and hence by the dominated convergence theorem the denominator approaches $\int\varphi(\hat{x})d\pi(\theta)=\varphi(\hat{x})$.

In the numerator of (27), for all θ, $\theta\varphi(\hat{x}-\lambda\theta)\to\theta\varphi(\hat{x})$ as $\lambda\to 0$. Moreover, $|\theta\varphi(\hat{x}-\lambda\theta)|\le|\theta|\varphi(0)$ for all θ and λ, and $\int\theta\varphi(0)d\pi(\theta)=\varphi(0)\int|\theta|d\pi(\theta)<\infty$ because π has a finite mean. So $|\theta|\varphi(0)$ is a dominating function for $\theta\varphi(\hat{x}-\lambda\theta)$ that is integrable with respect to π, and hence by the dominated convergence theorem the numerator approaches $\int\theta\varphi(\hat{x})d\pi(\theta)=\mu_1\varphi(\hat{x})$.

Taking the ratio, we have that $\hat{m}(\hat{x};\pi,\lambda)$ converges to $\mu_1\varphi(\hat{x})/\varphi(\hat{x})=\mu_1$ as $\lambda\to 0$, completing this step.

Step 2: *Show that there exists a dominating function $g:\mathbb{R}_+\to\mathbb{R}_+$ that is Lebesgue-integrable, such that for λ sufficiently small, $\int(\hat{m}(\hat{x};\pi,\lambda)-\mu_1)^2\varphi(\hat{x}-\lambda\theta)d\pi(\theta)\le g(\hat{x})$ for all $\hat{x}$.*

First, observe that

$$\begin{aligned}
\int(\hat{m}(\hat{x};\pi,\lambda)-\mu_1)^2\varphi(\hat{x}-\lambda\theta)d\pi(\theta) &= (\hat{m}(\hat{x};\pi,\lambda)-\mu_1)^2\int\varphi(\hat{x}-\lambda\theta)d\pi(\theta)\\
&= \left(\frac{\int\theta\varphi(\hat{x}-\lambda\theta)d\pi(\theta)}{\int\varphi(\hat{x}-\lambda\theta)d\pi(\theta)}-\mu_1\right)^2\cdot\int\varphi(\hat{x}-\lambda\theta)d\pi(\theta)\\
&\le \frac{\int(\theta-\mu_1)^2\varphi(\hat{x}-\lambda\theta)d\pi(\theta)}{\int\varphi(\hat{x}-\lambda\theta)d\pi(\theta)}\cdot\int\varphi(\hat{x}-\lambda\theta)d\pi(\theta)\\
&= \int(\theta-\mu_1)^2\varphi(\hat{x}-\lambda\theta)d\pi(\theta) \qquad (28)
\end{aligned}$$

where the inequality in the third line follows from Jensen's inequality: $(\mathbb{E}[\theta|\hat{X}^{(\lambda)}=\hat{x}]-\mu_1)^2=(\mathbb{E}[\theta-\mu_1|\hat{X}^{(\lambda)}=\hat{x}])^2\le\mathbb{E}[(\theta-\mu_1)^2|\hat{X}^{(\lambda)}=\hat{x}]$.

So it suffices to find an integrable function g for which $g(\hat{x})$ is everywhere larger than (28) for all $\lambda\in(0,1]$.

- *Constructing g for $\hat{x}\in[-2K,2K]$.*
  The expression (28) is uniformly bounded above by $\int(\theta-\mu_1)^2\varphi(0)d\pi(\theta)=\varphi(0)\operatorname{Var}_{\theta\sim\pi}(\theta)$. So, let

  $$g(\hat{x})=\varphi(0)\operatorname{Var}_{\theta\sim\pi}(\theta)\text{ for }\hat{x}\in[-2K,2K].$$

  It holds that $\int_{-2K}^{2K}g(\hat{x})d\hat{x}=4K\varphi(0)\operatorname{Var}_{\theta\sim\pi}(\theta)<\infty$.

- *Constructing g for $\hat{x} > 2K$.*
  Expanding out (28), we have

$$\int (\theta-\mu_1)^2 \varphi(\hat{x}-\lambda\theta) d\pi(\theta) = \underbrace{\int_{-\infty}^{\frac{\hat{x}}{2\lambda}} (\theta-\mu_1)^2 \varphi(\hat{x}-\lambda\theta) d\pi(\theta)}_{A} + \underbrace{\int_{\frac{\hat{x}}{2\lambda}}^{\infty} (\theta-\mu_1)^2 \varphi(\hat{x}-\lambda\theta) d\pi(\theta)}_{B} \tag{29}$$

First let us bound the term labeled A in (29). For $\theta \leq \frac{\hat{x}}{2\lambda}$, it holds that $\hat{x}-\lambda\theta \geq \hat{x}/2$. Therefore, assuming further that $\hat{x} \geq 2K$ – and in particular that $\hat{x} \geq 0$ – it holds that $\varphi(\hat{x}-\lambda\theta) \leq \varphi(\hat{x}/2)$. Hence,

$$\begin{aligned}
\underbrace{\int_{-\infty}^{\frac{\hat{x}}{2\lambda}} (\theta-\mu_1)^2 \varphi(\hat{x}-\lambda\theta) d\pi(\theta)}_{A} &\leq \int_{-\infty}^{\frac{\hat{x}}{2\lambda}} (\theta-\mu_1)^2 \varphi(\hat{x}/2) d\pi(\theta) \\
&\leq \int_{-\infty}^{\infty} (\theta-\mu_1)^2 \varphi(\hat{x}/2) d\pi(\theta) \\
&= \varphi(\hat{x}/2)\,\mathrm{Var}_{\theta\sim\pi}(\theta).
\end{aligned}$$

Now we move to the term labeled B in (29). By the fact that π is bounded by Pareto tails with finite variance,

$$\begin{aligned}
\underbrace{\int_{\frac{\hat{x}}{2\lambda}}^{\infty} (\theta-\mu_1)^2 \varphi(\hat{x}-\lambda\theta) d\pi(\theta)}_{B} &\leq \int_{\frac{\hat{x}}{2\lambda}}^{\infty} C\theta^{-\gamma}(\theta-\mu_1)^2 \varphi(\hat{x}-\lambda\theta) d\theta \\
&\leq \int_{\frac{\hat{x}}{2\lambda}}^{\infty} C\theta^{-\gamma}(\theta+|\mu_1|)^2 \varphi(\hat{x}-\lambda\theta) d\theta \\
&\leq C\frac{(\frac{\hat{x}}{2\lambda}+|\mu_1|)^2}{(\frac{\hat{x}}{2\lambda})^\gamma} \int_{\frac{\hat{x}}{2\lambda}}^{\infty} \varphi(\hat{x}-\lambda\theta) d\theta \\
&= C\frac{(\frac{\hat{x}}{2\lambda}+|\mu_1|)^2}{(\frac{\hat{x}}{2\lambda})^\gamma} \frac{1}{\lambda}(1-\Phi(-\frac{\hat{x}}{2})) \\
&= 2^{\gamma-2} C \lambda^{\gamma-3} \frac{(\hat{x}+2\lambda|\mu_1|)^2}{\hat{x}^\gamma} \Phi(\frac{\hat{x}}{2}) \\
&\leq 2^{\gamma-2} C \frac{(\hat{x}+2|\mu_1|)^2}{\hat{x}^\gamma} \text{ for } \lambda \in (0,1]
\end{aligned}$$

The inequality in the third line follows because $\theta^{-\gamma}(\theta+|\mu_1|)^2$ is decreasing in θ over $\theta > 0$ for any $\gamma > 2$, so we increase the expression when we plug in the lowest value of θ, i.e., $\theta = \hat{x}/(2\lambda)$. The inequality in the last line follows because $\lambda^{\gamma-3}(\hat{x}+2\lambda|\mu_1|)^2$ is increasing in λ over $\lambda > 0$ for any $\gamma > 3$, so we increase the expression relative to $\lambda \leq 1$ when we plug in $\lambda = 1$; and we also increase the expression when we replace $\Phi(\frac{\hat{x}}{2})$ by 1. These two observations about increasing and decreasing functions can be straightforwardly confirmed by taking derivatives.[6]

Putting the bounds on terms A and B together, let

$$g(\hat{x}) = \varphi(\hat{x}/2)\,\mathrm{Var}_{\theta\sim\pi}(\theta) + 2^{\gamma-2}C\frac{(\hat{x}+2|\mu_1|)^2}{\hat{x}^\gamma} \text{ for } \hat{x} > 2K.$$

As established, $g(\hat{x})$ is larger than (28) for all $\lambda \leq 1$. Moreover, $\int_{2K}^{\infty} g(\hat{x})d\hat{x}$ is finite: the first term is an integral of a normal pdf, and the second term is an integral of an expression that decays to zero as $\hat{x}$ goes to infinity at a rate of $\hat{x}^{2-\gamma}$, with the exponent $2-\gamma < -1$.

- *Constructing g for $\hat{x} < -2K$.*

  This case proceeds symmetrically to the construction for $\hat{x} > 2K$, now taking

  $$g(\hat{x}) = \varphi(\hat{x}/2)\,\mathrm{Var}_{\theta\sim\pi}(\theta) + 2^{\gamma-2}C\frac{(|\hat{x}|+2|\mu_1|)^2}{|\hat{x}|^\gamma} \text{ for } \hat{x} < 2K.$$

  Just as with $\hat{x} > 2K$, when $\hat{x} < -2K$ we have that $g(\hat{x})$ is an upper bound for (28) when $\lambda \leq 1$, and $\int_{-\infty}^{-2K} g(\hat{x})d\hat{x}$ is finite.

We have now established that $g(\hat{x})$ is an upper bound for (28) for all $\lambda \leq 1$ and for all $\hat{x}$, and that $\int g(\hat{x})d\hat{x} < \infty$, concluding the proof. $\square$

Proof of Lemma 8. Define $f^I_{X_2^{(s_2)}}(x) = \frac{1}{s_2}\int \varphi(\frac{x-\theta}{s_2})d\pi_1^I(\theta)$ and $f^0_{X_2^{(s_2)}}(x) = \frac{1}{s_2}\int \varphi(\frac{x-\theta}{s_2})d\pi_1^0(\theta)$ to be the marginal densities of $X_2^{(s_2)}$ under the respective distributions on θ of π_1^I and π_1^0.

Step 1: *Show that there exists $C' > 0$ such that $\frac{f^I_{X_2^{(s_2)}}(x)}{f^0_{X_2^{(s_2)}}(x)} \leq C'$ for all s_2.*

First observe that

$$\frac{f^I_{X_2^{(s_2)}}(x)}{f^0_{X_2^{(s_2)}}(x)} = \frac{\int \varphi(\frac{x-\theta}{s_2})d\pi_1^I(\theta)}{\int \varphi(\frac{x-\theta}{s_2})d\pi_1^0(\theta)} = \frac{\int \varphi(\frac{x-\theta}{s_2})\frac{d\pi_1^I(\theta)}{d\pi_1^0(\theta)}d\pi_1^0(\theta)}{\int \varphi(\frac{x-\theta}{s_2})d\pi_1^0(\theta)} \leq \sup_\theta \frac{d\pi_1^I(\theta)}{d\pi_1^0(\theta)}.$$

Next, recall that $\pi_1^I = \pi_1^{(x_1,s_1)}$, which is a posterior belief on θ given prior $\theta\sim\pi_0$ and some fixed signal realization $(X_1,S_1) = (x_1,s_1)$. Hence

$$\frac{d\pi_1^I(\theta)}{d\pi_0(\theta)} = \frac{\varphi(\frac{x_1-\theta}{s_1})}{\int \varphi(\frac{x_1-\theta'}{s_1})d\pi_0(\theta')}$$

$$\Rightarrow \sup_\theta \frac{d\pi_1^I(\theta)}{d\pi_0(\theta)} \leq \frac{\varphi(0)}{\int \varphi(\frac{x_1-\theta'}{s_1})d\pi_0(\theta').}$$

[6]The derivative of $\theta^{-\gamma}(\theta+|\mu_1|)^2$ with respect to θ evaluates to $-\theta^{-(1+\gamma)}(\theta+|\mu_1|)(|\mu_1|\gamma+(\gamma-2)\theta) < 0$. The derivative of $\lambda^{\gamma-3}(\hat{x}+2\lambda|\mu_1|)^2$ with respect to λ evaluates to $\lambda^{\gamma-4}(\hat{x}+2|\mu_1|\lambda)((\gamma-3)\hat{x}+2|\mu_1|(\gamma-1)\lambda) > 0$.

Under naive updating, $\pi_1^0 = \pi_0$, and thus $\sup_\theta \frac{d\pi_1^I(\theta)}{d\pi_1^0(\theta)} = \sup_\theta \frac{d\pi_1^I(\theta)}{d\pi_0(\theta)}$, bounded by the finite constant $C' = \frac{\varphi(0)}{\int \varphi(\frac{x_1-\theta'}{s_1})d\pi_0(\theta')}$. (Recall that x_1 and s_1 are taken as constants here.) Under Bayesian updating with study arrival probability $q < 1$, (3) implies that $\frac{d\pi_0(\theta)}{d\pi_1^0(\theta)} \leq \frac{1}{1-q}$ for all θ, and therefore that $\sup_\theta \frac{d\pi_1^I(\theta)}{d\pi_1^0(\theta)} = \sup_\theta \frac{d\pi_1^I(\theta)}{d\pi_0(\theta)} \frac{d\pi_0(\theta)}{d\pi_1^0(\theta)} \leq \frac{1}{1-q} \sup_\theta \frac{d\pi_1^I(\theta)}{d\pi_0(\theta)}$. Hence for Bayesian updating we have a bound $C' = \frac{1}{1-q} \frac{\varphi(0)}{\int \varphi(\frac{x_1-\theta'}{s_1})d\pi_0(\theta')}$.

In either case C' gives an upper bound on $\frac{f^I_{X_2^{(s_2)}2}(x)}{f^0_{X_2^{(s_2)}}(x)}$.

Step 2: *Show that* $\mathbb{E}_{\theta\sim\pi_1^I}\left[y\left(X_2^{(s_2)}\right)\right] \leq C'\mathbb{E}_{\theta\sim\pi_1^0}\left[y\left(X_2^{(s_2)}\right)\right]$.

Rewriting expectations in integral form,

$$\begin{aligned}
\mathbb{E}_{\theta\sim\pi_1^I}\left[y\left(X_2^{(s_2)}\right)\right] &= \int y\left(X_2^{(s_2)}\right) f^I_{X_2^{(s_2)}}(x)dx \\
&= \int y\left(X_2^{(s_2)}\right) \frac{f^I_{X_2^{(s_2)}}(x)}{f^0_{X_2^{(s_2)}}(x)} f^0_{X_2^{(s_2)}}(x)dx \\
&\leq \int y\left(X_2^{(s_2)}\right) C' f^0_{X_2^{(s_2)}}(x)dx \text{ (by Step 1)} \\
&= C'\mathbb{E}_{\theta\sim\pi_1^0}\left[y\left(X_2^{(s_2)}\right)\right].
\end{aligned}$$

□

E.3 Proofs for Online Appendix C

E.3.1 Proofs for Online Appendix C.2

Proof of Proposition 1. Recall that under normal priors, the variance of $\pi_1^{(X,S)}$ is independent of X. So fix $S = s$, and without loss of generality normalize the variance of $\pi_1^{(X,s)}$ to 1. Then given $X = x$ and $\theta \sim \pi_1^{(x,s)}$, the distribution of a random variable $Y = (x-\theta)^2$ is a noncentral chi-squared distribution with noncentrality parameter λ (equal to $(x - \mathbb{E}_{\theta\sim\pi_1^{(x,s)}}[\theta])^2$) that increases in $(x-\mu_0)^2$. The variable Y has CDF over realizations y given by $1 - Q_{1/2}(\sqrt{\lambda}, \sqrt{y})$ for Q the Marcum Q-function.[7] By Sun et al. (2010) Theorem 1(a), $Q_{1/2}(\sqrt{\lambda}, \sqrt{y})$ strictly increases in its first term $\sqrt{\lambda}$, implying that the distribution of $(x-\theta)^2$ under $\pi_1^{(x,s)}$ increases in the sense of FOSD as $(x-\mu_0)^2$ increases. Hence $\mathbb{E}_{\theta\sim\pi_1^{(x,s)}}[\delta((x-\theta)^2)]$ increases in $(x-\mu_0)^2$. A study $(X,S) = (x,s)$ is published if and only if $\mathbb{E}_{\theta\sim\pi_1^{(x,s)}}[\delta((x-\theta)^2)] \leq b$, so at standard error $S = s$ studies are published only if $(X-\mu_0)^2$ is sufficiently small. □

[7] See Wikipedia for details: `https://en.wikipedia.org/wiki/Noncentral_chi-squared_distribution`.

E.3.2 Proofs for Online Appendix C.3

Proof of Proposition 2. We first state a lemma that does not depend on Assumption 1.

Lemma 1. *In searching for an incentive-optimal publication rule, it is without loss of generality to restrict to rules $p(X, S)$ satisfying*

$$p(X,S) = \begin{cases} 1 & \text{if } S = \overline{s} \text{ and } \Delta(\pi_1^{(X,S)}, a^*(\pi_0)) > c - \lambda, \\ & \quad \text{or if } S < \overline{s} \text{ and } \Delta(\pi_1^{(X,S)}, a^*(\pi_0)) \geq c \\ 0 & \text{if } S > \overline{s}, \\ & \quad \text{or if } S = \overline{s} \text{ and } \Delta(\pi_1^{(X,S)}, a^*(\pi_0)) < c - \lambda \\ & \quad \text{or if } S < \overline{s} \text{ and } \Delta(\pi_1^{(X,S)}, a^*(\pi_0)) > c \end{cases}$$

for some $\overline{s} \in (0, \infty)$ and λ in $\mathbb{R} \cup \{-\infty, \infty\}$ in which the researcher chooses $S = \overline{s}$ if she conducts a study.

It remains only to show that in the incentive-optimal contract of the form in Lemma 1, the researcher chooses to conduct a study; that $\overline{s} \leq s^{\text{int}}$; and that $\lambda \geq 0$.

The facts that the researcher conducts a study and that $\overline{s} \leq s^{\text{int}}$ both follow from Assumption 1.

First, Assumption 1 guarantees that the journal prefers to follow the interim-optimal rule – at which the researcher conducts a study with $S = s^{\text{int}}$, and the journal only publishes studies with a nonnegative interim net benefit – than any rule that publishes nothing at all. (In the model without incentives in which $q = 1$ and S is deterministically equal to s^{int}, publishing no studies is feasible, but the interim-optimal rule is preferred.) So the incentive-optimal rule will induce the researcher to conduct a study, meaning that the researcher must be choosing $S = \overline{s}$.

Second, fix any publication rule of the form in Lemma 1 with $\overline{s} = s^h$ and $\lambda = \lambda^h$, for $s^h > s^{\text{int}}$. We claim that the publication rule of the same form with $\overline{s} = s^{\text{int}}$ and $\lambda = 0$ weakly improves payoffs. To see why this claim holds, note that the publication rule with $\overline{s} = s^h$ and $\lambda = \lambda^h$ would be weakly improved upon by one with $\overline{s} = s^{\text{int}}$ and $\lambda = 0$, supposing researcher participation. Recall that normal signals are Blackwell ordered by their standard errors: at standard error $S = s^{\text{int}}$, the findings X can be garbled into something informationally equivalent to findings from $S = s^h$. So some stochastic publication rule at $S = s^{\text{int}}$, combined with a garbling of these signals to the public, replicates the distribution of outcomes[8] that occur when a study arrives with $S = s^h$ and is published under the publication rule given by $\overline{s} = s^h$ and $\lambda = \lambda^h$. But the journal's payoffs given a study with $S = s^{\text{int}}$ are improved by removing the garbling to the public. Payoffs are further improved by publishing under the interim-optimal publication rule at $S = s^{\text{int}}$, which is exactly that given by a rule of the form

[8]I.e., the probability of publication at each state, and the joint distribution over public actions and states conditional on publication.

in Lemma 1 with $\overline{s} = s^{\text{int}}$ and $\lambda = 0$. Finally, by Assumption 1, the publication rule with $\overline{s} = s^{\text{int}}$ and $\lambda = 0$ does indeed get researcher to conduct a study, since the interim outcome satisfies the researcher's participation constraint.

The final step is to show that $\lambda \geq 0$. This is because, for any publication rule of the form of Lemma 1, increasing λ increases the publication probability at $S = \overline{s}$. Hence, it makes the researcher better off if she chooses $S = \overline{s}$ and slackens her incentive constraints. Moreover, starting from $\lambda < 0$, increasing λ to 0 improves the journal's payoff, since again $\lambda = 0$ is interim optimal and hence optimal conditional on a study being submitted at $S = \overline{s}$. □

Proof of Lemma 1. Take an arbitrary publication rule $\tilde{p}$. We will show that it can be replaced by a rule p of the desired form that weakly increases the journal's payoff.

First suppose that $\tilde{p}$ does not induce the researcher to conduct a study. Then define some p of the form in the statement of the Lemma by setting $\overline{s}$ arbitrarily and setting $\lambda = 0$. If the publication rule p induces the researcher not to participate, then the journal's payoffs are unchanged from $\tilde{p}$. If the rule p induces the researcher to conduct a study with standard error $S = s$, then the journal's payoffs are weakly higher than before, since under p the journal never publishes studies that give negative net interim payoff.

So, for the rest of the proof, assume that $\tilde{p}$ does in fact induce the researcher to conduct a study with S equal to some level $\overline{s}$. We show that there exists λ such that we can replace $\tilde{p}$ with a publication rule p satisfying the following properties and weakly improve the journal's payoff:

1. At $s > \overline{s}$, $p(x, s) = 0$:
   Let $p(x, s) = \tilde{p}(x, s)$ at $s \leq \overline{s}$ and 0 at $x > \overline{s}$. The publication rule p gives the researcher the same payoff from choosing $S = \overline{s}$ and weakly reduces her payoff from choosing other values of S, and so under p the researcher's behavior is unchanged. She continues to conduct a study with $S = \overline{s}$ and the journal's payoff given the choice of $S = \overline{s}$ is also unchanged.
2. At $s = \overline{s}$, $p(X, \overline{s}) = 1$ if $\Delta(\pi_1^{(X,\overline{s})}, a^*(\pi_0)) > c-\lambda$, and $p(X, \overline{s}) = 0$ if $\Delta(\pi_1^{(X,\overline{s})}, a^*(\pi_0)) < c - \lambda$:
   Let $p(x, s) = \tilde{p}(x, s)$ at all $s \neq \overline{s}$. Denote the probability of publication under $\tilde{p}$ at $S = \overline{s}$, given by $\mathbb{E}[\tilde{p}(X, S)|S = \overline{s}]$, by $y \in [0, 1]$. If $y = 0$ then $\tilde{p}$ is equivalent to a publication rule p of the appropriate form with $\lambda = \infty$. If $y = 1$ then $\tilde{p}$ is equivalent to a publication rule p of the appropriate form with $\lambda = -\infty$.
   For interior y, define $p(\cdot, \overline{s})$ so as to maximize the journal's payoff subject to accepting a share y of papers at this standard error. To do so, first set $\lambda \in \mathbb{R}$ as the supremum over values of l such that $P(\Delta(\pi_1^{(X,\overline{s})}, a^*(\pi_0)) > c - l|S = \overline{s}) \leq y$. Next, let $p(x, \overline{s}) = 0$ if $\Delta(\pi_1^{(x,\overline{s})}, a^*(\pi_0)) < c - \lambda$ and let $p(x, \overline{s}) = 1$ if $\Delta(\pi_1^{(x,\overline{s})}, a^*(\pi_0)) > c - \lambda$. Finally, if $\Delta(\pi_1^{(x,\overline{s})}, a^*(\pi_0)) = c - \lambda$, set $p(x, \overline{s})$ such that the publication probability at $S = \overline{s}$, $\mathbb{E}[p(X, S)|S = \overline{s}]$, is equal to y. (This last step is only relevant if $\Delta(\pi_1^{(X,\overline{s})}, a^*(\pi_0)) = c - \lambda$ with positive probability at

$S = \overline{s}$.)

The publication rules p and $\tilde{p}$ publish with the same probability as each other conditional on any choice S by the researcher. Hence, the researcher continues to be willing to pick $S = \overline{s}$. Moreover, given the constraint of publishing with probability y at $S = \overline{s}$, the journal's expected payoff given a researcher choice of $S = \overline{s}$ is maximized by p. Hence, the journal weakly prefers p to $\tilde{p}$ if the researcher is to choose $S = \overline{s}$.

3. At $s < \overline{s}$, $p(x, s) = 1$ if $\Delta(\pi_1^{(x,s)}, a^*(\pi_0)) \geq c$ and $p(x, s) = 0$ if $\Delta(\pi_1^{(x,s)}, a^*(\pi_0)) < c$:

   Let $p(x, s) = \tilde{p}(x, s)$ at $s \geq \overline{s}$; at $s < \overline{s}$, let $p(x, s) = 1$ if $\Delta(\pi_1^{(x,s)}, a^*(\pi_0)) \geq c$ and $p(x, s) = 0$ if $\Delta(\pi_1^{(x,s)}, a^*(\pi_0)) < c$.

   Under publication rule p, the researcher will either continue to choose $S = \overline{s}$ or will switch to $s' < \overline{s}$. If the researcher continues to choose $S = \overline{s}$, then the journal's payoffs are as before. If the researcher now chooses $s' < \overline{s}$, we claim that the journal must be weakly better off. (This argument exactly follows an argument in the proof of Proposition 2.) To show the claim, recall that normal signals are Blackwell ordered by their standard errors: at standard error $S = s'$, the finding X can be garbled into something informationally equivalent to a finding from $S = \overline{s}$. So some stochastic publication rule at $S = s'$, combined with a garbling of these signals to the public, replicates the distribution of outcomes (probability of publication at each state, and joint distribution over public actions and states conditional on publication) that occur when a study arrives with $S = \overline{s}$ and is published under the publication rule given by $p(X, \overline{s})$. But the journal's payoffs given a study that has been published with $S = s'$ are improved by removing the garbling to the public. Payoffs are further improved by publishing under the interim-optimal publication rule at $S = s'$, which is exactly that under p.

The only remaining item to prove is that it is without loss of generality to suppose that if the researcher chooses to conduct a study, she chooses $S = \overline{s}$; applying step 3 above could possibly have changed the researcher's choice of S to something below $\overline{s}$. However, iterating step 1 (with $\overline{s}$ redefined to the new choice of S) recovers a publication rule of the appropriate form in which the researcher does choose $S = \overline{s}$. $\square$

E.4 Proofs for Online Appendix D

E.4.1 Proofs for Online Appendix D.1

Proof of Proposition 3.

1. All comparative statics are immediate from the formula.
2. The only comparative static that is not immediate is that for the t-statistic cutoff, $\left(\frac{1}{S} + \frac{S}{\sigma_0^2}\right)(c - \mu_0)$, with respect to S. The argument for this result follows

identically as the argument for the analogous result in the proof of Proposition 1 part 3.

□

Proof of Proposition 4.

1. This result is immediate.
2. The derivative of the benefit with respect to σ_0 evaluates to

$$(1-\alpha)\frac{2s_2^4\sigma_0^7(S_1^4\sigma_0^4+2s_2^4(S_1^2+\sigma_0^2)(2S_1^2+\sigma_0^2)+s_2^2(5S_1^4\sigma_0^2+4S_1^2\sigma_0^4))}{(s_2^2+\sigma_0^2)^3(S_1^2+\sigma_0^2)^2(S_1^2\sigma_0^2+s_2^2S_1^2+s_2^2\sigma_0^2)^2}$$

which is positive. As $\sigma_0 \to 0$, the numerator goes to 0 while the denominator goes to a positive limit.

3. The derivative of the benefit with respect to S_1 evaluates to

$$-(1-\alpha)\frac{2s_2^4S_1\sigma_0^8(2S_1^2\sigma_0^2+\sigma_0^4+2s_2^2S_1^2+2s_2^2\sigma_0^2)}{(s_2^2+\sigma_0^2)^2(S_1^2+\sigma_0^2)^2(S_1^2\sigma_0^2+s_2^2S_1^2+s_2^2\sigma_0^2)^2}$$

which is negative. As $S_1 \to \infty$, the numerator is constant while the denominator goes to infinity.

4. The derivative of the benefit with respect to s_2 evaluates to

$$(1-\alpha)\frac{2\sigma_0^8s_2^3(-s_2^4(S_1^2+\sigma_0^2)+s_2^2\sigma_0^2(S_1^2+\sigma_0^2)+2S_1^2\sigma_0^4)}{(s_2^2+\sigma_0^2)^3(S_1^2+\sigma_0^2)(S_1^2\sigma_0^2+s_2^2S_1^2+s_2^2\sigma_0^2)^2}.$$

This has the sign of $-s_2^4(S_1^2+\sigma_0^2)+s_2^2\sigma_0^2(S_1^2+\sigma_0^2)+2S_1^2\sigma_0^4$. This expression is a concave quadratic in s_2^2, which is positive at $s_2^2=0$ and maximized at $s_2^2=\frac{\sigma_0^2}{2}>0$. In particular, the derivative in s_2 is positive and then negative. As $s_2 \to 0$, the numerator goes to zero while the denominator goes to a positive limit. As $s_2 \to \infty$, the numerator increases at a rate of s_2^4 while the denominator increases at a rate of s_2^6, so the ratio goes to 0. □

References

Dawid, A. P. (1973): "Posterior expectations for large observations," *Biometrika*, 60, 664–667.

Harbaugh, R., J. Maxwell, and K. Shue (2016): "Consistent good news and inconsistent bad news," Working paper.

Henry, E. and M. Ottaviani (2017): "Research and the approval process: the organization of persuasion," .

Libgober, J. (2015): "False Positives in Scientific Research," *Working Paper, Harvard University.*

McClellan, A. (2017): “Experimentation and Approval Mechanisms,” .

O’Hagan, A. (1979): “On outlier rejection phenomena in Bayes inference,” *Journal of the Royal Statistical Society. Series B (Methodological)*, 358–367.

Subramanyam, K. (1996): “Uncertain precision and price reactions to information,” *Accounting Review*, 207–219.

Sun, Y., Á. Baricz, and S. Zhou (2010): “On the Monotonicity, Log-Concavity, and Tight Bounds of the Generalized Marcum and Nuttall Q-Functions,” *IEEE Transactions on Information Theory*, 56, 1166–1186.

Tetenov, A. (2016): “An economic theory of statistical testing,” Tech. rep., cemmap working paper, Centre for Microdata Methods and Practice.

Yoder, N. (2018): “Designing Incentives for Heterogeneous Researchers,” Working paper.