

COMMUNICATING SCIENTIFIC UNCERTAINTY VIA APPROXIMATE POSTERiors

ISAIAH ANDREWS
MIT and NBER

JESSE M. SHAPIRO
Harvard University and NBER

We cast the problem of communicating scientific uncertainty as one of reporting a posterior distribution on an unknown parameter to an audience of Bayesian decision-makers. We establish novel bounds on the audience's regret when the analyst reports an approximation to a posterior that the audience treats as exact. Under a palatable restriction on the audience's decision problems, the bounds take an especially convenient form. Under a further restriction on the audience's priors, a bootstrap distribution can be used as a stand-in posterior. We propose a practical recipe for checking whether a conventional statistical report (say, a normal parameterized by a point estimate and standard error) is a good approximation, and for improving the report if it is not. We illustrate our proposals using the articles in the 2021 *American Economic Review* that use a bootstrap for inference.

KEYWORDS: Statistical communication, signed Kolmogorov distance, Bayes bootstrap.

1. INTRODUCTION

COMMUNICATING UNCERTAINTY about one's conclusions is a fundamental task of scientists. Within empirical economics, a conventional approach is to report a statistic, often a standard error, alongside the point estimate of a target parameter (Athey and Imbens (2023)). An important question is whether, and when, this convention accurately conveys the uncertainty about the target parameter.

In this paper, we propose to approach the problem of communicating uncertainty about a target parameter from a Bayesian perspective. Theoretically, we argue that this perspective leads to a natural notion of what is a “good” summary of the uncertainty about the target parameter. Practically, we argue that this perspective leads to a widely applicable toolkit for diagnosing and addressing failures of the conventional report.

We begin with the theoretical question of what constitutes a good summary of uncertainty for a Bayesian audience. We formulate this question in a model of statistical communication following Andrews and Shapiro (2021). An analyst makes a report about a target parameter to an audience of agents. Each agent faces a decision problem with an associated loss function, and holds a prior in a density ratio neighborhood of some *central prior*. Different agents may face different decision problems and hold different priors,

Isaiah Andrews: iandrews@mit.edu

Jesse M. Shapiro: jesse_shapiro@fas.harvard.edu

We acknowledge funding from the National Science Foundation (SES-1654234, SES-1949047) and the Semester Undergraduate Program for Economics Research at Harvard. We thank our dedicated research assistants for their contributions to this project, as well as the many authors who helped us to work with their code, data, and bootstrap replicates. We thank Giuseppe Cavaliere, Clément de Chaisemartin, Luca Fanelli, Anna Mikusheva, Jon Roth, and conference and seminar participants at the Princeton Day of Statistics, Harvard, MIT, the University of Tokyo, USC, UCLA, the IAAE, Yale, and the Reinhardt Lecture for helpful comments.

reflecting heterogeneity in the uses of scientific knowledge. Nevertheless, well-known results imply that reporting the *central posterior* that results from updating the central prior is as good—in terms of the loss it induces—as reporting the full data.

We ask what harm is done if, instead of reporting the central posterior, the analyst reports an approximation to it, and the audience treats the approximation as if it is exact. For an agent with a given prior and loss, we measure the harm by the agent's *posterior regret*: the expected loss from the action they will take based on the approximation, minus the expected loss from the best action they could take given the data. We establish a fundamental bound on the posterior regret. The bound has an intuitive economic structure, and is sharp when the agents' decision problems are sufficiently rich.

We turn next to the question of when to accept a given *conventional report*, such as a point estimate and standard error. We allow that there is a benefit to reporting a conventional approximation to the central posterior, for example because of the convention's simplicity and familiarity. Our theoretical results suggest a natural *procedure* for the analyst to follow in this case. Compute the conventional report, the central posterior, and the implied regret bound. If the regret bound is smaller than the benefit of adhering to the convention, make the conventional report. If the regret bound is larger than the benefit of adhering to the convention, report the central posterior. This procedure ensures that the audience's regret is no greater than the benefit of adhering to the convention.

To make this procedure a practical one, we need three ingredients: a conventional report, a class of audience decision problems, and a central posterior. For the conventional report, we focus on the practice of reporting a point estimate and standard error, which we model as reporting a normal approximation to the central posterior. We motivate this convention by recalling that Bernstein–von Mises (BvM) theory establishes a range of conditions on a prior and parameter under which the corresponding posterior is well approximated by a normal distribution in large samples. If the central posterior for the target parameter satisfies the conditions for a BvM theorem, then well-known results imply a variety of senses in which a normal distribution parameterized by the point estimate and standard error is a good approximation to the central posterior.

For the class of audience decision problems, we focus on the case where agents' loss functions and density ratio functions are bounded and of bounded increasing and decreasing variation. We argue that this class is sufficiently inclusive for use in practice. Under the proposed assumption, we prove that the bound on regret is itself bounded, both above and below, by scalar multiples of the signed Kolmogorov (SK) metric between the central posterior and the conventional approximation. The SK metric measures the distance between two distributions on the real line by adding together the magnitudes of the largest positive and negative vertical distances between their CDFs. Conveniently, the SK metric can therefore be read directly off of a p-p plot.

For the central posterior, the analyst may have a subjective prior whose implied posterior can serve in this role. To accommodate the remaining (and, we suspect, many) situations where this is not the case, we extend our results to show that if the analyst uses a *surrogate central posterior* that is close to the central posterior, the bounds we derive continue to apply up to a slack term that depends on the difference between the surrogate central posterior and the true central posterior.

We propose to use the Bayesian bootstrap distribution (Rubin (1981), Gasparini (1995), Chamberlain and Imbens (2003)) as a default surrogate central posterior. We characterize two situations in which this is reasonable. One situation is where the central prior is a Dirichlet process, in which case existing results show that the Bayes bootstrap distribution is the limit of a sequence of posteriors as the informativeness of the prior becomes low.

We adapt these results to show conditions under which the Bayes bootstrap distribution is close to the central posterior in the relevant metric. The other situation is where the target parameter is a functional of some other parameter, for example the CDF of the data or a vector of regression coefficients, that obeys a BvM theorem under both the bootstrap distribution and the central prior. In that case, we show that the Bayes bootstrap distribution serves as a surrogate with high probability as the sample grows large. Under these same conditions, we show that the Bayes bootstrap distribution is similar to that of other weighted bootstraps, such as the nonparametric bootstrap, so that these other bootstraps can reasonably be used when more convenient.

We thus arrive at a practical recipe for researchers in empirical economics. Calculate the point estimate and standard error. Sample from a bootstrap distribution. Make a p-p plot comparing the two distributions. If the distributions are close in SK metric, make the conventional report. If not, report the bootstrap distribution directly, for example with a histogram or CDF. An accompanying python package (<https://github.com/JMSLab/BootstrapReport>) and web app (<https://jmslab.shinyapps.io/bootstrapreport/>) called `BootstrapReport` facilitate adoption.

We illustrate practicality by applying the recipe to the universe of articles in the 2021 *American Economic Review* which bootstrap some target parameter, and for which we were able to recover bootstrap replicates. These articles cover a wide range of fields and methods. Our proposed approach applies readily to all of them. For some target parameters, the conventional report approximates the bootstrap distribution well. For others, it does not. The quality of the approximation can differ meaningfully even across different parameters targeted in the same article.

This paper makes two main contributions. The main theoretical contribution is a set of novel bounds on the posterior regret from reporting an approximate central posterior to an audience of decision-makers who believe they have received an exact report.¹ In addition to those we explore here, these bounds may have other uses. For example, in situations where the analyst performs an explicitly Bayesian analysis, and wishes to provide a low-dimensional summary of the posterior distribution for a target parameter (e.g., a posterior mean and standard deviation), our results can be used to bound the harm from doing so.²

The main practical contribution is a procedure for evaluating the quality of a normal (or other) approximation to the uncertainty in an estimated parameter. Previous work in the frequentist tradition (e.g., [Beran \(1997\)](#), [Zhan \(2018\)](#), [Angelini, Cavaliere, and Fanelli \(2022, 2024\)](#)) has proposed using a bootstrap to diagnose failures of conventional asymptotic approximations.³ In contrast to this work, and to much other previous work on uncertainty quantification in economics, our approach neither aims for, nor achieves,

¹Our focus on communicating a useful summary of data to an audience of decision-makers is related to research on omniprediction (e.g., [GTKRS+ \(2022\)](#)) and sequential calibration (e.g., [Noarov, Ramalingam, Roth, and Xie \(2025\)](#)), though our setting, problem statement and, consequently, proposed solutions, are different.

²[Cohen and Einav \(2007\)](#), for example, summarize many marginal posterior distributions with two statistics, the posterior mean and standard deviation. [Smets and Wouters \(2007\)](#) summarize with four statistics, the mode, mean, and 5th and 95th quantiles. In each case, if we think of these statistics as parameterizing some family of approximating distributions, our bound can be used directly to evaluate the quality of the resulting approximation to the sampled posterior.

³See also [Cavaliere, Fanelli, and Georgiev \(2025\)](#). Previous work in the frequentist tradition also recommends reporting richer bootstrap information than the standard error (see, e.g., [Efron \(1982\)](#), [Hall \(1992\)](#), and [Hahn and Liao \(2021\)](#)). [Wang \(2025\)](#) suggests diagnosing approximation failures for GMM via a comparison to a quasi-Bayes posterior distribution.

frequentist guarantees such as size control.⁴ Instead, our approach focuses on limiting the potential for bad decision-making by Bayesian agents treating the reported approximation as a central posterior.

There are two main reasons we believe the Bayesian communication perspective we adopt here is useful for empirical economics. First, the description of agents who treat the scientist's report as a summary of a revised belief (posterior) about the target parameter matches well with how we and, we suspect, others, consume empirical research. Second, the use of a Bayesian perspective allows us to make certain statements without reference to asymptotic or other approximations. In particular, if the bootstrap distribution differs meaningfully from the default report, our results establish that there exists *some* agent who would be meaningfully better off basing their decision-making directly on the bootstrap distribution.

Wide applicability of our procedure does require substantive assumptions. When these assumptions are reasonable, our procedure affords a recipe for quantifying uncertainty about a target parameter without assuming approximate normality of the sampling distribution of the estimator, and without specializing the recipe for particular use cases. This is in contrast to the large frequentist literature on failures of the conventional normal approximation, which requires specialized treatments depending on the nature of the target parameter and the reason for the failure of the conventional normal approximation, and which aims at goals, such as size control, that differ from those we pursue here.

The remainder of the paper is organized as follows. Section 2 introduces our communication framework and our main theoretical result bounding the regret from reporting an approximate posterior. Section 3 introduces the conventional report and the procedure to control regret. Section 4 introduces the class of audience decision problems, and the theoretical results, that justify the use of the SK metric. Section 5 introduces the idea of a surrogate central posterior and develops the use of the bootstrap in this role. Section 6 presents the findings from our census of articles in the 2021 *American Economic Review* that use the bootstrap. A running example throughout the article serves to illustrate the main ideas and limitations. An Appendix contains proofs of all results stated in the main text. The Supplemental Material (Andrews and Shapiro (2026)) contains additional theoretical results, illustrations for the running example, and findings from the bootstrap census.

2. A BOUND ON REGRET FROM APPROXIMATE POSTERIOR REPORTS

In this section, we lay out our abstract framework, which follows ideas in Andrews and Shapiro (2021). We use the framework to derive a fundamental bound on the regret from reporting an approximate posterior belief on the target parameter. We develop practical implications in later sections.

2.1. *Communicating a Posterior Report*

Consider an analyst who observes data $X \in \mathcal{X}$. The analyst reports some function of the data to an audience of agents, who each must take an action $a \in \mathcal{A}$.⁵ The consequences of each agent's decision depend on the value of an unknown *target parameter*

⁴Indeed, for the frequentist goal of describing the sampling distribution of the target parameter, results in Singh (1981) and Weng (1989) suggest that the nonparametric bootstrap is more accurate than the Bayes bootstrap that we recommend as the default surrogate for the central posterior (see also Hall (1992)).

⁵As the dimension of $\mathcal{A}$ is unrestricted, the assumption that all agents share a common action space $\mathcal{A}$ is without loss of generality.

$\theta \in \Theta$. Specifically, each agent is endowed with a loss function $L : \mathcal{A} \times \Theta \rightarrow [0, \lambda]$ in some class $\mathcal{L}$, so that an agent with loss function $L \in \mathcal{L}$ who takes action $a \in \mathcal{A}$ realizes loss $0 \leq L(a, \theta) \leq \lambda < \infty$ when the target parameter's true value is $\theta \in \Theta$. The assumption that the loss is bounded by λ implies that the decisions taken based on the research findings have bounded stakes.

Each agent is further endowed with a prior $\pi \in \Delta(\Theta \times \mathcal{X})$ that is dominated by some *central prior* π^* . We can describe the agent's prior with a marginal prior $\pi(\theta) \in \Delta(\Theta)$ on the target parameter, and a conditional prior $\pi(X|\theta) \in \Delta(\mathcal{X})$ on the data given the target parameter. We assume that each agent's conditional prior given θ agrees with that of the central prior, so that $\pi(X|\theta) = \pi^*(X|\theta)$ for all $\theta \in \Theta$.⁶ All prior disagreement in the audience therefore concerns the marginal prior, and we can compactly describe any prior π by a density ratio function $w(\theta) = \frac{d\pi}{d\pi^*}(\theta)$. We let $\mathcal{W}$ denote a class of such density ratio functions, $w \in \mathcal{W}$. We assume that the class $\mathcal{W}$ includes $w(\theta) = 1$, which corresponds to an agent whose prior agrees with the central prior.

EXAMPLE—Get Out the Vote: Our running example follows [Nickerson, Friedrichs, and King \(2006\)](#). The analyst observes the results of an experiment in which (N^T, N^C) citizens are chosen at random to either receive (T) or not receive (C) a phone call encouraging them to vote. The data consist of the number $X = (X^T, X^C)$ of citizens in each group who vote. All agents believe that the data are independently binomially distributed according to $X^T|N^T, \beta^T \sim B(N^T, \beta^T)$ and $X^C|N^C, \beta^C \sim B(N^C, \beta^C)$ where $\beta^T, \beta^C \in [0, 1]$ are probabilities.

The agents in the audience are campaign managers for local candidates who must decide what share $a \in [0, 1] = \mathcal{A}$ of a fixed budget to devote to phone calls instead of other campaign activities. For a given campaign manager, the loss $L(a, \theta) \in [0, 1]$ is the probability that their candidate is defeated.

The loss depends on the effectiveness of a phone call, which is encoded in the probabilities (β^T, β^C) . If the target parameter θ is maximally rich, $\theta = (\beta^T, \beta^C)$, then the assumption that all disagreement concerns the marginal prior is without loss. If instead the parameter of interest is more restrictive, then our assumptions on the form of disagreement are more substantive.

To illustrate, suppose that the loss depends on $\theta^{TC} = \beta^T - \beta^C \in [-1, 1] = \Theta$, the effect of a phone call on the probability of voting. Suppose further that, under the central prior π^* , we have that $\theta^{TC} \sim U[\underline{\theta}, 1]$, for $\underline{\theta} > 0$ a small value, and $\beta^C|\theta^{TC} \sim U[0, 1 - \theta^{TC}]$, which together suffice to also describe $\beta^T|\theta^{TC}$ and therefore, given the binomial likelihood, the distribution of the data X . Figure 1 depicts the density implied by the central prior.

Under the assumption that $\pi(X|\theta) = \pi^*(X|\theta)$, audience members all agree that $\beta^C|\theta^{TC} \sim U[0, 1 - \theta^{TC}]$, but may disagree that $\theta^{TC} \sim U[\underline{\theta}, 1]$. Notice that, under all such prior distributions, we have that $\beta^T > \beta^C$ a.s., that is, that the intervention increases turnout.

Certain forms of disagreement can be incorporated directly into the loss function, rather than the prior. Suppose, for example, that a particular campaign manager believes that the experiment is not externally valid, so that θ^{TC} overstates the true causal effect of a phone call by a factor of $\eta \geq 1$. We can write such a manager's loss as

⁶Here and throughout, we use $\pi(\theta)$ and $\pi(X|\theta)$ as shorthand for the measures $\pi_\theta(B)$ for $B \subseteq \Theta$ and $\pi_{X|\theta}(B|\theta)$ for $B \subseteq \mathcal{X}$, respectively. Other similar notation may be read analogously, and our assumption that each agent's conditional prior agrees with that of the central prior may be read as $\pi_{X|\theta}(B|\theta) = \pi_{X|\theta}^*(B|\theta)$ for all measurable sets $B \subseteq \mathcal{X}$.

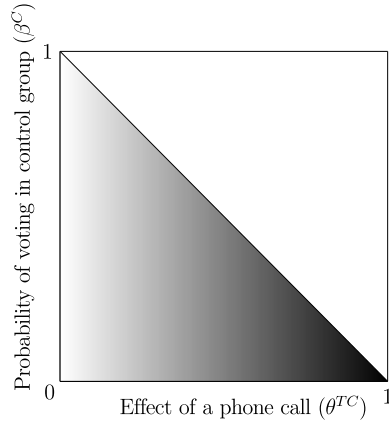


FIGURE 1.—Central prior in the example. *Note:* The plot shows the density of the joint central prior on the probability β^C of voting for citizens in the control group, and the effect θ^{TC} of a phone call on the probability of voting.

$L(a, \theta^{TC}) = \tilde{L}(a, \theta^{TC}/\eta)$ for $\tilde{L}$ a loss that is parameterized by the true causal effect θ^{TC}/η .

What remains is to describe the analyst's report. The analyst's information consists of the data X . If an agent with prior π were to directly observe the data X , their posterior belief about the target parameter would be $\pi(\theta|X)$. It follows that, from the agent's standpoint, no report can improve upon a report of $\pi(\theta|X)$. Importantly, the same is also true of a report that consists of the *central posterior* $\pi^*(\theta|X)$. The reason is that each agent's posterior is just a reweighting of the central posterior, that is, $\pi(\theta|X) \propto w(\theta)\pi^*(\theta|X)$.⁷ This is a classical justification for reporting posterior distributions (see, e.g., [Raiffa and Schlaifer \(1961\)](#), [Hildreth \(1963\)](#), and [Geweke \(1997\)](#)).

A leading possibility is, then, that the analyst reports $\pi^*(\theta|X)$. For reasons we discuss below, the analyst may instead prefer to report some other distribution $\hat{\pi}^*(\theta|X)$, for example as an approximation to $\pi^*(\theta|X)$. We assume that $\hat{\pi}^*(\theta|X)$ is absolutely continuous with respect to the central prior π^* , and denote by $\Delta_{\pi^*}(\Theta)$ the set of such distributions. We next measure the harm to the agents if $\hat{\pi}^*(\theta|X) \neq \pi^*(\theta|X)$.

2.2. Regret From an Approximate Posterior Report

We assume that each agent treats the analyst's report as if it is a report of the central posterior $\pi^*(\theta|X)$. Each agent therefore forms their perceived posterior belief $\hat{\pi}(\theta|X)$, with $\hat{\pi}(\theta|X) \propto w(\theta)\hat{\pi}^*(\theta|X)$, by reweighting the report. Each agent then seeks to minimize their perceived posterior expected loss $E_{w(\theta)\hat{\pi}^*(\theta|X)}[L(a, \theta)]$.⁸ A perceived optimal action is then

$$\hat{a} \in \arg \min_{a \in \mathcal{A}} E_{w(\theta)\hat{\pi}^*(\theta|X)}[L(a, \theta)].$$

⁷Here and throughout, we use $w(\theta)\pi^*(\theta|X)$ to denote the measure defined by $\mu(A) = \int_A w(\theta) d\pi^*(\theta|X)$ for measurable $A \subseteq \Theta$.

⁸Here and throughout, for any measure μ on Θ and any function $f: \Theta \rightarrow \mathbb{R}$, we let the operator $E_\mu[f(\theta)] = \frac{\int_\Theta f(\theta) d\mu(\theta)}{\int_\Theta d\mu(\theta)}$ denote the expectation of f with respect to the normalized measure $\mu / \int_\Theta d\mu(\theta)$. If μ is a probability measure, so that $\int_\Theta d\mu(\theta) = 1$, then $E_\mu[\cdot]$ corresponds to the usual expectation.

For concreteness, we proceed here as if an optimal action exists. The proofs in the [Appendix](#) cover the more general case where an optimal action may not exist, in which case the bounds we derive below continue to hold up to an arbitrarily small slack term.

If the analyst's report differs from the central posterior, an agent's perceived optimal action $\hat{a}$ may differ from their true posterior optimal action conditional on the data. We define an agent's *posterior regret* to be the increase in expected loss from taking the perceived optimal action $\hat{a}$ rather than the true optimal action. Formally, given data $X \in \mathcal{X}$, an agent with loss $L \in \mathcal{L}$ and density ratio $w \in \mathcal{W}$ has regret from the report $\hat{\pi}^*(\theta|X)$ given by

$$R(\hat{\pi}^*(\theta|X); X, L, w, \pi^*) = \mathbb{E}_{w(\theta)\pi^*(\theta|X)}[L(\hat{a}, \theta)] - \min_{a \in \mathcal{A}} \mathbb{E}_{w(\theta)\pi^*(\theta|X)}[L(a, \theta)],$$

where, importantly, both expectations are taken with respect to the agent's true posterior $\pi(\theta|X) \propto w(\theta)\pi^*(\theta|X)$.⁹

2.3. A Fundamental Bound on Regret

Our goal is to bound $R(\hat{\pi}^*(\theta|X); X, L, w, \pi^*)$. As notation, for any class $\mathcal{F}$ of functions $\mathcal{Y} \times \Theta \rightarrow \mathbb{R}$, we let $\mathcal{F}(\mathcal{Y})$ denote the class of functions formed by fixing a value of $y \in \mathcal{Y}$ and a function $f \in \mathcal{F}$, that is, functions of the form $f(y, \cdot)$, and we let $f(\mathcal{Y}) \subseteq \mathcal{F}(\mathcal{Y})$ denote the set of such functions formed using a particular $f \in \mathcal{F}$. For ease of exposition, we suppose that the set of pairs of expectations of differences of pairs of functions in $\mathcal{L}(\mathcal{A})$ under the true and perceived posteriors is compact.¹⁰ If this condition fails, as we discuss in the [Appendix](#), the sharpness result that we state below continues to hold up to an arbitrarily small slack term.

We then have the following result.

THEOREM 1—Fundamental Bound on Regret: *For any report $\hat{\pi}^*(\theta|X) \in \Delta_{\pi^*}(\Theta)$, data $X \in \mathcal{X}$, loss $L \in \mathcal{L}$, and density ratio $w \in \mathcal{W}$, the regret $R(\hat{\pi}^*(\theta|X); X, L, w, \pi^*)$ is bounded above by*

$$\begin{aligned} \bar{R}(\hat{\pi}^*(\theta|X); X, \mathcal{L}, \mathcal{W}, \pi^*) &= \sup_{\ell_1, \ell_0 \in \mathcal{L}(\mathcal{A}), w \in \mathcal{W}} \mathbb{E}_{w(\theta)\pi^*(\theta|X)}[\ell_1(\theta) - \ell_0(\theta)] \\ \text{s.t. } \mathbb{E}_{w(\theta)\hat{\pi}^*(\theta|X)}[\ell_1(\theta) - \ell_0(\theta)] &\leq 0. \end{aligned}$$

Moreover, there exist $\bar{w} \in \mathcal{W}$ and $\bar{L} : \mathcal{A} \times \Theta \rightarrow \mathbb{R}$ such that $\bar{L}(\mathcal{A}) \subseteq \mathcal{L}(\mathcal{A})$ and $R(\hat{\pi}^*(\theta|X); X, \bar{L}, \bar{w}, \pi^*) = \bar{R}(\hat{\pi}^*(\theta|X); X, \mathcal{L}, \mathcal{W}, \pi^*)$.

Intuitively, the regret bound $\bar{R}$ measures the difference between the approximation $\hat{\pi}^*(\theta|X)$ and the true central posterior $\pi^*(\theta|X)$, an interpretation that we return to later. To motivate the form of the bound, observe that if we fix an action $a \in \mathcal{A}$, any loss function $L \in \mathcal{L}$ implies a function $\ell(\theta) = L(a, \theta)$ only of the target parameter. Consider an agent

⁹If multiple optimal actions $\hat{a}$ exist, we consider the selection among them that yields the largest regret, that is, the selection that maximizes the value of $\mathbb{E}_{w(\theta)\pi^*(\theta|X)}[L(\hat{a}, \theta)]$.

¹⁰That is, we assume that the set

$$\{\mathbb{E}_{w(\theta)\pi^*(\theta|X)}[\ell_1(\theta) - \ell_0(\theta)], \mathbb{E}_{w(\theta)\hat{\pi}^*(\theta|X)}[\ell_1(\theta) - \ell_0(\theta)] : \ell_1, \ell_0 \in \mathcal{L}(\mathcal{A}), w \in \mathcal{W}\}$$

is compact.

choosing between actions a_0 and a_1 , with corresponding $\ell_0, \ell_1 \in \mathcal{L}(\mathcal{A})$. The worst case for this agent arises when ℓ_1 has a weakly lower perceived posterior expectation than ℓ_0 , but a much larger true posterior expectation, since in this case the agent is willing to make a choice yielding a large regret. The proof of Theorem 1 shows that characterizing the worst-case two-action decision problem is sufficient to bound the regret over all decision problems. The theorem also states that the bound is sharp in the sense that, if the audience is sufficiently rich, then there exists some agent whose regret achieves the bound.

The bound $\bar{R}(\hat{\pi}^*(\theta|X); X, \mathcal{L}, \mathcal{W}, \pi^*)$ can be practical to work with in some cases. For example, when $\mathcal{L}(\mathcal{A})$ is convex and $w(\theta) = 1$, the bound solves a convex programming problem. When $|\Theta| < \infty$, the program is finite-dimensional. When $|\Theta| = \infty$, the program is infinite-dimensional, but approximation methods exist under suitable restrictions on $\mathcal{L}(\mathcal{A})$ (see, e.g., Devolder, Glineur, and Nesterov (2010)). When $w(\theta)$ is non-constant, the program becomes non-convex but may remain tractable in some situations. As we seek a method that is generally computationally light, we prefer to offer a bound that can be read directly from the distributions $\hat{\pi}^*(\theta|X)$ and $\pi^*(\theta|X)$. We later derive such a bound based on an integral probability metric.

REMARK 1: Theorem 1 fixes a particular realization of the data and imagines that agents treat the analyst's report $\hat{\pi}^*(\theta|X)$ as if it is a report of the central posterior $\pi^*(\theta|X)$. We may instead imagine that, as in Andrews and Shapiro (2021), the analyst commits in advance to a reporting procedure and each agent updates according to Bayes's rule with full knowledge of the procedure. In that case, Supplemental Appendix A shows that each agent's expected regret is bounded above by the agent's expectation of the bound derived in Theorem 1, though this bound is not sharp.

3. CONTROLLING THE REGRET FROM A CONVENTIONAL REPORT

Within empirical economics, it is conventional to report a point estimate and standard error (Athey and Imbens (2023)). Scientific conventions are often imperfect but they do have value, for example because of their simplicity and familiarity (see, e.g., the discussion in BBJN+ (2018)). Here, we consider the situation of an analyst who values adhering to a convention but wishes to limit the harm from a poor approximation. We show how the analyst can use the bound in Theorem 1 to balance these goals. Lastly, we formalize the convention of reporting a point estimate and standard error as a normal (Gaussian) approximation to the central posterior, and discuss how this convention can be justified via Bernstein–von Mises theory.

3.1. A Procedure for Controlling Regret

Consider the situation of an analyst who must choose between reporting the central posterior $\pi^*(\theta|X)$ and reporting some particular approximation $\pi^0(\theta|X) \in \Delta_{\pi^*}(\Theta)$ prescribed by convention. There is an intrinsic value of $\kappa \geq 0$ to reporting the conventional approximation, in the sense that, all else equal, each agent would be willing to accept an additional loss of $\kappa \geq 0$ to receive the conventional report.

We continue to assume that agents treat any report as if it is a report of $\pi^*(\theta|X)$, and update accordingly. This may be because agents trust that $\pi^0(\theta|X)$ approximates $\pi^*(\theta|X)$, or because agents find it difficult to compute the posterior $\pi(\theta|\pi^0(\theta|X))$ that results from updating their beliefs directly from the conventional report.

The analyst seeks a procedure that, for any $X \in \mathcal{X}$, satisfies the following criteria:

- (C1) The analyst reports $\pi^0(\theta|X)$ only when the regret from doing so is no more than κ for all agents.
- (C2) The analyst reports $\pi^*(\theta|X)$ only when the regret from reporting $\pi^0(\theta|X)$ exceeds κ for some agent.

A procedure that satisfies Criterion (C1) avoids using the conventional report when doing so imposes, on some agent, regret that exceeds the intrinsic value κ of adhering to the convention. A procedure that satisfies Criterion (C2) conserves the conventional report when doing so does not impose a large regret on any agent.

Consider the following procedure for the analyst.

Procedure 1: Controlling regret from a conventional report.

- **Calculate**
 - central posterior $\pi^*(\theta|X)$
 - conventional approximation $\pi^0(\theta|X)$
 - regret bound $\bar{R}(\pi^0(\theta|X); X, \mathcal{L}, \mathcal{W}, \pi^*)$
 - **Choose** threshold τ
 - **Report** threshold τ and
 - $\hat{\pi}^*(\theta|X) = \pi^0(\theta|X)$ if $\bar{R}(\pi^0(\theta|X); X, \mathcal{L}, \mathcal{W}, \pi^*) \leq \tau$
 - $\hat{\pi}^*(\theta|X) = \pi^*(\theta|X)$ otherwise
-

We then have the following corollary of Theorem 1, taking $\bar{L}$ to be the loss function defined in the statement of the theorem.

COROLLARY 1: *If $\tau \leq \kappa$, then Procedure 1 satisfies Criterion (C1). If $\tau \geq \kappa$ and $\bar{L} \in \mathcal{L}$, then Procedure 1 satisfies Criterion (C2). If $\tau = \kappa$ and $\bar{L} \in \mathcal{L}$, then any procedure for choosing between reports $\pi^*(\theta|X)$ and $\pi^0(\theta|X)$ that satisfies Criteria (C1) and (C2) is equivalent to Procedure 1.*

Corollary 1 states that for τ sufficiently small, Procedure 1 satisfies Criterion (C1). Corollary 1 further states that, for τ sufficiently large, Procedure 1 satisfies Criterion (C2) if the class of losses is sufficiently rich, in the sense that it contains the loss $\bar{L}$ that attains the bound in Theorem 1. When both conditions hold, so $\tau = \kappa$, Procedure 1 is the unique procedure for choosing between reporting the central posterior and reporting the conventional approximation that satisfies both criteria.

Implementing Procedure 1 requires specifying a conventional approximation. We focus on a conventional normal approximation, which can be justified by Bernstein–von Mises (BvM) theory.

3.2. Justification for Conventional Normal Reports

One way to interpret the common convention of reporting a point estimate and standard error is as a normal approximation to the central posterior, that is,

$$\pi^0(\theta|X) = \pi^N(\theta|X) = N(\hat{\theta}(X), \hat{\Sigma}(X)),$$

where $\hat{\theta}(X)$ is a point estimate, $\hat{\Sigma}(X)$ is an estimated variance, and Θ is a finite-dimensional Euclidean space. One way to justify such a convention is via Bernstein–von

Mises (BvM) theory, which provides a range of conditions on the prior and target parameter under which $\pi^N(\theta|X) \approx \pi^*(\theta|X)$ in large samples.¹¹ Such approximations appear reasonable in many economic settings.

EXAMPLE: Get Out the Vote, Continued] The first row of plots in Figure 2 shows five marginal prior distributions on θ^{TC} . The middle plot in this row corresponds to the central prior π^* , which is uniform. The left plots correspond to audience priors that are “pessimistic” about the intervention in the sense that they put more density on lower values of the effect θ^{TC} than does the central prior. The right plots correspond to audience priors that are “optimistic” about the intervention in the sense that they put more density on higher values of the effect θ^{TC} than does the central prior.

The second row of plots shows, for each prior distribution in the first row, a histogram of MCMC draws from the corresponding posterior distribution, conditional on the realized data X reported in Nickerson, Friedrichs, and King (2006, Table 3), pooled across congressional districts. The middle plot in this row corresponds to the central posterior $\pi^*(\theta^{TC}|X)$. The left and right plots in this row correspond to the audience posteriors $\pi(\theta^{TC}|X)$. The audience posteriors are more similar to the central posterior than the audience priors are to the central prior. But, the audience posteriors reflect the pessimism and optimism of the corresponding priors.

In the second row, we overlay on each histogram the corresponding perceived posterior $\hat{\pi}(\theta^{TC}|X)$ formed by reweighting the normal approximation $\pi^N(\theta^{TC}|X)$ that is parameterized by the maximum likelihood point estimate $\hat{\theta}^{TC}(X)$ and estimated variance $\hat{\Sigma}^{TC}(X)$, and truncated to $\theta^{TC} \in [\theta, 1]$. For the central posterior, the normal approximation applies directly, $\hat{\pi}(\theta^{TC}|X) = \pi^N(\theta^{TC}|X)$. For the other posteriors, the normal approximation must be reweighted according to $\hat{\pi}(\theta^{TC}|X) \propto w(\theta)\pi^N(\theta^{TC}|X)$. In all cases, the perceived posterior is a close visual match to the true posterior.

An alternative way for agents to update would be to treat a report of $\hat{\theta}^{TC}(X)$, $\hat{\Sigma}^{TC}(X)$ as describing a quadratic approximation to the log-likelihood for θ^{TC} rather than as a normal approximation to the posterior for θ^{TC} . Using a point estimate and standard error as the basis for a quadratic approximation to a log-likelihood can be justified by large-sample approximations similar to those underlying BvM theory (see, e.g., Geyer (2013)). If the agents treat the log-likelihood as quadratic with mode $\hat{\theta}^{TC}(X)$ and Hessian $\hat{\Sigma}^{TC}(X)$, they arrive at posteriors identical to the perceived posteriors appearing in the second row of plots in Figure 2. Interpreting the conventional normal report as an approximation to the likelihood therefore offers an alternative justification for the model of agents’ belief updating that we encode in our definition of regret.

3.3. Failures of Conventional Normal Reports

While there are many economic settings where the approximation $\pi^N(\theta|X) \approx \pi^*(\theta|X)$ appears reasonable, there are others where it does not. Such situations are an important motivation for Procedure 1.

EXAMPLE—Get Out the Vote, Continued: Nickerson, Friedrichs, and King (2006, Table 6) report the cost per vote obtained via phone calls. We can write this parameter as

¹¹See, for example, Theorems 12.1 and 12.8 of Ghosal and van der Vaart (2017).

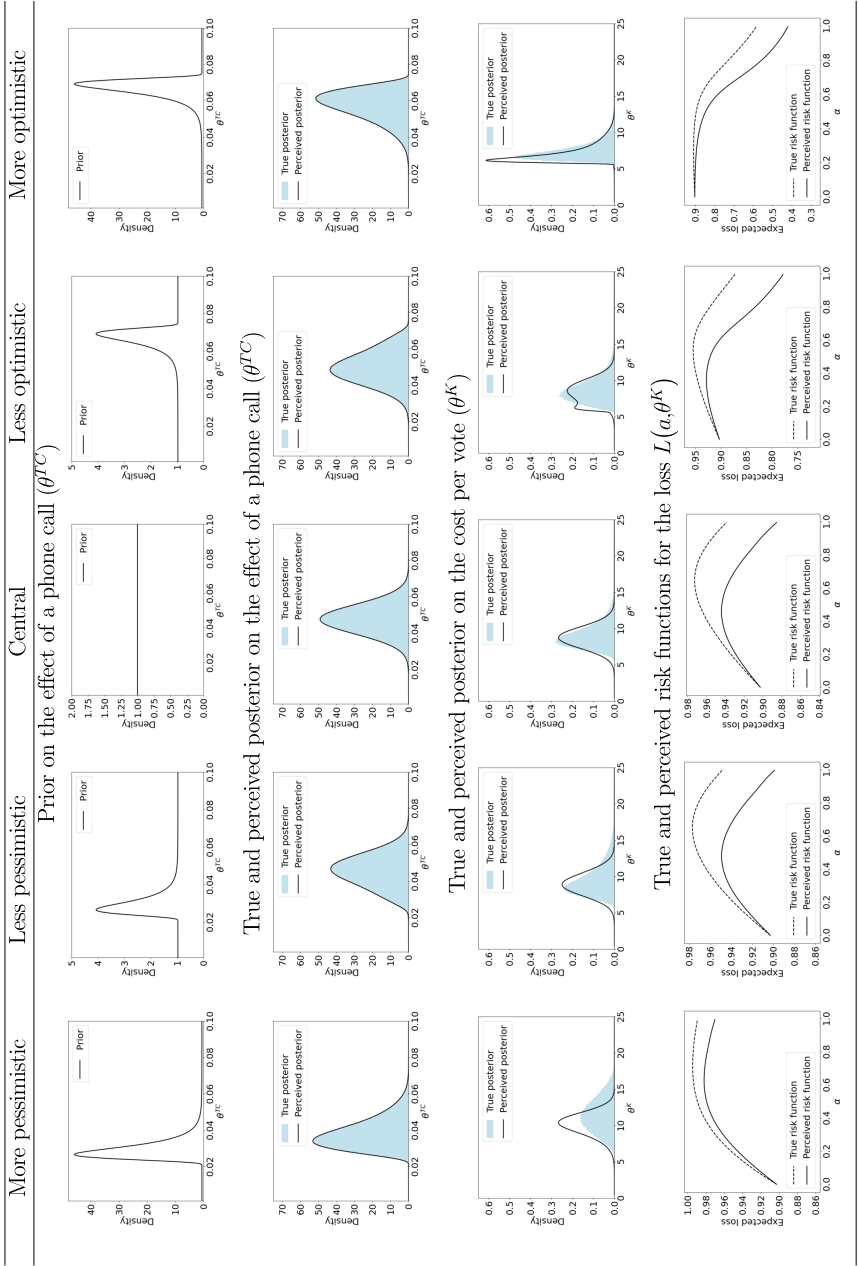


FIGURE 2.—Example prior distributions, posterior distributions, and decision problems. *Notes:* The first row shows the central prior, and four possible audience priors, on the effect θ^C of a phone call on the probability of voting. The histograms in the second row describe the corresponding true posteriors on θ^C given the observed data, approximated by MCMC. The histograms in the third row describe the corresponding true posteriors on θ^K given the observed data, approximated by MCMC. We overlay on each histogram the density of the perceived posterior formed by reweighing the normal distribution parameterized by the maximum likelihood point estimate and (delta-method) standard error. The fourth row describes the expected value of a particular loss function with respect to the true and perceived posteriors.

$\theta^K = K/(\beta^T - \beta^C)$ where $K > 0$ is the cost per phone call and where we recall that β^T and β^C are the probabilities of voting in the treatment and control groups, respectively. Previously, we took the target parameter to be $\theta^{TC} = (\beta^T - \beta^C)$, but the target parameter $\theta^K = K/(\beta^T - \beta^C)$ may be more directly useful to local campaign managers interested in optimally allocating their budgets.

The third row of plots in Figure 2 shows, for each prior distribution on θ^{TC} in the first row, a histogram of MCMC draws from the corresponding posterior distribution $\pi(\theta^K|X)$. We overlay on each histogram the corresponding perceived posterior $\hat{\pi}(\theta^K|X)$ formed by reweighting the normal approximation $\pi^N(\theta^K|X)$ that is parameterized by the maximum likelihood point estimate $\hat{\theta}^K(X)$ and the delta-method estimated variance $\hat{\Sigma}^K(X)$, and truncated to $\theta^K \in [K, \frac{K}{\underline{\theta}}]$. In all cases, the perceived posterior is visually distinct from the true posterior.

The difference between the reweighted normal approximation and the agent's posterior opens up the possibility of poor decision-making. To illustrate this possibility, we imagine a local campaign manager with a fixed budget $M > 0$, a share $a \in [0, 1]$ of which is allocated to phone calls. The remaining funds are spent on a last-minute rally the night before the election, which generates a return of Y votes per dollar, where Y is uncertain and drawn according to a known distribution with CDF $G(\cdot)$. The opponent is ahead by V votes, and the loss is the probability of defeat. We can therefore write the loss as

$$L(a, \theta^K) = G\left(\frac{V/M - a/\theta^K}{1 - a}\right).$$

For any prior π , we can define both the true risk function $E_{\pi(\theta^K|X)}[L(a, \theta^K)]$ and the perceived risk function obtained by replacing $\pi(\theta^K|X)$ with a reweighting of $\pi^N(\theta^K|X)$. The fourth row of plots in Figure 2 shows, for an illustrative parameterization of the loss, the true and perceived risk functions associated with each example prior. Under both the true and perceived posteriors, the optimal action for the more pessimistic agent is to devote no resources to phone calls, $a = \hat{a} = 0$, and the optimal action for the more optimistic agent is to devote all resources to phone calls, $a = \hat{a} = 1$. Because their optimal actions are identical under the true and perceived posteriors, these agents experience no regret from the conventional normal report. By contrast, under the true posterior, the optimal action for the agents with the less pessimistic and central priors is to devote no resources to phone calls, $a = 0$, whereas these agents' optimal action under the perceived posterior is to devote all resources to phone calls, $\hat{a} = 1$. These agents experience regret of, respectively, 0.046 and 0.036, from the conventional normal report.

Procedure 1 can avoid this regret. We have learned that there is an agent who experiences regret of at least 0.046 from the conventional report. If this regret exceeds the threshold τ , then Procedure 1 prescribes reporting $\pi^*(\theta^K|X)$ rather than $\pi^N(\theta^K|X)$.

The example shows, in a particular setting, how Procedure 1 can help the analyst to control the regret from reporting a poor approximation to the central posterior. To enable practitioners to apply Procedure 1 in a wide range of economic settings, we would ideally like to avoid the need to calculate the (possibly infinite-dimensional) program in Theorem 1. We next propose a more convenient bound that avoids that computation.

4. A CONVENIENT BOUND USING THE SK METRIC

To facilitate adoption of Procedure 1, we would like a more convenient bound that avoids the need to solve the (possibly infinite-dimensional) program in Theorem 1. In

this section, we first derive such a bound using an integral probability metric that depends on the classes of losses $\mathcal{L}$ and density ratio functions $\mathcal{W}$ in the audience. We then propose default choices for these classes that we argue are reasonable for many economic settings, and show that these choices lead to an especially convenient procedure.

4.1. A Bound in Terms of an Integral Probability Metric

For probability measures $\mu, \nu \in \Delta(\Theta)$ and a class $\mathcal{F}$ of functions $f : \Theta \rightarrow \mathbb{R}$ that are absolutely integrable w.r.t. μ, ν , let

$$d_{\mathcal{F}}(\mu, \nu) = \sup_{f \in \mathcal{F}} |E_{\mu}[f(\theta)] - E_{\nu}[f(\theta)]|$$

denote the largest difference in the expected values that the two measures assign to any function in $\mathcal{F}$. When $|E_{\mu}[f(\theta)] - E_{\nu}[f(\theta)]| \neq 0$ for some $f \in \mathcal{F}$ whenever $\mu \neq \nu$, $d_{\mathcal{F}}(\cdot, \cdot)$ is an integral probability metric (IPM).¹² IPMs play a role in many areas of probability theory. For example, if $\mathcal{F}$ is the set of all functions $f : \Theta \rightarrow [0, 1]$, then $d_{\mathcal{F}}(\mu, \nu) = d_{\text{TV}}(\mu, \nu)$ is the total variation distance, also defined as the largest difference in measure assigned to any set.¹³ To concisely describe other function classes, for scalars $\underline{\phi} \leq \bar{\phi}$ we let $\underline{\phi} \leq \mathcal{F} \leq \bar{\phi}$ denote that $f : \Theta \rightarrow [\underline{\phi}, \bar{\phi}]$ for all $f \in \mathcal{F}$. Further, for function classes, $\mathcal{G}$ and $\mathcal{F}$ we let $\mathcal{G} \cdot \mathcal{F}$ denote the class formed by the product of $g \in \mathcal{G}$ and $f \in \mathcal{F}$, that is, functions of the form $g(\cdot)f(\cdot)$.

We now state a bound on regret that takes the form of an IPM. This bound will be the basis of our practical recommendations.

PROPOSITION 1—Convenient Bound on Regret: *For any report $\hat{\pi}^*(\theta|X) \in \Delta_{\pi^*}(\Theta)$, data $X \in \mathcal{X}$, loss class $0 \leq \mathcal{L} \leq \lambda$, scalar $\omega \geq 1$, and density ratio class $\omega^{-1} \leq \mathcal{W} \leq \omega$, the bound $\bar{R}(\hat{\pi}^*(\theta|X); X, \mathcal{L}, \mathcal{W}, \pi^*)$ is itself bounded above by $2\omega \cdot d_{\mathcal{W} \cdot \mathcal{L}(\mathcal{A})}(\hat{\pi}^*(\theta|X), \pi^*(\theta|X))$. If, further, $\mathcal{L}(\mathcal{A})$ is centrosymmetric around $\frac{\lambda}{2}$ and $\mathcal{L}(\mathcal{A})$ contains all constant functions with values in $[0, \lambda]$, then $\bar{R}(\hat{\pi}^*(\theta|X); X, \mathcal{L}, \mathcal{W}, \pi^*)$ is bounded below by $d_{\mathcal{L}(\mathcal{A})}(\hat{\pi}^*(\theta|X), \pi^*(\theta|X))$.*

Proposition 1 states that, if the density ratios in the audience are bounded, the fundamental regret bound $\bar{R}$ is no larger than a scalar multiple of the IPM for the function class $\mathcal{W} \cdot \mathcal{L}(\mathcal{A})$ that depends on products of the density ratios $w \in \mathcal{W}$ and losses $L(\mathcal{A}) \in \mathcal{L}(\mathcal{A})$ in the audience. The scalar multiple depends, in turn, on the bound ω on the density ratio neighborhood. Proposition 1 further states that, under suitable conditions, the bound $\bar{R}$ is bounded below by the IPM for the function class $\mathcal{L}(\mathcal{A})$ that depends only on the class of losses in the audience. When the bound $\bar{R}$ is sharp, it follows directly that some agent has regret at least as large as the IPM for $\mathcal{L}(\mathcal{A})$.

4.2. Losses and Density Ratios of Bounded Signed Variation

To use Proposition 1 in practice, we must select a class of losses $\mathcal{L}$ and density ratios $\mathcal{W}$. In some settings, these choices may be clear from the communication problem faced by

¹²Otherwise, it is a pseudometric, as it always satisfies symmetry, nonnegativity, and the triangle inequality.

¹³That is, $d_{\text{TV}}(\mu, \nu) = \sup_{\Theta' \subseteq \Theta} |\mu(\Theta') - \nu(\Theta')|$.

the analyst. For the remaining settings where this is not the case, we recommend default choices.

For our recommended default, we consider the case where θ is scalar, $\Theta \subseteq \mathbb{R}$, and take $\mathcal{L}$ and $\mathcal{W}$ to consist of bounded functions of bounded signed variation. For any function $f : \Theta \rightarrow \mathbb{R}$, let

$$V(f) = \max \left\{ \sup_{K \in \mathbb{N}, t_0 \leq \dots \leq t_K \in \Theta} \sum_{k=1}^K (f(t_k) - f(t_{k-1}))_+, \sup_{K \in \mathbb{N}, t_0 \leq \dots \leq t_K \in \Theta} \sum_{k=1}^K (f(t_{k-1}) - f(t_k))_+ \right\}$$

denote the signed variation of the function, where we take $(x)_+ = \max\{x, 0\}$ to be the positive part of $x \in \mathbb{R}$. The signed variation is akin to the usual notion of total variation, but takes separate account of increasing and decreasing regions. For any class of functions $\mathcal{F}$, we let $V(\mathcal{F})$ denote the largest signed variation of all functions in the class, that is, $V(\mathcal{F}) = \sup_{f \in \mathcal{F}} V(f)$.

We suppose losses in $\mathcal{L}$, which we recall are bounded by λ , also have signed variation bounded by λ , $V(\mathcal{L}) \leq \lambda$. Bounded signed variation seems like a palatable restriction. Any loss function with $L(a, \theta) \in [0, \lambda]$ that is, for each $a \in \mathcal{A}$, either quasi-concave or quasi-convex in θ must have signed variation in θ bounded by λ . This admits a wide range of losses under which, for instance, there is a best or worst value $\theta(a)$ for each $a \in \mathcal{A}$, and the loss increases or decreases with distance from this value. It also admits monotone losses under which, for each $a \in \mathcal{A}$, the agent always prefers either higher or lower values of θ .

We likewise suppose that the density ratio functions in $\mathcal{W}$ are bounded, $\omega^{-1} \leq \mathcal{W} \leq \omega$, and that they have bounded signed variation, $V(\mathcal{W}) \leq (\omega - \omega^{-1})$. This class of density ratio functions is again permissive. Any density ratio function with $w(\theta) \in [\omega^{-1}, \omega]$ that is quasi-concave or quasi-convex in θ has signed variation bounded by $\omega - \omega^{-1}$. This admits density ratios that are monotone in the distance from some peak $\bar{\theta}$ or trough $\underline{\theta}$, such that the agent finds values near to $\bar{\theta}$ or far from $\underline{\theta}$ more likely than under the central prior. It also admits any monotone density ratio function such that the agent finds higher or lower values of θ more likely than under the central prior.

What bounded signed variation does rule out are losses and density ratio functions that oscillate repeatedly as θ changes, for instance functions that are highly cyclical in θ , or that depend sensitively on the parity (even or odd) of the digits in θ .

EXAMPLE—Get Out the Vote, Continued: Recall that the loss is $L(a, \theta^K) = G(\frac{V/M - a/\theta^K}{1-a})$, which is bounded by 1. Because for any $G(\cdot)$ the loss is monotone in θ^K for any a , this loss also has signed variation bounded by 1. For the signed variation of $L(a', \theta^K)$ to be greater than 1 for some $a' \in [0, 1]$ would require that $L(a', \theta^K)$ is non-monotone in θ^K , for example because a more effective phone campaign induces more turnout from the opponent's voters. And, it would require that $L(a', \theta^K)$ is not single-peaked in θ^K , so that, over the range of θ^K , the loss switches multiple times from increasing to decreasing in θ^K . Turning to the density ratios $w(\theta)$, because the density ratios implied by the priors in Figure 2 are quasi-concave, their signed variation is bounded above by $\omega - \omega^{-1}$. For the agent experiencing greatest regret, this bound is 3.8.

4.3. Equivalence to the Signed Kolmogorov Metric

The class of functions of bounded signed variation pairs with a convenient IPM. We first define this metric and then show its connection to the class of functions of bounded signed variation.

DEFINITION 1: For measures $\mu, \nu \in \Delta(\mathbb{R})$, the *signed Kolmogorov (SK) metric* is

$$d_{\text{SK}}(\mu, \nu) = \sup_{t \in \Theta} (\mu(\theta \leq t) - \nu(\theta \leq t))_+ + \sup_{t \in \Theta} (\nu(\theta \leq t) - \mu(\theta \leq t))_+,$$

where $\mu(\theta \leq t) = \mu(\{\theta \in \Theta : \theta \leq t\})$.

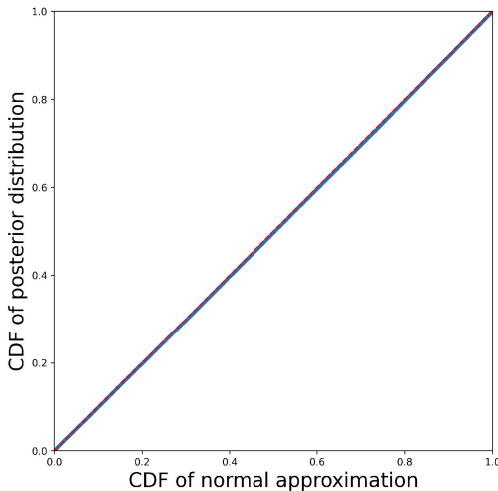
The signed Kolmogorov (SK) distance between two measures is found by adding together the largest positive and negative vertical distances between their respective CDFs.¹⁴ The SK metric can be read off of a p-p plot, which, for two measures μ, ν , is a plot of the CDF of one measure against the CDF of the other, that is, of $\{\mu(\theta \leq t), \nu(\theta \leq t)\}_{t \in \Theta}$.

EXAMPLE—Get Out the Vote, Continued: Figure 3a shows a p-p plot comparing the central posterior $\pi^*(\theta^{TC}|X)$ to the normal approximation $\pi^N(\theta^{TC}|X)$ for the effect θ^{TC} of a phone call. If the plot were on the 45-degree line, we would have $d_{\text{SK}}(\pi^*(\theta^{TC}|X), \pi^N(\theta^{TC}|X)) = 0$. In fact, the SK metric is small, at approximately 0.005.

Figure 3b shows a p-p plot comparing the central posterior $\pi^*(\theta^K|X)$ to the normal approximation $\pi^N(\theta^K|X)$ for the cost per vote θ^K . Here, the SK metric is larger, at approximately 0.052.

The SK metric is equivalent to the IPM generated by the class of non-negative functions with value and signed variation both bounded by 1. The following lemma shows this by extending a result of Müller (1997) for the Kolmogorov metric.

(a) Target parameter is the effect of a phone call



(b) Target parameter is the cost per vote

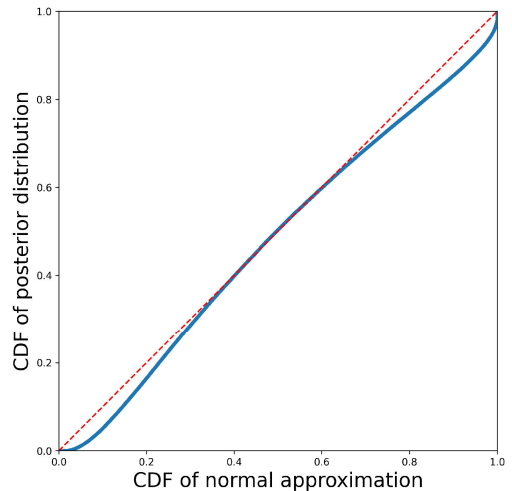


FIGURE 3.—Example p-p plots. Notes: Panel (a) shows a p-p plot comparing the central posterior $\pi^*(\theta^{TC}|X)$ to the normal approximation $\pi^N(\theta^{TC}|X)$ for the effect θ^{TC} of a phone call. Panel (b) shows a p-p plot comparing the central posterior $\pi^*(\theta^K|X)$ to the normal approximation $\pi^N(\theta^K|X)$ for the cost per vote θ^K . In both panels, the dashed line is the 45-degree line.

¹⁴This metric is distinct from what Filion (2015) terms the “signed Kolmogorov–Smirnov test.”

LEMMA 1: If $\mathcal{F}$ is the set of all functions $f : \mathbb{R} \rightarrow [0, 1]$ with $V(f) \leq 1$, then $d_{\mathcal{F}}(\mu, \nu) = d_{\text{SK}}(\mu, \nu)$ for any probability measures $\mu, \nu \in \Delta(\mathbb{R})$.

When the classes of losses $\mathcal{L}$ and density ratio functions $\mathcal{W}$ are bounded and of bounded signed variation, Lemma 1 immediately implies a connection between the SK metric and the IPMs that bound the regret in Proposition 1.

COROLLARY 2: Fix any probability measures $\mu, \nu \in \Delta(\Theta)$. If $0 \leq \mathcal{L} \leq \lambda$ and $V(\mathcal{L}) \leq \lambda$, then $d_{\mathcal{L}(\mathcal{A})}(\mu, \nu) \leq \lambda d_{\text{SK}}(\mu, \nu)$, with equality when $\mathcal{L}$ is otherwise unrestricted. If, further, $\omega^{-1} \leq \mathcal{W} \leq \omega$ and $V(\mathcal{W}) \leq (\omega - \omega^{-1})$, then $d_{\mathcal{W} \cdot \mathcal{L}(\mathcal{A})}(\mu, \nu) \leq \lambda(2\omega - \omega^{-1})d_{\text{SK}}(\mu, \nu)$.

Corollary 2 states that if the function classes $\mathcal{L}$ and $\mathcal{W}$ take the form we propose in Section 4.2, then the IPMs appearing in the bounds in Proposition 1 can be replaced with ones proportional to the SK metric. We now make use of this fact to specialize Procedure 1.

4.4. Procedure Based on the SK Metric

We propose the following specialization of Procedure 1 for the case where $\Theta \subseteq \mathbb{R}$.

Procedure 2: Controlling regret with the SK metric.

- | | |
|--|---|
| <ul style="list-style-type: none"> • Calculate <ul style="list-style-type: none"> – SK metric $d_{\text{SK}}(\pi^0(\theta X), \pi^*(\theta X))$ | <ul style="list-style-type: none"> • Report threshold τ and <ul style="list-style-type: none"> – $\hat{\pi}^*(\theta X) = \pi^0(\theta X)$ if $d_{\text{SK}} \leq \tau$ – $\hat{\pi}^*(\theta X) = \pi^*(\theta X)$ otherwise |
|--|---|

All other details follow Procedure 1.

We then have the following corollary of Proposition 1 and Lemma 1.

COROLLARY 3: Suppose that $\mathcal{L}$ and $\mathcal{W}$ satisfy the conditions of Corollary 2. Then if $\tau \leq \frac{\kappa}{2\lambda\omega(2\omega - \omega^{-1})}$, Procedure 2 satisfies Criterion (C1). Moreover, whenever Procedure 2 recommends reporting $\pi^*(\theta|X)$, if $\mathcal{L}$ is maximally rich, then there exists some agent in the audience whose regret from reporting $\pi^0(\theta|X)$ is at least $\lambda\tau$.

Because the upper bounds in Proposition 1 and Corollary 2 are not generally sharp, Procedure 2 will not generally satisfy both Criterion (C1) and Criterion (C2) for the same value of κ . However, because Proposition 1 establishes a lower bound on regret, if Procedure 2 recommends abandoning the conventional report, and the class of losses is sufficiently rich, then there is *some* agent whose regret from the conventional report is at least $\lambda\tau$.

EXAMPLE—Get Out the Vote, Continued: A loss function $L(a, \theta^K)$ maps the budget share $a \in [0, 1]$ and cost per vote $\theta^K > 0$ to the probability of defeat. A loss of 1 is certain defeat and a loss of 0 is certain victory. If all loss functions have signed variation bounded by 1 and the class $\mathcal{L}$ is otherwise unrestricted, then making the conventional report $\pi^N(\theta^K|X)$ about the cost per vote θ^K implies a willingness to increase some agent's probability of defeat by 0.052, relative to their posterior optimal action, in order to preserve the scientific convention.

5. A BOOTSTRAP SURROGATE FOR THE CENTRAL POSTERIOR

To facilitate adoption of Procedure 2, we would ideally like a default choice of central posterior $\pi^*(\theta|X)$ that is both convenient and reasonable for many economic settings. Because no one choice of posterior is likely exactly right for many different settings, we instead aim for a posterior that serves as a good approximation, in an appropriate sense, to posteriors from many reasonable priors. In this section, we first formalize the idea of such a surrogate posterior. We then propose to use a bootstrap distribution as a default surrogate posterior, and we describe two classes of posteriors for which its surrogacy is justified.

5.1. Use of a Surrogate Central Prior

Procedure 2 requires that the analyst calculate the central posterior. Suppose that the analyst has access instead to a *surrogate central posterior* $\pi^S(\theta|X)$ that is close to the central posterior in the sense that $d_{\text{SK}}(\pi^S(\theta|X), \pi^*(\theta|X)) \leq \delta$ for some small $\delta \in [0, 1]$. The triangle inequality implies that

$$|d_{\text{SK}}(\pi^0(\theta|X), \pi^S(\theta|X)) - d_{\text{SK}}(\pi^0(\theta|X), \pi^*(\theta|X))| \leq \delta.$$

Therefore, if the surrogate is far from the conventional report, then so is the central posterior. Moreover, under the conditions of Corollary 2, $d_{\mathcal{W}, \mathcal{L}(\mathcal{A})}(\pi^S(\theta|X), \pi^*(\theta|X)) \leq \lambda(2\omega - \omega^{-1})\delta$, so that by Proposition 1, reporting the surrogate controls regret even when the conventional report is not a good approximation to the central posterior.

We can therefore modify Procedure 2 as follows.

Procedure 3: Controlling regret with a surrogate central posterior.

- | | |
|--|---|
| <ul style="list-style-type: none"> • Calculate <ul style="list-style-type: none"> – surrogate posterior $\pi^S(\theta X)$ – SK metric $d_{\text{SK}}(\pi^0(\theta X), \pi^S(\theta X))$ | <ul style="list-style-type: none"> • Report threshold τ and <ul style="list-style-type: none"> – $\hat{\pi}^*(\theta X) = \pi^0(\theta X)$ if $d_{\text{SK}} \leq \tau$ – $\hat{\pi}^*(\theta X) = \pi^S(\theta X)$ otherwise |
|--|---|

All other details follow Procedure 1.

We then have the following corollary of Proposition 1 and Lemma 1.

COROLLARY 4: *Suppose that $\mathcal{L}$ and $\mathcal{W}$ satisfy the conditions of Corollary 2 and that $d_{\text{SK}}(\pi^S(\theta|X), \hat{\pi}^*(\theta|X)) \leq \delta$. Then Procedure 3 satisfies Criterion (C1) for $\tau \leq \frac{\kappa}{2\lambda\omega(2\omega - \omega^{-1})} - \delta$. Moreover, whenever Procedure 3 recommends reporting $\pi^S(\theta|X)$, if $\mathcal{L}$ is maximally rich, then there exists some agent in the audience whose regret from reporting $\pi^0(\theta|X)$ is at least $\lambda(\tau - \delta)$.*

5.2. Two Classes of Central Priors With a Bootstrap Surrogate Posterior

We now describe two classes of central priors for which a bootstrap can serve as a surrogate posterior. To do this, we specialize to the case where $X = (X_1, \dots, X_n) \in \mathcal{X}_0^n = \mathcal{X}$ consists of $n \in \mathbb{N}$ i.i.d. draws from some (possibly unknown) probability distribution $P \in \Delta(\mathcal{X}_0)$. We further suppose that $\theta = \theta(P)$ for $\theta(\cdot)$ a known functional, so that θ is

identified from the distribution of the data, as with a point-identified parameter, pseudotrue value from a GMM procedure, or bound on an identified set. A prior π specifies a distribution for P , which in turn implies a distribution for θ . We let $P_n \in \Delta(\mathcal{X}_0)$ denote the empirical distribution of the sample.

5.2.1. Dirichlet Process Central Priors

Suppose that $\pi_\alpha^*(P) = \text{DP}(\alpha, Q)$ where $\text{DP}(\cdot, \cdot)$ denotes a Dirichlet process (e.g., Ghosal and van der Vaart (2017, Chapter 4)), $\alpha > 0$ is a scalar, and $Q \in \Delta(\mathcal{X}_0)$ is a distribution for a single observation X_i . The parameter α controls the informativeness of the prior, with a smaller value corresponding to a less informative prior. The parameter Q controls the central tendency of the prior, and in particular corresponds to the marginal distribution of a single observation X_i under the prior.

EXAMPLE—Get Out the Vote, Continued: An observation $X_i \in \{0, 1\}^2$ encodes both whether citizen i votes and whether citizen i is assigned to receive a phone call. A measure Q describes a contingency table for these two events. Any given sample X implies a sample contingency table P_n . A prior is a distribution over contingency tables.

A prior of the form $\pi_\alpha^*(P) = \text{DP}(\alpha, Q)$ implies a posterior of the form

$$\pi_\alpha^*(P|X) = \text{DP}\left(\alpha + n, \frac{\alpha}{\alpha + n}Q + \frac{n}{\alpha + n}P_n\right).$$

The centering measure of the posterior distribution is a weighted average of the prior contingency table Q and the sample contingency table P_n . As the informativeness of the prior becomes large, $\alpha \rightarrow \infty$, the posterior centers on the prior contingency table. As the informativeness of the prior becomes small, $\alpha \rightarrow 0$, the posterior centers on the sample contingency table.

Just as the informativeness of the prior is controlled by the parameter α , the informativeness of the sample is controlled by the sample size n . As the informativeness of the sample becomes large, $n \rightarrow \infty$, the posterior centers on the sample contingency table, and with increasing precision, so that draws from the posterior become increasingly concentrated on the observed distribution of the data. As the informativeness of the sample becomes small, $n \rightarrow 0$, the posterior converges to the prior distribution. As the informativeness of the prior also becomes small, $\alpha \rightarrow 0$, the prior distribution remains centered on the prior contingency table, but becomes less concentrated, so that draws from the prior typically place almost all mass in one cell (e.g., almost all citizens are treated and vote), with the location of that cell drawn randomly from the prior contingency table.

The *Bayes bootstrap distribution* is defined as $\pi^B(P|X) = \text{DP}(n, P_n)$. The Bayes bootstrap distribution can be conveniently sampled by calculating the functional $\theta(\cdot)$ on data X whose observations have been reweighted by a standard Dirichlet (Rubin (1981)). Well-known properties of this bootstrap, together with conjugacy properties of Dirichlet processes, imply that the Bayes bootstrap distribution is a surrogate posterior for a Dirichlet process central prior with small informativeness α .

LEMMA 2: If $\pi_\alpha^*(P) = \text{DP}(\alpha, Q)$ for some Q , $\theta(P)$ is continuous with respect to convergence in distribution almost everywhere in the support of $\pi^B(P|X)$, and either (i) $\pi^B(\theta|X)$ is a continuous distribution or (ii) the range of $\theta(P)$ is countable, then

$$\lim_{\alpha \rightarrow 0} d_{\text{SK}}(\pi^B(\theta|X), \pi_\alpha^*(\theta|X)) = 0.$$

Lemma 2 states that, under continuity conditions, the Bayes bootstrap distribution for θ is close, in SK metric, to the central posterior $\pi_\alpha^*(\theta|X)$ when the corresponding central prior $\pi_\alpha^*(P)$ is not very informative.

Lemma 2 justifies Procedure 3 under Dirichlet process central priors $\pi_\alpha^*(P) = \text{DP}(\alpha, Q)$ with small α . We think this justification is reasonable for situations where audience members have limited prior information about the target parameter, and so are happy to adopt a nonparametric prior with low informativeness. For other situations, we turn to an alternative justification for the use of $\pi^B(P|X)$ as a surrogate.

5.2.2. Central Posteriors That Agree on a BvM Parameter

Suppose that $\theta(P) = \theta(\beta(P))$ for known functions $\beta : \Delta(\mathcal{X}_0) \rightarrow \mathcal{B}$ and $\theta : \mathcal{B} \rightarrow \mathbb{R}$, where $\mathcal{B}$ may be infinite-dimensional.¹⁵ On its own, this restriction has no content, as we can always take $\beta(P) = P$ or $\beta(P) = \theta(P)$.

EXAMPLE—Get Out the Vote, Continued: The parameter $\beta = (\beta^T, \beta^C)$ describes the probability of voting with and without a phone call. The target parameters $\theta^{TC}(\beta) = \beta^T - \beta^C$ and $\theta^K(\beta) = K/(\beta^T - \beta^C)$ can both be written as functions of β .

If the central posterior $\pi^*(\beta|X)$ and Bayes bootstrap distribution $\pi^B(\beta|X)$ are close in a metric which implies closeness of $\pi^*(\theta|X)$ and $\pi^B(\theta|X)$ in SK, then the Bayes bootstrap distribution can serve as a surrogate. To state this formally, for a class of functions $\mathcal{F}$ let $\mathcal{F} \circ \theta$ denote the class of functions of the form $f \circ \theta = f(\theta(\cdot))$ for $f \in \mathcal{F}$.

LEMMA 3: If $d_G(\pi^B(\beta|X), \pi^*(\beta|X)) \leq \delta$ for a class of functions $\mathcal{G}$, and $\mathcal{F} \circ \theta \subseteq \mathcal{G}$ for $\mathcal{F}$ the class of functions with value and signed variation bounded by 1, that is, the maximal class satisfying $0 \leq \mathcal{F} \leq 1$ and $V(\mathcal{F}) \leq 1$, it follows that $d_{\text{SK}}(\pi^B(\theta|X), \pi^*(\theta|X)) \leq \delta$.

Agreement on an underlying parameter β can, in turn, be justified by BvM theory. BvM theory provides conditions on the central prior π^* , parameters β , and the function class $\mathcal{G}$ under which $\pi^*(\beta|X)$ approaches the conventional normal approximation in the metric d_G as n grows large. When β is finite-dimensional, standard results require only that $\mathcal{G}$ is bounded (see, e.g., van der Vaart (1998, Chapter 10); Ghosal and van der Vaart (2017, Chapter 12.3)). When β is infinite-dimensional, existing results further require that $\mathcal{G}$ have a bounded Lipschitz constant (see, e.g., Castillo and Nickl (2014)).

EXAMPLE—Get Out the Vote, Continued: For $\beta = (\beta^T, \beta^C)$, the left plot in Supplemental Appendix Figure 1 shows that the central posterior $\pi^*(\beta|X)$ is visually similar to the conventional normal approximation $N(\hat{\beta}(X), \hat{\Sigma}_{\hat{\beta}}(X))$.

Importantly, BvM theory can also be applied to the Bayes bootstrap distribution. Supplemental Appendix C shows a variety of conditions under which the hypotheses of Lemma 3 hold with probability tending to 1 as n grows large.

EXAMPLE—Get Out the Vote, Continued: The right plot in Supplemental Appendix Figure 1 shows that the Bayes bootstrap distribution $\pi^B(\beta|X)$ is visually similar to the conventional normal approximation $N(\hat{\beta}(X), \hat{\Sigma}_{\hat{\beta}}(X))$.

¹⁵Formally, $\theta(P) = \theta_\beta(\beta(P))$, where $\theta : \Delta(\mathcal{X}_0) \rightarrow \mathbb{R}$ and $\theta_\beta : \mathcal{B} \rightarrow \mathbb{R}$. We suppress the distinction between $\theta(\cdot)$ and $\theta_\beta(\cdot)$ in the main text to ease notation.

5.3. Procedure With a Bootstrap Surrogate

The preceding arguments suggest the following refinement to Procedure 3.

Procedure 4: Controlling regret with Bayes bootstrap and a conventional normal report.

- **Calculate**

- Bayes bootstrap distribution $\pi^B(\theta|X)$
- Conventional normal report $\pi^0(\theta|X) = \pi^N(\theta|X)$
- SK metric $d_{\text{SK}}(\pi^N(\theta|X), \pi^B(\theta|X))$

- **Report** threshold τ and

- $\hat{\pi}^*(\theta|X) = \pi^N(\theta|X)$ if $d_{\text{SK}} \leq \tau$
- $\hat{\pi}^*(\theta|X) = \pi^B(\theta|X)$ otherwise

All other details follow Procedure 1.

COROLLARY 5: Suppose that $\mathcal{L}$ and $\mathcal{W}$ satisfy the conditions of Corollary 2 and that either

- (i) $\pi_\alpha^*(P) = \text{DP}(\alpha, Q)$ for α such that $d_{\text{SK}}(\pi^B(\theta|X), \pi_\alpha^*(\theta|X)) \leq \delta$ or
- (ii) $d_{\mathcal{F} \circ \theta}(\pi^B(\beta|X), \pi^*(\beta|X)) \leq \delta$ for $\mathcal{F}$ the maximal class satisfying $0 \leq \mathcal{F} \leq 1$ and $V(\mathcal{F}) \leq 1$.

Then Procedure 4 satisfies Criterion (C1) for $\tau \leq \frac{\kappa}{2\lambda\omega(2\omega-1)} - \delta$. Moreover, whenever Procedure 4 recommends reporting $\pi^B(\theta|X)$, if $\mathcal{L}$ is maximally rich, then there exists some agent in the audience whose regret from reporting $\pi^N(\theta|X)$ is at least $\lambda(\tau - \delta)$.

EXAMPLE—Get Out the Vote, Continued: Figure 4a shows a p-p plot comparing the Bayes bootstrap distribution $\pi^B(\theta^{TC}|X)$ to the normal approximation $\pi^N(\theta^{TC}|X)$ for the effect of the intervention. The SK metric is small, at approximately 0.003. Figure 4b shows a p-p plot comparing the Bayes bootstrap distribution $\pi^B(\theta^K|X)$ to the normal approximation $\pi^N(\theta^K|X)$ for the cost per vote. The SK metric is larger, at approximately 0.052. The conclusions from Figure 4, which uses the bootstrap as a surrogate, are therefore similar to those from Figure 3, which uses the central posterior.

Notice that here the Bayes bootstrap distribution is a surrogate for the central posterior but not necessarily for all agents' posteriors. Nevertheless, Procedure 4 allows that agents may reweight the Bayes bootstrap distribution as if it were the central posterior. To us, central priors of the form envisioned in Lemma 3 seem easier to contemplate, and hence to reweight, than those of the form envisioned in Lemma 2. We therefore think that the justification via Lemma 3 is more appealing for situations where the audience has substantive beliefs about the parameter of interest.

REMARK 2: We recommend using the Bayes bootstrap distribution because it admits a direct interpretation following Lemma 2. Weng (1989) shows a sense in which the nonparametric bootstrap approximates the posterior distribution corresponding to a Dirichlet process prior in a large sample, albeit not as accurately as does the Bayes bootstrap. Supplemental Appendix D shows that, under the same BvM conditions that can make the Bayes bootstrap a useful surrogate for the central posterior, the Bayes bootstrap distribution will be close to that of other weighted bootstraps, including the nonparametric bootstrap. The analyst may therefore alternatively substitute the nonparametric bootstrap as a surrogate.

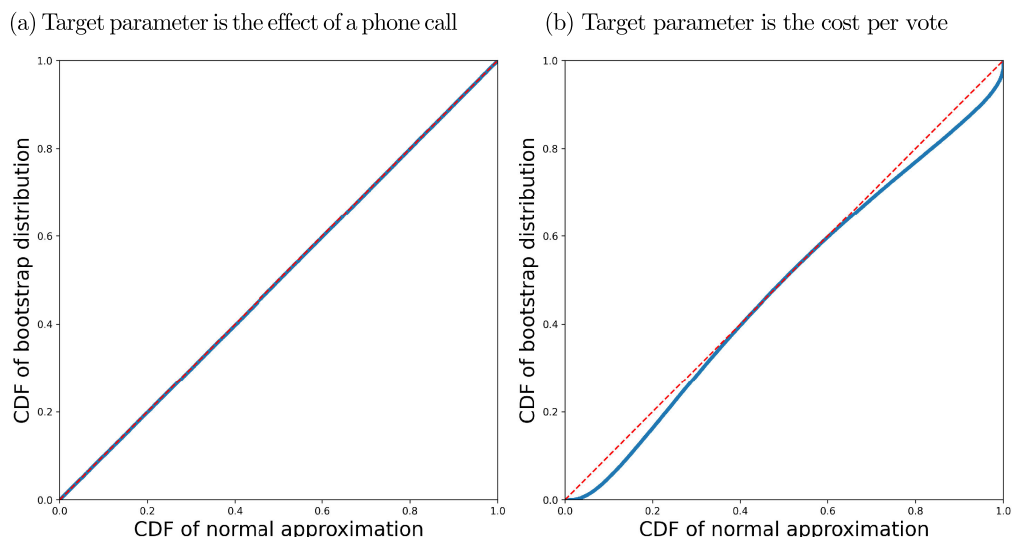


FIGURE 4.—Comparing the Bayes bootstrap distribution to the normal approximation. *Notes:* Panel (a) shows a p-p plot comparing the Bayes bootstrap distribution $\pi^B(\theta^{TC}|X)$ to the normal approximation $\pi^N(\theta^{TC}|X)$ for the effect θ^{TC} of a phone call. Panel (b) shows a p-p plot comparing the Bayes bootstrap distribution $\pi^B(\theta^K|X)$ to the normal approximation $\pi^N(\theta^K|X)$ for the cost per vote θ^K . In both panels, the dashed line is the 45-degree line.

6. APPLICATION TO A BOOTSTRAP CENSUS

Procedure 4 is applicable to a wide range of economic settings. To demonstrate the procedure’s applicability, we applied it to the papers in the 2021 *American Economic Review* that use a bootstrap (Andrews and Shapiro (2025)). We now describe our methods and findings.

6.1. Census of Papers Using a Bootstrap

We used a Google Scholar query (https://scholar.google.com/scholar?q=bootstrap+source%3AAmerican+source%3AEconomic+source%3AReview+&hl=en&as_sdt=0%2C10&as_ylo=2021&as_yhi=2021) to identify papers published in the *American Economic Review* in 2021 that use a bootstrap. For each paper, we identified the main objects of interest, which we define to be objects for which a quantitative or a qualitative description appears in the abstract or introduction. We focus on objects of interest for which the bootstrap is used for inference. We excluded from our census papers that are primarily methodological, papers that use the bootstrap only to calculate a p -value, or papers that use a bootstrap exclusively for objects reported in appendices. Supplemental Appendix Table 1 lists the papers we include in our census along with the number of objects of interest in each paper that we include in our analysis.

For each paper, we attempted to reproduce the bootstrap replicates for all objects of interest using the published replication code and data. When this was not feasible (e.g., due to confidential data), we contacted the authors to request the bootstrap replicates. We succeeded in obtaining bootstrap replicates covering 81 objects of interest across 14 papers, with only 1 paper for which we were unable to obtain the replicates.

Consistent with the wide applicability of Procedure 4, the papers in the census cover a range of topics including public economics, labor economics, macroeconomics, behavioral

economics, industrial organization, and development economics. The objects of interest include parameters describing technology, welfare calculations from a structural model, impulse responses, and transformations of regression coefficients.

Our main calculations will use the replicates that we obtained from the replication materials or from the authors. Given these replicates, applying Procedure 4 is trivial and can be done in a web app (<https://jmslab.shinyapps.io/bootstrapreport/>).

Although our theoretical development focuses on the Bayes bootstrap, only 1 paper in the census actually uses the Bayes bootstrap. Among the others, the most popular form of bootstrap is the nonparametric bootstrap (10 articles including block bootstraps), followed by various forms of parametric bootstrap (3 articles).¹⁶ As we discuss in Section 5, there are reasons to expect different bootstrap procedures to yield similar distributions in large samples. For a subset of articles using a nonparametric bootstrap, we were able to use the published replication code and data to compute a Bayes bootstrap. In no case do we reject that the SK distance to the default normal differs between the two bootstrap distributions; see Supplemental Appendix Figure 2 for details.

Applying Procedure 4 requires specifying a conventional report. For our main analysis, we use the normal approximation $\pi^0(\theta|X) = \pi^N(\theta|X) = N(\hat{\theta}(X), \hat{\Sigma}(X))$ implied by the reported point estimate and the bootstrap standard error, defined as the standard deviation of the bootstrap distribution. All of the papers in the census report point estimates for all objects of interest, and 10 out of 14 report bootstrap standard errors. The remaining 4 out of 14 papers report bootstrap confidence intervals. To better capture such situations, Supplemental Appendix Figure 3 shows results where we take the variance of the default normal to match the difference between the 97.5th and 2.5th percentiles of the bootstrap distribution, rather than its standard deviation.

6.2. Findings From the Bootstrap Census

Figure 5 shows a series of p-p plots comparing the CDF of bootstrap replicates, that is, samples from $\pi^B(\theta|X)$, to the conventional normal report $\pi^N(\theta|X)$. Each plot reports the maximum positive and negative vertical distance between the CDF of the bootstrap replicates and of the conventional report. The sum of these distances is an estimator of $SK(\pi^B(\theta|X), \pi^N(\theta|X))$ that is consistent in the number of bootstrap replicates. For each paper in our census, we depict two p-p plots, corresponding to the objects with the smallest and largest estimated SK metric among the objects of interest in the paper. The figure shows that, as in our running example, the estimated SK metric can differ meaningfully even across objects of interest reported in the same paper.

Across all of the objects of interest in our census, we find that the estimated SK metric is greater than 0.1 in 72 percent of cases, and is greater than 0.2 in 32 percent of cases. In these cases, if the Bayes bootstrap is a good surrogate for the central posterior, then under the conditions of Corollary 4, the value of the conventional report must represent a share of at least 0.1 or 0.2, respectively, of the worst-case loss to justify reporting the conventional normal report. Supplemental Appendix Figure 3 shows the full distribution of the estimated SK metric across the objects of interest in our census.

¹⁶We include in the category of parametric bootstraps procedures that treat some statistics as exactly normally distributed. In all but 1 of the papers in the census, the authors' bootstrap procedure implicitly treats the data as i.i.d. across some observed units.

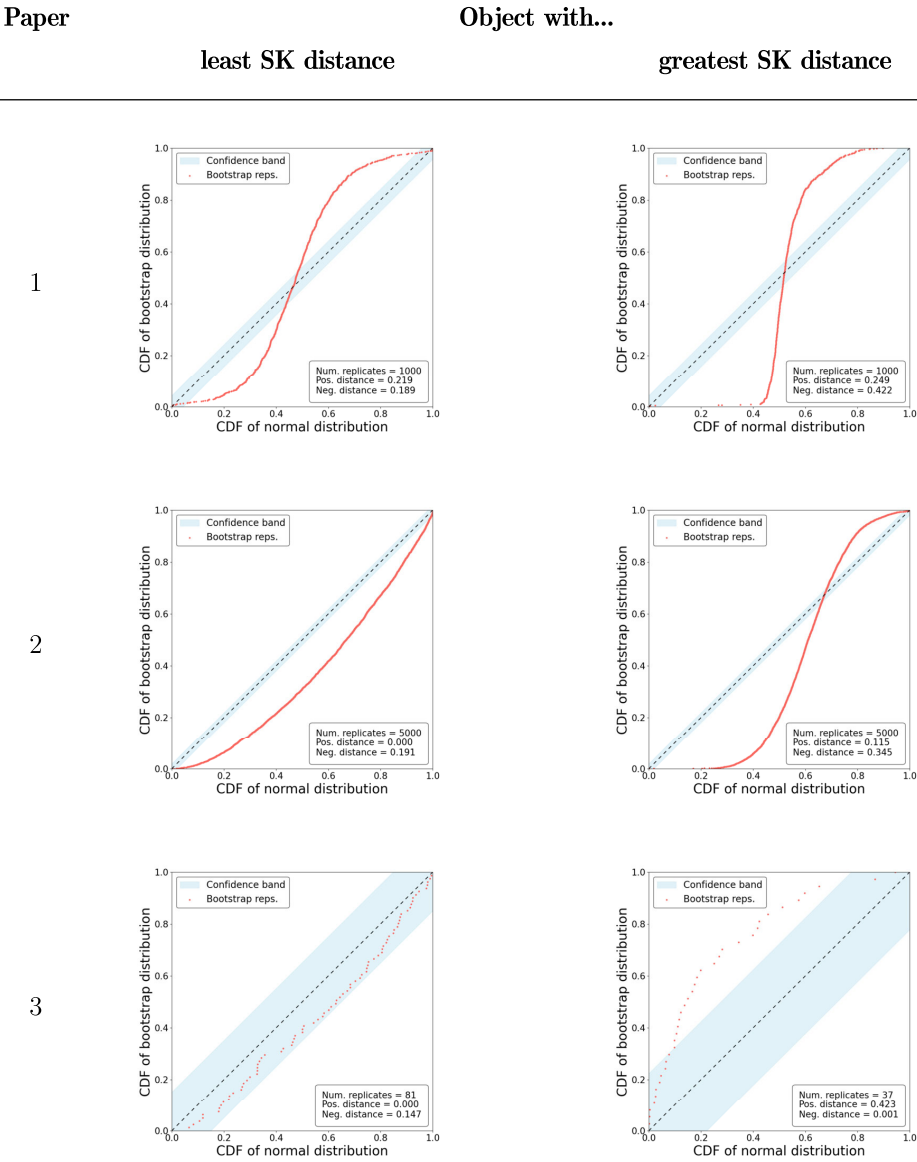


FIGURE 5.—Comparison of bootstrap distribution to default normal report. *Notes:* Each row corresponds to a paper in our bootstrap census. Each plot is a p-p plot comparing the distribution of the bootstrap replicates to the distribution of the default normal report whose mean is given by the point estimate and whose standard deviation is given by the bootstrap standard error. The shaded region is a uniform confidence band designed to contain the empirical CDF of the bootstrap replicates with probability at least 0.95 whenever the true bootstrap distribution is given by the default normal report. Each plot legend reports the maximum positive and negative vertical distances between the two distributions. Each row includes two plots, one for the object of interest with the smallest sum of maximum positive and negative distances (“least SK distance”) and one for the object of interest with the largest sum of distances (“greatest SK distance”). Rows (papers) are in descending order according to their greatest SK distance.

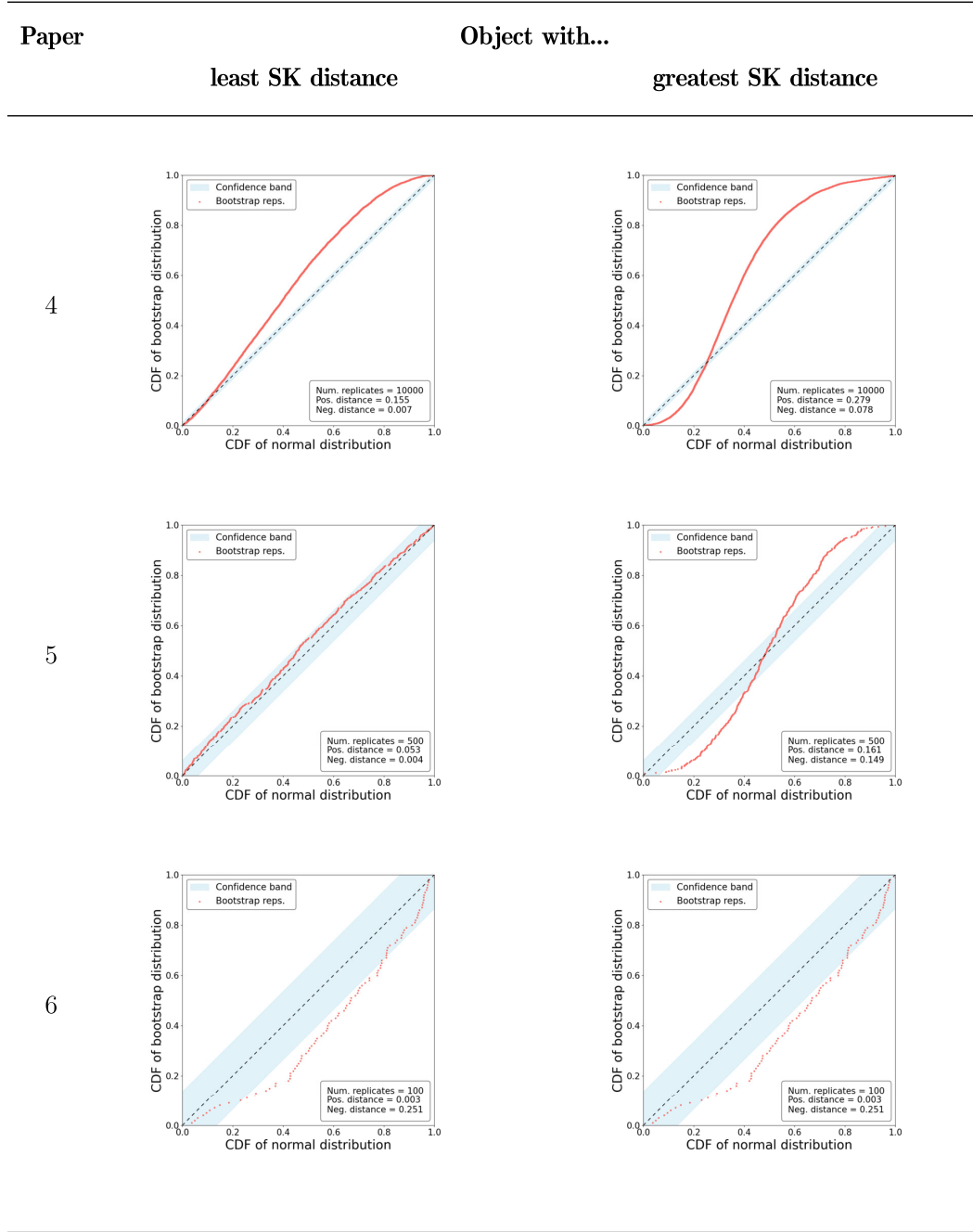


FIGURE 5.—Continued.

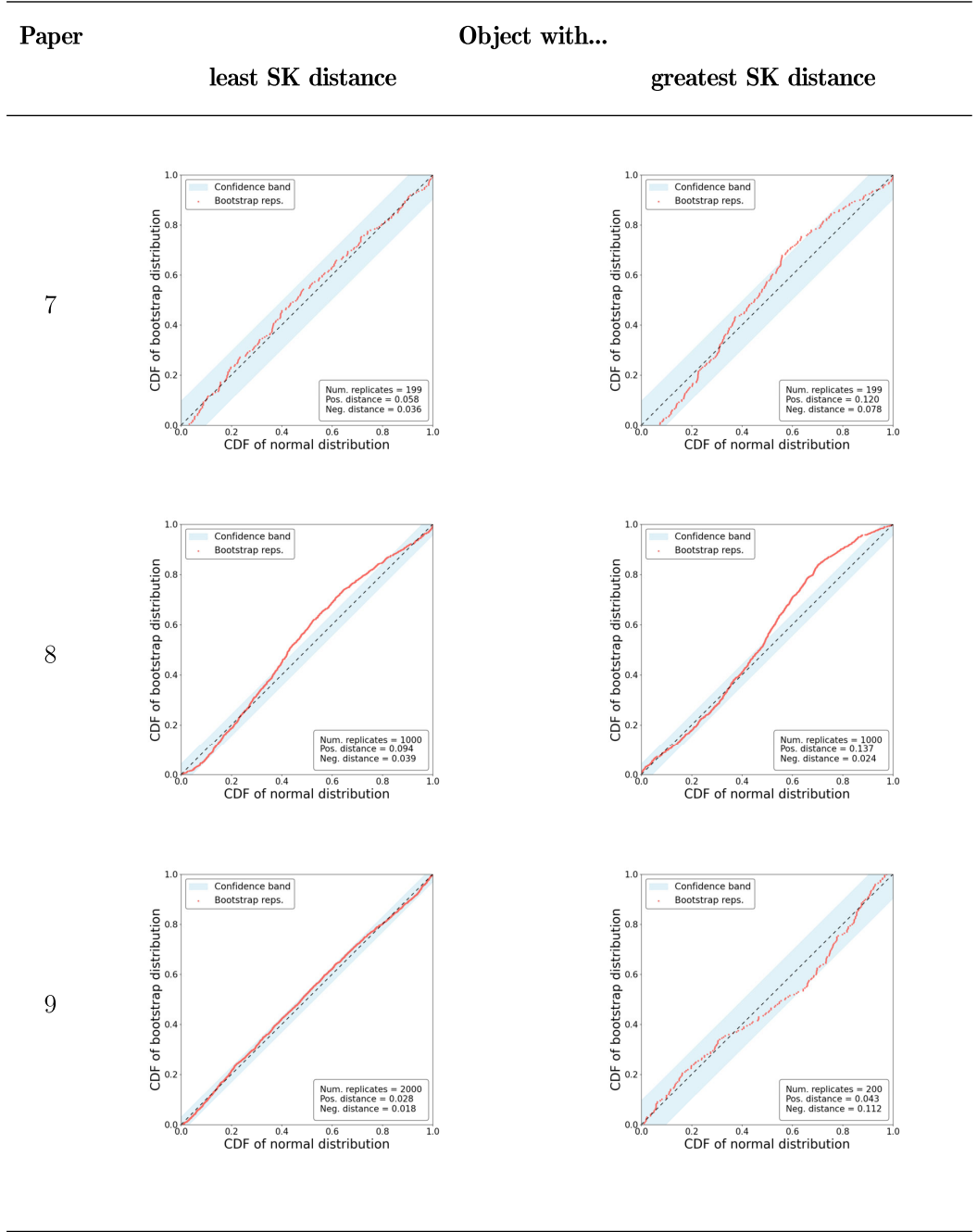


FIGURE 5.—Continued.

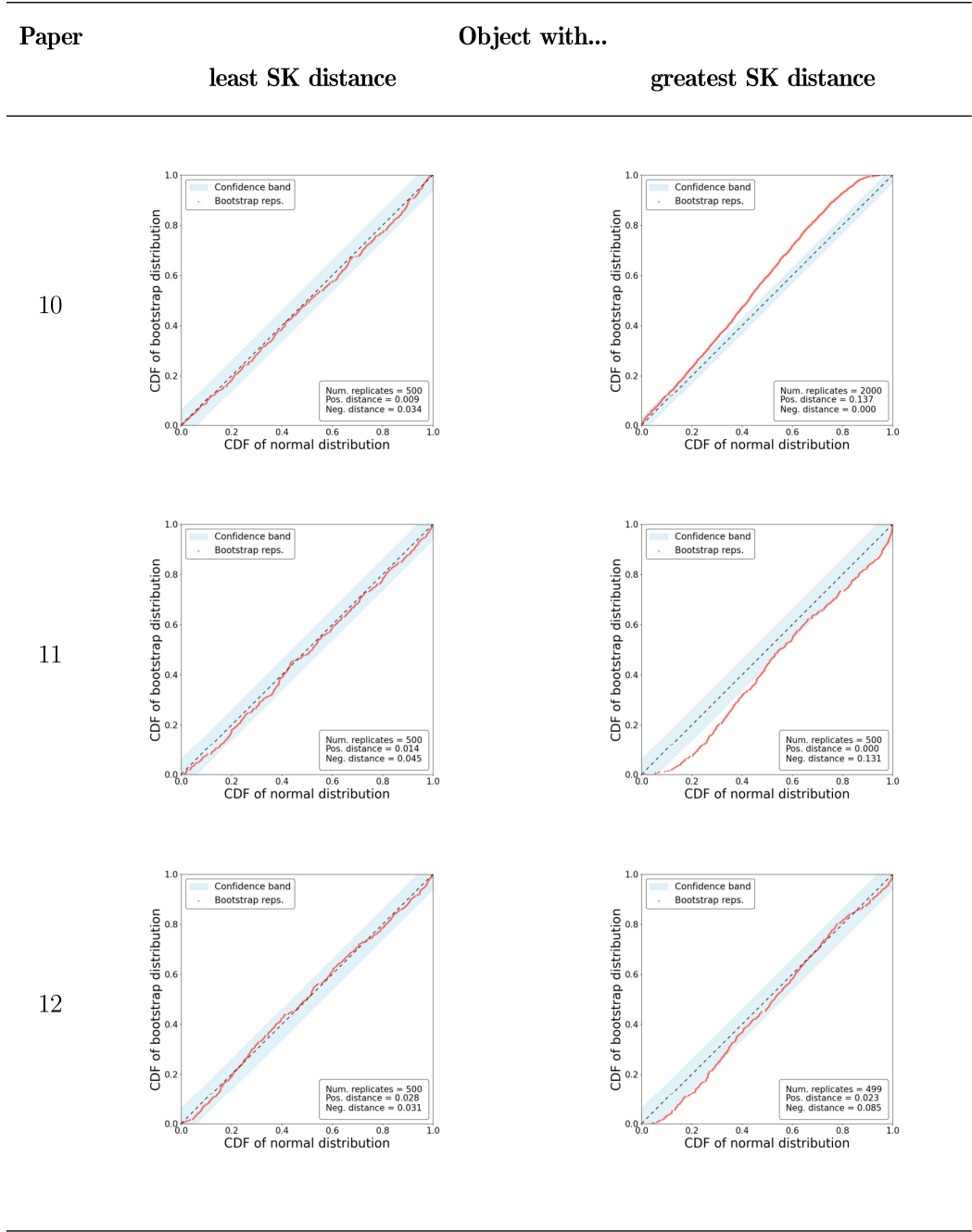


FIGURE 5.—Continued.

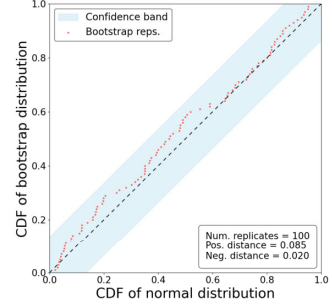
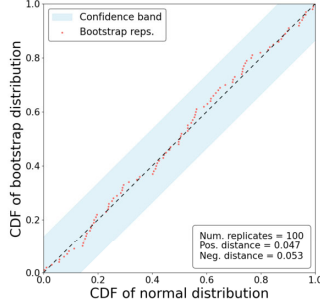
Paper

Object with...

least SK distance

greatest SK distance

13



14

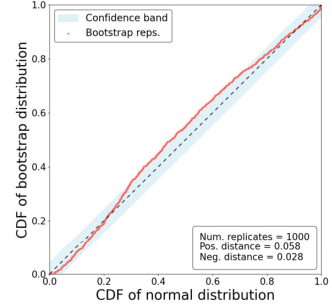
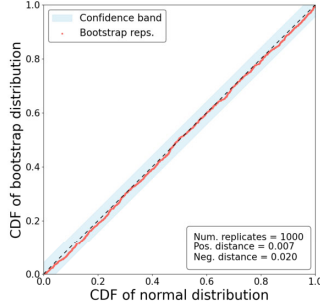


FIGURE 5.—Continued.

APPENDIX: PROOFS

To extend Theorem 1 to the case where an optimal action may not exist, we need to define regret for this case. For $\varepsilon > 0$, define the set of ε -optimal actions under (π, L) , $\mathcal{A}_\varepsilon(\pi, L) = \{\hat{a} \in \mathcal{A} : E_\pi[L(\hat{a}, \theta)] \leq \inf_{a \in \mathcal{A}} E_\pi[L(a, \theta)] + \varepsilon\}$. We consider the worst-case regret over ε -optimal actions as $\varepsilon \rightarrow 0$,

$$R(\hat{\pi}^*(\theta|X); X, L, w, \pi^*) = \lim_{\varepsilon \rightarrow 0} \left(\sup \left\{ E_{w(\theta)\pi^*(\theta|X)}[L(\hat{a}, \theta) - L(a, \theta)] : \begin{array}{l} \hat{a} \in \mathcal{A}_\varepsilon(w(\cdot)\hat{\pi}^*(\cdot|X), L), a \in \mathcal{A}, \\ E_{w(\theta)\hat{\pi}^*(\theta|X)}[L(\hat{a}, \theta) - L(a, \theta)] \leq 0 \end{array} \right\} \right).$$

THEOREM 2: For R as defined above,

$$\sup_{w \in \mathcal{W}} \sup_{L \in \mathcal{L}} R(\hat{\pi}^*(\theta|X); X, L, w, \pi^*) \leq \bar{R}(\hat{\pi}^*(\theta|X); X, \mathcal{L}, \mathcal{W}, \pi^*)$$

for $\bar{R}$ as defined in Theorem 1.

PROOF OF THEOREM 2: For any $L \in \mathcal{L}$ and $w \in \mathcal{W}$, consider any pair of actions $\hat{a}$ and a such that the agent weakly prefers $\hat{a}$ to a under $\hat{\pi}$, $E_{w(\theta)\hat{\pi}^*(\theta|X)}[L(\hat{a}, \theta) - L(a, \theta)] \leq 0$. Note that since $L(a, \theta), L(\hat{a}, \theta) \in \mathcal{L}(\mathcal{A})$, it is immediate that $E_{\pi(\theta|X)}[L(\hat{a}, \theta) - L(a, \theta)]$ is bounded above by

$$\sup_{\ell_1, \ell_0 \in \mathcal{L}(\mathcal{A}), w \in \mathcal{W}} E_{w(\theta)\pi^*(\theta|X)}[\ell_1(\theta) - \ell_0(\theta)] \quad \text{subject to: } E_{w(\theta)\hat{\pi}^*(\theta|X)}[\ell_1(\theta) - \ell_0(\theta)] \leq 0.$$

Hence, we see that for any $\varepsilon > 0$ and any $\hat{a} \in \mathcal{A}_\varepsilon(\hat{\pi}, L)$, $a \in \mathcal{A}$ such that $E_{w(\theta)\hat{\pi}^*(\theta|X)}[L(\hat{a}, \theta) - L(a, \theta)] \leq 0$,

$$E_{w(\theta)\pi^*(\theta|X)}[L(\hat{a}, \theta) - L(a, \theta)] \leq \bar{R}(\hat{\pi}^*(\theta|X); X, \mathcal{L}, \mathcal{W}, \pi^*(\theta|X)),$$

from which the result is immediate. Q.E.D.

PROOF OF THEOREM 1: The first part of the result is immediate from Theorem 2. For the second part of the result, note that for any $\ell_1, \ell_0 \in \mathcal{L}(\mathcal{A})$, $w \in \mathcal{W}$ such that

$$E_{w(\theta)\hat{\pi}^*(\theta|X)}[\ell_1(\theta) - \ell_0(\theta)] \leq 0, \tag{1}$$

if we consider $\mathcal{A} = \{0, 1\}$ and the loss

$$L(a, \theta) = \begin{cases} \ell_1(\theta) & \text{if } a = 1, \\ \ell_0(\theta) & \text{if } a = 0, \end{cases}$$

then the agent weakly prefers $\hat{a} = 1$ based on their perceived posterior, but will incur regret $E_{w(\theta)\pi^*(\theta|X)}[\ell_1(\theta) - \ell_0(\theta)]$ from this action. Hence, if we take the worst case over $\ell_1, \ell_0 \in \mathcal{L}(\mathcal{A})$, $w \in \mathcal{W}$ satisfying (1), we obtain risk $\bar{R}(\hat{\pi}^*(\theta|X); X, \mathcal{L}, \mathcal{W}, \pi^*)$, where the conditions of the theorem ensure that this worst case is attained at some $\bar{\ell}_1, \bar{\ell}_0, \bar{w}$, and we may take

$$\bar{L} = \begin{cases} \bar{\ell}_1(\theta) & \text{if } a = 1, \\ \bar{\ell}_0(\theta) & \text{if } a = 0. \end{cases}$$

Note that $\bar{L}(\mathcal{A}) \subseteq \mathcal{L}(\mathcal{A})$, but we may have $\bar{L} \notin \mathcal{L}$. If, contrary to our assumptions, the set

$$\{E_{w(\theta)\pi^*(\theta|X)}[\ell_1(\theta) - \ell_0(\theta)], E_{w(\theta)\hat{\pi}^*(\theta|X)}[\ell_1(\theta) - \ell_0(\theta)] : \ell_1, \ell_0 \in \mathcal{L}(\mathcal{A}), w \in \mathcal{W}\}$$

is not compact, we can select $\ell_1, \ell_0 \in \mathcal{L}(\mathcal{A})$, $w \in \mathcal{W}$ to come arbitrarily close to the worst-case regret $\bar{R}(\hat{\pi}^*(\theta|X); X, \mathcal{L}, \mathcal{W}, \pi^*)$. Q.E.D.

PROOF OF COROLLARY 1: If $\tau \leq \kappa$, Procedure 1 reports $\pi^0(\theta|X)$ only if $\bar{R}(\pi^0(\theta|X); X, \mathcal{L}, \mathcal{W}, \pi^*) \leq \kappa$. Criterion (C1) is thus satisfied by Theorem 1. If $\tau \geq \kappa$ and $\bar{L} \in \mathcal{L}$, Procedure 1 reports $\pi^*(\theta|X)$ only if $\bar{R}(\pi^0(\theta|X); X, \mathcal{L}, \mathcal{W}, \pi^*) > \kappa$. Criterion (C2) is thus satisfied by Theorem 1. By Theorem 1, if $\bar{L} \in \mathcal{L}$, any procedure satisfying Criteria (C1) and (C2) must report $\pi^0(\theta|X)$ if and only if $\bar{R}(\pi^0(\theta|X); X, \mathcal{L}, \mathcal{W}, \pi^*) \leq \kappa$, and thus is equivalent to Procedure 1 with $\tau = \kappa$. Q.E.D.

PROOF OF PROPOSITION 1: Since $w(\theta) \geq \omega^{-1}$ by assumption, for any $w \in \mathcal{W}$ and any $\ell_1, \ell_0 \in \mathcal{L}(\mathcal{A})$, $\int w(\theta) d\pi^*(\theta|X) \geq \omega^{-1}$ and

$$\mathbb{E}_{w(\theta)\pi^*(\theta|X)}[\ell_1(\theta) - \ell_0(\theta)] \leq \omega \int (\ell_1(\theta) - \ell_0(\theta))w(\theta) d\pi^*(\theta|X).$$

Note, next, that $\mathbb{E}_{w(\theta)\hat{\pi}^*(\theta|X)}[\ell_1(\theta) - \ell_0(\theta)] \leq 0$ implies that

$$\begin{aligned} & \omega \int (\ell_1(\theta) - \ell_0(\theta))w(\theta) d\pi^*(\theta|X) \\ & \leq \omega \left(\int (\ell_1(\theta) - \ell_0(\theta))w(\theta) d\pi^*(\theta|X) - \int (\ell_1(\theta) - \ell_0(\theta))w(\theta) d\hat{\pi}^*(\theta|X) \right) \\ & \leq 2\omega \cdot \sup_{f \in \mathcal{W} \cdot \mathcal{L}(\mathcal{A})} \left| \int f(\theta) d\pi^*(\theta|X) - \int f(\theta) d\hat{\pi}^*(\theta|X) \right| \\ & = 2\omega \cdot d_{\mathcal{W} \cdot \mathcal{L}(\mathcal{A})}(\hat{\pi}^*(\theta|X), \pi^*(\theta|X)). \end{aligned}$$

Since this holds for all $w \in \mathcal{W}$ and any $\ell_1, \ell_0 \in \mathcal{L}(\mathcal{A})$ with $\mathbb{E}_{w(\theta)\hat{\pi}^*(\theta|X)}[\ell_1(\theta) - \ell_0(\theta)] \leq 0$, the first part of the result is immediate from Theorem 1.

For the second part of the result, centrosymmetry implies that

$$\begin{aligned} & \sup_{f \in \mathcal{L}(\mathcal{A})} \left| \int f(\theta) d\pi^*(\theta|X) - \int f(\theta) d\hat{\pi}^*(\theta|X) \right| \\ & = \sup_{f \in \mathcal{L}(\mathcal{A})} \left(\int f(\theta) d\pi^*(\theta|X) - \int f(\theta) d\hat{\pi}^*(\theta|X) \right). \end{aligned}$$

Consider $\tilde{f}$ which attains the preceding supremum (which exists by our compactness assumptions), $\mathcal{A} = \{0, 1\}$, and the loss

$$L(a, \theta) = \begin{cases} \tilde{f}(\theta) & \text{if } a = 1, \\ \int \tilde{f}(\theta) d\hat{\pi}^*(\theta|X) & \text{if } a = 0. \end{cases}$$

By construction, the agent is indifferent between the two actions based on their perceived posterior, but choosing $\hat{a} = 1$ yields a regret of

$$\int \tilde{f}(\theta) d\pi^*(\theta|X) - \int \tilde{f}(\theta) d\hat{\pi}^*(\theta|X) \geq d_{\mathcal{L}(\mathcal{A})}(\pi^*(\theta|X), \hat{\pi}^*(\theta|X)).$$

The result is immediate. Q.E.D.

PROOF OF LEMMA 1: We first show that for $\mathcal{I}$ the set of indicator functions for (open, closed, and half-open) intervals in $\mathbb{R}$, $d_{\mathcal{I}}(\mu, \nu) = d_{\text{SK}}(\mu, \nu)$.

To see that this is the case, note that since SK is a metric, $|d_{\text{SK}}(\mu, \nu)| = d_{\text{SK}}(\mu, \nu)$. Note that for all $t_\mu, t_\nu \in \mathbb{R}$,

$$\begin{aligned} & |\mu(\theta \leq t_\mu) - \mu(\theta \leq t_\nu) - \nu(\theta \leq t_\mu) + \nu(\theta \leq t_\nu)| \\ & = |\mu(\theta \in (t_\mu \wedge t_\nu, t_\mu \vee t_\nu]) - \nu(\theta \in (t_\mu \wedge t_\nu, t_\mu \vee t_\nu])|, \end{aligned}$$

where $a \wedge b = \min\{a, b\}$ and $a \vee b = \max\{a, b\}$. Hence, we can rewrite d_{SK} as the IPM generated by indicator functions for half-open intervals

$$d_{\text{SK}}(\mu, \nu) = \sup_{l, u} \left| \int 1\{t \in (l, u]\} d\mu(t) - \int 1\{t \in (l, u]\} d\nu(t) \right|.$$

By definition, $d_{\mathcal{I}}(\mu, \nu) \geq d_{\text{SK}}(\mu, \nu)$, so to prove the result it suffices to show that we cannot have a strict inequality.

To this end, note that for any $-\infty \leq a \leq b \leq \infty$ and any measure $\mu \in \Delta(\mathbb{R})$,

$$\lim_{\varepsilon \downarrow 0} \int 1\{t \in (a - \varepsilon, b]\} d\mu(t) = \int 1\{t \in [a, b]\} d\mu(t),$$

while for any $-\infty \leq a < b \leq \infty$,

$$\begin{aligned} \lim_{\varepsilon \downarrow 0} \int 1\{t \in (a, b - \varepsilon]\} d\mu(t) &= \int 1\{t \in (a, b)\} d\mu(t), \\ \lim_{\varepsilon \downarrow 0} \int 1\{t \in (a - \varepsilon, b - \varepsilon]\} d\mu(t) &= \int 1\{t \in [a, b)\} d\mu(t). \end{aligned}$$

Since the same argument applies when integrating over ν , it is immediate that $d_{\mathcal{I}}(\mu, \nu) = d_{\text{SK}}(\mu, \nu)$. The result of Lemma 1 thus follows provided $d_{\mathcal{I}}(\mu, \nu) = d_{\mathcal{F}}(\mu, \nu)$, as we prove in the following lemma.

LEMMA 4: *For $\mathcal{I}$ the set of indicator functions for (open, closed, and half-open) intervals in $\mathbb{R}$, and $\mathcal{F}$ the set of all functions $f: \mathbb{R} \rightarrow [0, 1]$ with $V(f) \leq 1$, we have $d_{\mathcal{I}}(\mu, \nu) = d_{\mathcal{F}}(\mu, \nu)$ for all measures μ, ν .*

The proof of Lemma 4 is broadly similar to arguments in Müller (1997), and is provided in the Supplemental Appendix. *Q.E.D.*

PROOF OF COROLLARY 2: Under the assumptions of the corollary, note that for $\ell \in \mathcal{L}(\mathcal{A})$ and $w \in \mathcal{W}$, $0 \leq \ell(\theta)w(\theta) \leq \lambda\omega$. Note further that for $\theta' \leq \theta$,

$$\begin{aligned} (\ell(\theta)w(\theta) - \ell(\theta')w(\theta'))_+ &= (\ell(\theta)w(\theta) - \ell(\theta')w(\theta) + \ell(\theta')w(\theta) - \ell(\theta')w(\theta'))_+ \\ &\leq \ell(\theta')(w(\theta) - w(\theta'))_+ + w(\theta)(\ell(\theta) - \ell(\theta'))_+. \end{aligned}$$

Hence, for an increasing sequence $t_0, \dots, t_K$,

$$\begin{aligned} &\sum_{k=1}^K (\ell(t_k)w(t_k) - \ell(t_{k-1})w(t_{k-1}))_+ \\ &\leq \sum_{k=1}^K \ell(t_k)(w(t_k) - w(t_{k-1}))_+ + \sum_{k=1}^K w(t_k)(\ell(t_{k-1}) - \ell(t_k))_+ \\ &\leq \omega \sum_{k=1}^K (\ell(t_{k-1}) - \ell(t_k))_+ + \lambda \sum_{k=1}^K (w(t_k) - w(t_{k-1}))_+ \leq \omega\lambda + \lambda(\omega - \omega^{-1}). \end{aligned}$$

Since $\omega \geq 1$, it follows that $0 \leq \mathcal{W} \cdot \mathcal{L}(\mathcal{A}) \leq \lambda(2\omega - \omega^{-1})$ and $V(\mathcal{W} \cdot \mathcal{L}(\mathcal{A})) \leq \lambda(2\omega - \omega^{-1})$, from which the result is immediate noting that, for the first part of the result, we can take $\omega = 1$. *Q.E.D.*

PROOF OF COROLLARY 3: By Corollary 2,

$$d_{\mathcal{W} \cdot \mathcal{L}(\mathcal{A})}(\hat{\pi}^*(\theta|X), \pi^*(\theta|X)) \leq \lambda(2\omega - \omega^{-1})d_{\text{SK}}(\hat{\pi}^*(\theta|X), \pi^*(\theta|X)).$$

By the first part of Proposition 1, we thus know that

$$\bar{R}(\hat{\pi}^*(\theta|X); X, \mathcal{L}, \mathcal{W}, \pi^*) \leq 2\omega\lambda(2\omega - \omega^{-1})d_{\text{SK}}(\hat{\pi}^*(\theta|X), \pi^*(\theta|X)).$$

Procedure 2 reports $\pi^0(\theta|X)$ only when $d_{\text{SK}}(\pi^0(\theta|X), \pi^*(\theta|X)) \leq \tau$, and hence when $\bar{R}(\pi^0(\theta|X); X, \mathcal{L}, \mathcal{W}, \pi^*) \leq 2\omega\lambda(2\omega - \omega^{-1})\tau$, so the first part of the result follows from Corollary 1.

For the second part of the result, note that Procedure 2 reports $\pi^*(\theta|X)$ only when $d_{\text{SK}}(\pi^0(\theta|X), \pi^*(\theta|X)) > \tau$. By our assumption that $\mathcal{L}$ is maximally rich, there exists an agent with the loss used to construct the lower bound in the proof of Proposition 1. Hence, there exists an agent with regret at least

$$d_{\mathcal{L}(\mathcal{A})}(\pi^0(\theta|X), \pi^*(\theta|X)) = \lambda d_{\text{SK}}(\pi^0(\theta|X), \pi^*(\theta|X)) > \lambda\tau,$$

which implies the result. *Q.E.D.*

PROOF OF COROLLARY 4: Procedure 3 reports π^0 only when $d_{\text{SK}}(\pi^0(\theta|X), \pi^S(\theta|X)) \leq \tau$. By the triangle inequality, this implies that $d_{\text{SK}}(\pi^0(\theta|X), \pi^*(\theta|X)) \leq \tau + \delta$, and thus by the argument in the proof of Corollary 3 above that

$$\bar{R}(\pi^0(\theta|X); X, \mathcal{L}, \mathcal{W}, \pi^*) \leq 2\omega\lambda(2\omega - \omega^{-1})(\tau + \delta),$$

so the first part of the result follows from Proposition 1.

Procedure 3 reports π^S only when $d_{\text{SK}}(\pi^0(\theta|X), \pi^S(\theta|X)) > \tau$, in which case the triangle inequality implies that $d_{\text{SK}}(\pi^0(\theta|X), \pi^*(\theta|X)) > \tau - \delta$. By our assumption that $\mathcal{L}$ is maximally rich, there exists an agent with the loss used to construct the lower bound in the proof of Proposition 1. Hence, there exists an agent with regret at least

$$d_{\mathcal{L}(\mathcal{A})}(\pi^0(\theta|X), \pi^*(\theta|X)) = \lambda d_{\text{SK}}(\pi^0(\theta|X), \pi^*(\theta|X)) > \lambda(\tau - \delta),$$

which implies the result. *Q.E.D.*

PROOF OF LEMMA 2: By Theorem 4.6 in Ghosal and van der Vaart (2017), for any $Q \in \Delta(\mathcal{X}_0)$ and $\pi_\alpha^* = \text{DP}(\alpha, Q)$, $\pi_\alpha^*(P|X) = \text{DP}(\alpha + n, \frac{\alpha}{\alpha+n}Q + \frac{n}{\alpha+n}P_n)$. By Theorem 4.16 of Ghosal and van der Vaart (2017), $\pi_\alpha^*(P|X) \rightarrow_d \pi^B(P|X) = \text{DP}(n, P_n)$ as $\alpha \rightarrow 0$, where $\rightarrow_d$ denotes convergence in distribution. By assumption, θ is continuous in P (with respect to the topology of weak convergence) almost everywhere on the support of $\pi^B(P|X)$. Hence, by the continuous mapping theorem (e.g., Theorem 18.11 of van der Vaart (1998)), $\pi_\alpha^*(\theta|X) \rightarrow_d \pi^B(\theta|X)$. By the definition of convergence in distribution, it follows that the distribution function of $\pi_\alpha^*(\theta|X)$ converges to that of $\pi^B(\theta|X)$ at every continuity point of the latter distribution function. If $\pi^B(\theta|X)$ is continuous, then all values of θ are continuity points for the distribution function of $\pi^B(\theta|X)$. If instead the

range of $\theta(P)$ is countable, then for any point t in the range there is a value $\varepsilon(t) > 0$ such that $t + \varepsilon(t)$ is smaller than any point in the range above t , and is trivially a continuity point of $\pi^B(\theta|X)$. Hence, convergence in distribution implies that the distribution function of $\pi_\alpha^*(\theta|X)$ at $t + \varepsilon(t)$ converges to that of $\pi^B(\theta|X)$, and thus that the same holds at t . For both cases, convergence in distribution thus implies point-wise convergence of distribution functions. However, point-wise convergence of distribution functions implies uniform convergence (by, e.g., Lemma 2.11 of [van der Vaart \(1998\)](#)), and uniform convergence of distribution functions implies convergence in d_{SK} . *Q.E.D.*

PROOF OF LEMMA 3: By assumption, $f(\theta(\cdot)) \in \mathcal{G}$ for all $f \in \mathcal{F}$, so by Lemma 1,

$$\begin{aligned} d_{SK}(\pi^B(\theta|X), \pi^*(\theta|X)) &= \sup_{f \in \mathcal{F}} |E_{\pi^B(\theta|X)}[f(\theta)] - E_{\pi^*(\theta|X)}[f(\theta)]| \\ &= \sup_{f \in \mathcal{F}} |E_{\pi^B(\beta|X)}[f(\theta(\beta))] - E_{\pi^*(\beta|X)}[f(\theta(\beta))]| \\ &\leq \sup_{g \in \mathcal{G}} |E_{\pi^B(\beta|X)}[g(\beta)] - E_{\pi^*(\beta|X)}[g(\beta)]|, \end{aligned}$$

from which the result is immediate. *Q.E.D.*

PROOF OF COROLLARY 5: Follows immediately from Lemmas 2 and 3 and Corollary 4. *Q.E.D.*

REFERENCES

- Andrews, Isaiah and Jesse Shapiro (2025), “Bootstrap Census.” Zenodo. <https://doi.org/10.5281/zenodo.17085552>. [0863]
- Andrews, Isaiah and Jesse M. Shapiro (2026), “Supplement to ‘Communicating Scientific Uncertainty via Approximate Posteriors’.” *Econometrica Supplemental Material*, 94, <https://doi.org/10.3982/ECTA22613>. [0846]
- Andrews, Isaiah and Jesse M. Shapiro (2021), “A Model of Scientific Communication.” *Econometrica*, 89 (5), 2117–2142. [0843,0846,0850]
- Angelini, Giovanni, Giuseppe Cavaliere, and Luca Fanelli (2022), “Bootstrap Inference and Diagnostics in State Space Models: With Applications to Dynamic Macro Models.” *Journal of Applied Econometrics*, 37 (1), 3–22. [0845]
- Angelini, Giovanni, Giuseppe Cavaliere, and Luca Fanelli (2024), “An Identification and Testing Strategy for Proxy-SVARs With Weak Proxies.” *Journal of Econometrics*, 238 (2), 105604. [0845]
- Athey, Susan and Guido W. Imbens (2023), “In Search of the Third Number: What to Report Beyond Point Estimates and Standard Errors.” https://youtu.be/i-K9a6UhtsE?si=-rBsjVgY2p3c_DLc, Berkeley Initiative for Transparency in the Social Sciences Keynote Lecture; video accessed in December 2023. [0843,0850]
- [BBJN+] Benjamin, Daniel J., James O. Berger, Magnus Johannesson, Brian A. Nosek, Eric-Jan Wagenmakers, Richard Berk, Kenneth A. Bollen, Björn Brembs, Lawrence Brown, Colin Camerer et al. (2018), “Redefine Statistical Significance.” *Nature Human Behaviour*, 2 (1), 6–10. [0850]
- Beran, Rudolf (1997), “Diagnosing Bootstrap Success.” *Annals of the Institute of Statistical Mathematics*, 49 (1), 1–24. [0845]
- Castillo, Ismaël and Richard Nickl (2014), “On the Bernstein–von Mises Phenomenon for Nonparametric Bayes Procedures.” *Annals of Statistics*, 42 (5), 1941–1969. [0861]
- Cavaliere, Giuseppe, Luca Fanelli, and Iliyan Georgiev (2025), “Bootstrap Diagnostic Tests.” *arXiv econ.EM*. arXiv:2509.01351. Accessed on 03 September 2025. [0845]
- Chamberlain, Gary and Guido W. Imbens (2003), “Nonparametric Applications of Bayesian Inference.” *Journal of Business and Economic Statistics*, 21 (1), 12–18. [0844]
- Cohen, Alma and Liran Einav (2007), “Estimating Risk Preferences From Deductible Choice.” *American Economic Review*, 97 (3), 745–788. [0845]

- Devolder, Olivier, François Glineur, and Yurii Nesterov (2010), “Solving Infinite-Dimensional Optimization Problems by Polynomial Approximation.” In *Recent Advances in Optimization and Its Applications in Engineering: The 14th Belgian-French-German Conference on Optimization*, 31–40, Springer. [0850]
- Efron, Bradley (1982), *The Jackknife, the Bootstrap and Other Resampling Plans*. CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics. [0845]
- Filion, Guillaume J. (2015), “The Signed Kolmogorov-Smirnov Test: Why It Should not Be Used.” *GigaScience*, 4 (1), 9. [0857]
- Gasparini, Mauro (1995), “Exact Multivariate Bayesian Bootstrap Distributions of Moments.” *Annals of Statistics*, 23 (3), 762–768. [0844]
- Geweke, John (1997), “Posterior Simulators in Econometrics.” In *Advances in Economics and Econometrics: Theory and Applications, Seventh World Congress*, Vol. 3 (David M. Kreps and Kenneth F. Wallis, eds.), 128–165, Cambridge University Press, Cambridge. [0848]
- Geyer, Charles J. (2013), “Asymptotics of Maximum Likelihood Without the LLN or CLT or Sample Size Going to Infinity.” In *Advances in Modern Statistical Theory and Applications: A Festschrift in Honor of Morris L. Eaton*. Institute of Mathematical Statistics, Vol. 10, 1–25. [0852]
- Ghosal, Subhashis and Aad van der Vaart (2017), *Fundamentals of Nonparametric Bayesian Inference*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge. [0852, 0860, 0861, 0873]
- [GTKRS+] Gopalan, Parikshit, Adam Tauman Kalai, Omer Reingold, Vatsal Sharan, and Udi Wieder (2022), “Omnipredictors.” In *13th Innovations in Theoretical Computer Science Conference (ITCS 2022)* (Mark Braverman, ed.). Leibniz International Proceedings in Informatics (LIPIcs), Vol. 215, 79:1–79:21, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany. [0845]
- Hahn, Jinyong and Zhipeng Liao (2021), “Bootstrap Standard Error Estimates and Inference.” *Econometrica*, 89 (4), 1963–1977. [0845]
- Hall, Peter (1992), *The Bootstrap and Edgeworth Expansion*. Springer Series in Statistics. Springer, New York, NY. [0845, 0846]
- Hildreth, Clifford (1963), “Bayesian Statisticians and Remote Clients.” *Econometrica*, 31 (3), 422–438. [0848]
- Müller, Alfred (1997), “Integral Probability Metrics and Their Generating Classes of Functions.” *Advances in Applied Probability*, 29 (2), 429–443. [0857, 0872]
- Nickerson, David W., Ryan D. Friedrichs, and David C. King (2006), “Partisan Mobilization Campaigns in the Field: Results From a Statewide Turnout Experiment in Michigan.” *Political Research Quarterly*, 59 (1), 85–97. [0847, 0852]
- Noarov, Georgy, Ramya Ramalingam, Aaron Roth, and Stephan Xie (2025), “High-Dimensional Prediction for Sequential Decision Making.” In *Proceedings of the 42nd International Conference on Machine Learning*, PMLR 267. [0845]
- Raiffa, Howard and Robert Schlaifer (1961), *Applied Statistical Decision Theory*. Harvard University Graduate School of Business Administration, Boston, MA. [0848]
- Rubin, Donald B. (1981), “The Bayesian Bootstrap.” *Annals of Statistics*, 9 (1), 130–134. [0844, 0860]
- Singh, Kesar (1981), “On the Asymptotic Accuracy of Efron’s Bootstrap.” *Annals of Statistics*, 9 (6), 1187–1195. [0846]
- Smets, Frank and Rafael Wouters (2007), “Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach.” *American Economic Review*, 97 (3), 586–606. [0845]
- van der Vaart, Aad W. (1998), *Asymptotic Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge. [0861, 0873, 0874]
- Wang, Andrew (2025), “A General Diagnostic of the Normal Approximation in GMM Models.” *Econometrics Journal*, 28 (2), 261–275. [0845]
- Weng, Chung-Sing (1989), “On a Second-Order Asymptotic Property of the Bayesian Bootstrap Mean.” *Annals of Statistics*, 17 (2), 705–710. [0846, 0862]
- Zhan, Zhaoguo (2018), “Detecting Weak Identification by Bootstrap.” <https://web.archive.org/web/20230806184309/https://sites.google.com/site/zhaoguo/zhao/>. Accessed on 07 August 2023. [0845]

Co-editor Guido W. Imbens handled this manuscript.

Manuscript received 6 January, 2024; final version accepted 21 January, 2026; available online 26 January, 2026.

The replication package for this paper is available at <https://doi.org/10.5281/zenodo.18203246>. The Journal checked the data and codes included in the package for their ability to reproduce the results in the paper and approved online appendices.

All authors assume responsibility for all aspects of the paper.