Political Science Research and Methods (2026), **14**, 414–429
doi:10.1017/psrm.2025.16

ORIGINAL ARTICLE

Bayesian reasoning for qualitative replication analysis: Examples from climate politics

Tasha Fairfield[1] and Andrew Charman[2]

[1]Department of International Development, London School of Economics, London, UK and [2]Department of Physics, UC Berkeley, Berkeley, CA, USA
Corresponding author: Tasha Fairfield; Email: tasha.fairfield@gmail.com

(Received 19 August 2024; revised 31 December 2024; accepted 20 February 2025)

Abstract
This paper demonstrates how Bayesian reasoning can be used for an analog of replication analysis with qualitative research that conducts inference to best explanation. We overview the basic mechanics of Bayesian reasoning with qualitative evidence and apply our approach to recent research on climate change politics, a matter of major importance that is beginning to attract greater interest in the discipline. Our re-analysis of illustrative evidence from a prominent article on global collective-action versus distributive politics theories of climate policy largely accords with the authors' conclusions, while illuminating the value added of Bayesian analysis. In contrast, our in-depth examination of scholarship on oil majors' support for carbon pricing yields a Bayesian inference that diverges from the authors' conclusions. These examples highlight the potential for Bayesian reasoning not only to improve inferences when working with qualitative evidence but also to enhance analytical transparency, facilitate communication of findings, and promote knowledge accumulation.

Keywords: Bayesian inference; climate politics; reliability of inference; qualitative methods

1. Introduction

Qualitative evidence can make vital contributions to scholarship that strives for inference to best explanation. Diverse sources of empirical evidence that cannot easily be quantified, including but hardly limited to detailed information from interviews with key informants, ethnographic observations, news reports, meeting notes, and archival records, provide "clues" that help adjudicate between alternative explanations, in the same way that a detective endeavors to figure out which suspect(s) committed the crime, how, and why. Yet qualitative studies that make causal arguments do not always draw clearly reasoned and well justified conclusions from the evidence presented. Among other common pitfalls, the authors often tell a causal story that seems consistent with the evidence, but without considering how well the evidence might fit with alternative explanations. This practice can lead to overstated claims and misses the opportunity for knowledge accumulation that comes from asking how much better the author's theory, hypothesis, or explanatory framework performs relative to others in the literature.

This paper advocates Bayesian reasoning as a tool for conducting a qualitative analog of replication analysis that involves re-examining the empirical evidence authors provide and re-assessing how strongly it supports the authors' argument relative to rival hypotheses. We apply this approach

© The Author(s), 2025. Published by Cambridge University Press on behalf of EPS Academic Ltd. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/psrm.2025.16 Published online by Cambridge University Press

to recent scholarship on climate politics, an area of crucial substantive importance that has begun to attract more attention from social scientists. We suggest that the growing body of qualitative research in this domain and beyond could usefully harness Bayesian reasoning to arrive at more systematic, more transparent, and more defensible inferences.

By way of introduction, Bayesian reasoning is an intuitive process that begins by assessing relative *prior odds* on rival hypotheses, drawing on any relevant initial information we possess. We gather evidence and evaluate its inferential weight by asking which hypothesis makes that evidence more expected, and how much more expected relative to a rival. We then update to obtain *posterior odds* on our hypotheses—following Bayes' rule, we gain more confidence in whichever hypothesis makes the evidence more expected.

Bayesian reasoning offers multiple advantages. Beyond helping scholars make more rigorous inferences and avoid common cognitive biases, it opens up the author's analysis for scrutiny and facilitates evaluation of the reliability of findings—our preferred terminology for analogs of replication in qualitative research—namely, does the evidence presented justify the conclusions? Bayesianism further provides a natural framework for knowledge accumulation—posterior odds based on what we know so far simply become "prior" odds when we proceed to gather more data. In addition, Bayesianism facilitates communication of findings, by allowing us to more clearly express our judgements about which hypothesis is more plausible and how much more plausible relative to rivals or, equivalently, how much uncertainty surrounds our conclusions.

Moving forward, Section 2 reviews the fundamentals of Bayesian inference (Bennett, 2015; Fairfield and Charman, 2022) and explicates our approach to scrutinizing the reliability of findings. Section 3 provides an illustration drawing on Aklin and Mildenberger's (2020) comparison of global collective action versus distributive politics theories of climate change. Section 4 introduces a simple additive form of Bayes' rule and defines the weight of evidence, an intuitive concept that captures how strongly an empirical observation supports one hypothesis over a rival. Section 5 applies this approach to conduct a more extensive evaluation of reliability that examines how strongly evidence from Vormedal et al.'s (2020) research on oil majors' support for carbon pricing favors the competitive advantage argument they advance over a salient alternative from business politics literature on strategic accommodation. We have chosen Aklin and Mildenberger (2020) and Vormedal *et al.* (2020) for scrutiny to illustrate the contribution that Bayesian reasoning can make toward improving reliability of inference and analytic transparency even in leading qualitative scholarship that takes rival explanations seriously. Our Bayesian reanalysis of some of Aklin and Mildenberger's evidence largely supports their argument, whereas our more comprehensive re-evaluation of evidence in the study by Vormedal *et al.* yields an inference that diverges from their conclusions.

2. Bayesian inference

Bayes' rule provides a rigorous framework for updating our judgements about which hypothesis provides the best explanation when we learn new information:

$$\begin{aligned} \textit{posterior odds} &= \textit{prior odds} \times \textit{likelihood ratio} \\ \frac{P(H_i|E\,\mathcal{J})}{P(H_j|E\,\mathcal{J})} &= \frac{P(H_i|\mathcal{J})}{P(H_j|\mathcal{J})} \times \frac{P(E|H_i\,\mathcal{J})}{P(E|H_j\,\mathcal{J})}, \end{aligned} \tag{1}$$

where H_i represents a hypothesis, E represents an evidentiary observation, and $\mathcal{J}$ represents relevant initial knowledge or background information. The *prior odds* represent the scholar's degree of belief in H_i relative to H_j in light of their background information $\mathcal{J}$. The *posterior odds* in turn express how much more plausible the scholar finds H_i relative to H_j in light of not just their background information but also evidence E; that is, the posterior odds represent the updated assessment of the relative plausibility of the hypotheses after learning E. The *likelihood ratio* gives the relative plausibility, or likelihood, of the evidence under one hypothesis compared to the rival. Here we evaluate how much

https://doi.org/10.1017/psrm.2025.16 Published online by Cambridge University Press

more (or less) expected the evidence is assuming that H_i is true, as compared to assuming instead that H_j is true. Assessing likelihood ratios is the central inferential step that tells us how to update from prior odds to posterior odds. The likelihood ratio in essence determines how strongly the evidence favors H_i vs. H_j—the more expected H_i makes E relative to H_j, the more strongly E favors H_i over H_j and, regardless of our prior odds, the more confidence we subsequently gain in H_i relative to H_j.

In qualitative research, where we heuristically follow Bayesian principles to guide our reasoning (rather than conducting quantitative Bayesian model comparison[1]), we evaluate likelihood ratios by "mentally inhabit[ing] the world" of each hypothesis (Hunter, 1984). We imagine the kinds of events that might naturally occur in a world governed by hypothesis H_i. We ask if evidence E seems like the kind of thing we would expect in that world (high probability), if it would have to be regarded as an unexpected coincidence or a surprising fluke (low probability), or if its likelihood falls somewhere in between on the probability spectrum. We then imagine an alternative hypothetical world governed by hypothesis H_j and reason about whether E seems more or less expected than in the H_i world. All that matters for Bayesian inference is the *relative* likelihood of the evidence. Evidence that is unlikely under H_i will nevertheless favor that hypothesis over H_j if it is even *less* likely under H_j. Similarly, evidence that is expected under H_i will actually undermine that hypothesis relative to H_j if it is even *more* expected under H_j. In sum, we evaluate how strongly evidence supports a hypothesis over a rival by thinking about the predictions each hypothesis makes and assessing to what extent the evidence seems more expected in the world of one hypothesis relative to that of the rival.

This Bayesian approach differs from other qualitative methodologies in the importance placed on rival hypotheses and the centrality of uncertainty. Working within a frequentist paradigm, King *et al.* (KKV, 1994) focus on assessing whether evidence is consistent with the observable implications of a single hypothesis of interest, while approaches to process tracing that emphasize causal mechanisms often view "tracing the process" as adequate for inference, without considering rival hypotheses. Crisp-set QCA introduces ad hoc procedures to relax the rules of deductive logic, rather than applying probabilistic reasoning, and fuzzy set theory enshrines vagueness (degree of membership in a vague category), rather than uncertainty, as its core framework for reasoning. None of these approaches address the central critique of qualitative scholarship that KKV (1994:32) raised but ultimately left unresolved: "the pervasive failure to provide reasonable estimates of the uncertainty of the investigator's inferences." Bayesianism does just that—posterior odds express how much confidence we have in one hypothesis relative to rivals considering all information analyzed so far: the greater the odds, the lower the uncertainty.

We now turn to the framework that Bayesianism can provide for analogs of replication analysis. As Freese and Peterson (2020) observe, replication can mean very different things to different scholars. For frequentist statistical analysis, their typology helpfully distinguishes between approaches that use (i) the same vs. new data and (ii) the same vs. different procedures with respect to the study of interest. We will use the term *reliability* of inference to encompass all these activities, with the idea that each ultimately aims to contribute to the overarching goal of assessing "whether scientific claims based on scientific results are true" (Goodman, Fanelli, and Ioannidis 2016:1), or in more Bayesian terms, how much confidence in our conclusions does the data merit.

In frequentist statistics, we would consider standard replication analysis to fall under Freese and Peterson's (2020) new-data/same-procedures category, where the same analysis is conducted with a new random sample of data from the original population. But qualitative evidence cannot be analyzed within a frequentist framework, which by its own strictures applies only to stochastic data. Bayesianism applies the same inferential principle to all kinds of data, whether quantitative or qualitative, stochastic or nonstochastic—we evaluate likelihood ratios under competing hypotheses. In a Bayesian framework, examining new data with the same procedures would simply involve analyzing more evidence beyond the study in question and updating the posterior odds accordingly, which

[1] See Jeffreys (1939), Jaynes (2003), and Gregory (2005) on logical Bayesian data analysis.

https://doi.org/10.1017/psrm.2025.16 Published online by Cambridge University Press

while valuable does not involve any direct scrutiny of the study itself and hence does not comport well with the frequentist notion of new-data/same-procedures replication to evaluate a study's "specific results" (Freese and Peterson, 2020).

Our focus will instead be "same-data/Bayesian-procedures" re-evaluation of qualitative scholarship that may or may not have employed Bayesian reasoning. This task involves scrutinizing how strongly evidence from the study weighs in favor of the authors' argument relative to rival explanations. It is especially important for qualitative research, because analyzing likelihood ratios when working with plain-language hypotheses and detailed empirical observations that cannot naturally be quantified will always involve some degree of subjectivity. Furthermore, different scholars have different background information that can matter for how they interpret the likelihood of evidence under each hypothesis. Ideally, a "same-data/Bayesian-procedures" exercise of the sort we present in Sections 3 and 5 would be the first step in a larger process of discussion and debate among scholars, with the goal of establishing a reasonable level of agreement regarding the inferential import of the evidence.

Readers may notice that in this context, we have not placed much importance on prior odds—this is the component of Bayesian analysis that is most sensitive to background information, so priors can vary widely among scholars.[2] We recommend allowing scholars to supply their own priors and focusing instead on the inferential import of evidence in the study of interest when conducting Bayesian re-evaluation. If we wish to scrutinize the broader body of existing knowledge from relevant literature that should inform our priors, we enter the realm of meta-analysis, which we return to in Section 6.[3] That said, our examples in Sections 3 and 5 include high-level assessments of prior odds, primarily to illustrate the broader Bayesian reasoning process.

3. Application: Climate-change politics

Despite the importance of rival hypotheses, their role has been under-appreciated in qualitative research that aims to provide causal explanations. In the environmental politics domain for example, of 56 relevant articles that we identified (since 2015) in two leading journals (*Environmental Politics* and *Global Environmental Politics*), only 10 included some effort to evaluate rival explanations in light of concrete evidence. Three of those ruled out the alternatives almost immediately, and none evaluated which hypothesized explanation made the evidence more expected. A Bayesian lens reveals that ignoring rival hypotheses is dangerous, because even if the evidence seems to fit with the causal process or mechanism that a hypothesis proposes, that evidence could fit even better with a rival. Using an accepted theory to explain cases is a valid endeavor. However, few social science theories are widely accepted, and there will be readers who legitimately wonder whether the author's chosen theory or framework performs better than any number of alternatives in the literature or possible explanations one might conjure.

In contrast to the relative neglect of rival hypotheses in qualitative work, Aklin and Mildenberger's (2020) prominent discussion of climate change explicitly compares competing theories emphasizing global collective action vs. distributive conflict. Their approach to adjudicating between these alternatives has some Bayesian flavor, in that they compare each theory's empirical predictions with available evidence. By comparison, scholars often discuss "observable implications" for only a single theory of interest (e.g., Kelsey, 2021:72–73; Sun and Yu, 2023).

But instead of evaluating *relative likelihoods* for each evidentiary observation under the respective theories, Aklin and Mildenberger (2020, Table 1) classify high-level summary statements of evidence as "consistent", not consistent, or "unlikely" to be consistent with each theory. These three categories are inadequate for assessing evidentiary import, which is always a matter of degree. Evidence deemed

[2] Appendix A.2 expounds on priors.

[3] Bayesian meta-analysis would begin with uninformative priors reflecting a state of minimal background information and then incorporate existing knowledge by evaluating likelihood ratios for empirical observations described in the literature.

https://doi.org/10.1017/psrm.2025.16 Published online by Cambridge University Press

"inconsistent" with both H_1 and H_2 might still support H_1 over H_2, if the evidence is even less likely under H_2 as compared to H_1. Likewise, evidence that seems "consistent" with each hypothesis may nevertheless favor H_1 over H_2, if it is even more likely under H_1 relative to H_2. That is, evidence judged "consistent" (or "inconsistent") with both hypotheses need not be "observationally equivalent" and may well contribute to inference. Despite Aklin and Mildenberger's (2022:146) assertion that they adopt a "Bayesian perspective," their analysis accordingly falls short of Bayesian inference, even if in a loose sense they use "observable implications" to "inform our posterior belief" about the plausibility of alternative theories. We demonstrate how to apply a fully Bayesian approach below.

We will work with the following rival hypotheses:

> *Global Collective Action* (H_{GCA}): *Progress on climate mitigation depends first and foremost on global institutions that discourage free-riding, with mechanisms to monitor and punish non-cooperation. States have a strong tendency to enact climate policies reciprocally, while defection from international agreements by pivotal countries motivates other states to defect as well, for fear of bearing disproportionate costs while the largest carbon polluting nations shirk and continue propelling climate change.*

> *Domestic Organized Interests* (H_{DOI}): *Progress on climate mitigation depends first and foremost on domestic distributive politics. Specifically, the balance of power among organized interests with divergent preferences (e.g., political parties, business associations, social movements) shapes climate policy outcomes. Free-riding at the global level is not a strong constraint; defection from climate commitments by pivotal countries is not automatically reciprocated.*

The wording we employ in H_{GCA} aims to convey that the hypothesis make probabilistic predictions, rather than deterministic assertions about political behavior, which in our view are rarely appropriate for social science. Rather than invoking a broad family of distributive conflict theories (as do Aklin and Mildenberger), we take one specific member of this family as our rival hypothesis, drawing on Hacker and Pierson (2010). We do so for the sake of clarity, since the casual logic that a hypothesis proposes needs to be articulated in enough detail for scholars to envision the corresponding world and reason about the likelihood of evidence in hand.

For simplicity, we assume minimal background knowledge about climate politics, which justifies even prior odds on the hypotheses. We next scrutinize three pieces of evidence that include adequate empirical detail for likelihood reasoning.

> E_1 = *Many states have unilaterally enacted climate mitigation policies in the absence of a binding climate treaty. For example, European countries enacted carbon taxes before the Kyoto protocol was adopted in 1997. And states around the world continued enacting climate policies after the U.S. withdrew from Kyoto in 2001.* (Aklin and Mildenberger, 2020:12–13)

Intuitively, E_1 favors the domestic politics hypothesis. Applying Bayesian reasoning, first imagine a world where the global collective-action hypothesis is true. Here E_1 would be surprising. This evidence runs counter to H_{GCA}'s predictions—we would instead expect policy adoption to follow establishment of a binding treaty, and considering that the USA is a pivotal country for curtailing climate change, its decision to withdraw from Kyoto should have motivated other states to halt their climate efforts, rather than enacting new policies—although we would need to know how many states have done so before arriving at a more precise judgement about how surprising E_1 is under H_{GCA}.

Now consider the alternative world of H_{DOI}. Absent details about how key aspects of domestic politics vary around the world (which for simplicity we assume is not part of our background information), the hypothesis does not make clear predictions about where or how extensively we should

https://doi.org/10.1017/psrm.2025.16 Published online by Cambridge University Press

observe climate policy adoption. Nevertheless, E_1 is much more plausible under H_{DOI}, where international free-riding is not a salient constraint on climate action, than it is under H_{GCA}, where E_1 displays a surprising lack of concern over US free-riding.

Accordingly, E_1 fairly strongly favors the domestic organized interests hypothesis. Because E_1 is much more probable under H_{DOI} relative to H_{GCA}—that is, $P(E_1|H_{\text{DOI}}\,\mathcal{J}) \gg P(E_1|H_{\text{GCA}}\,\mathcal{J})$—it significantly boosts our confidence in H_{DOI}.

> E_2 = *In Australia, John Howard's Liberal government (1996–2007) resisted Kyoto ratification by invoking US nonparticipation.* (Aklin and Mildenberger, 2020:18)

Imagining a world where H_{GCA} is true, the Liberal government's behavior is expected—they invoke international free-riding to explain their position against Kyoto, reflecting the very concerns that the hypothesis posits. Turning to H_{DOI}, the Howard government's opposition to Kyoto is also plausible if we invoke background knowledge that this was a pro-business, ideologically center-right government. Domestic politics is what matters for climate policy in the H_{DOI} world, yet denouncing US free-riding could be a good way for the government to legitimate its position among the broader electorate, beyond business constituencies that oppose carbon regulations. But E_2 is not as expected as it is under H_{GCA}, because there are other more or less equally plausible rationales that the government might invoke when opposing Kyoto in the H_{DOI} world (e.g., compliance would make energy too expensive). That is, H_{DOI} spreads the probability more evenly over other possible justifications for disregarding Kyoto, whereas H_{GCA} concentrates more probability on the US free-riding rationale. Accordingly, E_2 favors H_{GCA}, but in our view only weakly to moderately.

> E_3 = *A Labor government won power in Australia in 2007 with a promise to ratify the Kyoto Protocol and followed through on that promise at COP 13 in Bali.* (Aklin and Mildenberger, 2020:18)

Under H_{GCA}, Labor's behavior is quite surprising: this new government ignores free-rider concerns and signs the treaty despite non-participation from the USA, a pivotal nation for climate change. E_3 is much more expected under H_{DOI}, where a center-left party displacing a center-right party in government is exactly the kind of change in balance of power among organized domestic actors that should produce shifts in climate policy. We therefore judge E_3 to strongly favor H_{DOI} over H_{GCA}.

Our Bayesian reasoning about E_2 and E_3 clarifies precisely where the inferential weight from these observations about Australian politics arises. It is the behavior of the Labor government that lends significant weight to the distributive politics perspective, while the Howard government's actions do much less to discriminate between the hypotheses.

We now evaluate posterior odds. Starting from even prior odds, E_1 strongly boosts our confidence in H_{DOI}, E_2 then lowers our confidence in that hypothesis a bit relative to H_{GCA}, while E_3 contributes another substantial boost in favor of H_{DOI}. In light of these three evidentiary observations, we end up with very strong posterior odds favoring H_{DOI} vs. H_{GCA}. By contrast, readers with substantial background knowledge drawn from the global collective-action literature on climate change might start with prior odds moderately to strongly favoring H_{GCA} and would hence arrive at posterior odds that perhaps only moderately favor H_{DOI} (assuming they agree with our analysis of likelihood ratios).

The posterior odds in light of $E_1E_2E_3$ would in turn become the "prior odds" when moving forward to consider additional evidence, which may of course change our views about which hypothesis is more plausible. While we have only scrutinized a few empirical observations from the article, we concur that this evidence on balance supports Aklin and Mildenberger's argument in favor of distributive politics (even though E_2 taken alone favors the rival hypothesis). A more thorough re-assessment would of course consider the remaining evidence.

https://doi.org/10.1017/psrm.2025.16 Published online by Cambridge University Press

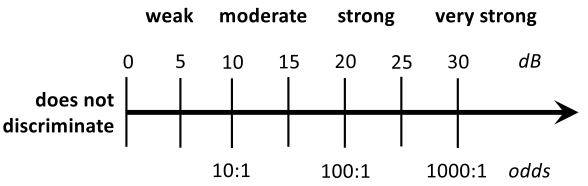


Figure 1. Decibel scale for quantifying the weight of evidence. For calibration, sound files are available at https://tashafairfield.wixsite.com/home/bayes-resources.

4. Bayes' rule in log-odds form

Taking the logarithm of Bayes' rule (1), we obtain a simple, additive relationship that is easy to remember and easy to use:

$$\begin{aligned} \textit{posterior log-odds} \;&=\; \textit{prior log-odds} \;+ \textit{weight of evidence} \\ \log\left[\frac{P(H_j|E\,\mathcal{J})}{P(H_k|E\,\mathcal{J})}\right] &= \log\left[\frac{P(H_j|\mathcal{J})}{P(H_k|\mathcal{J})}\right] + \log\left[\frac{P(E|H_j\,\mathcal{J})}{P(E|H_k\,\mathcal{J})}\right]. \end{aligned} \tag{2}$$

The weight of evidence (the logarithm of the likelihood ratio, Good, 1983) conveys its probative value—namely, how much it supports one hypothesis over a rival. We denote the weight of evidence in favor of H_j relative to H_k as $WoE\,(H_j : H_k)$. When the evidence contains multiple empirical observations, $E = (E_1 E_2 \cdots E_N)$, the overall (net) weight of evidence becomes the sum of the components:

$$WoE\,(H_j : H_k) = WoE_1(H_j{:}H_k) + WoE_2(H_j{:}H_k) + \cdots + WoE_N(H_j{:}H_k)\,. \tag{3}$$

Appendix B.1 addresses a technical nuance (possible dependence among evidentiary observations) that can matter when working with multiple pieces of evidence, but readers should be able to follow the example in Section 5 without delving into those details.

We can fully leverage Eqs. 2 and 3 by quantifying in decibels, a familiar logarithmic scale, and using an analogy to sound (Fairfield and Charman, 2022). For weights of evidence, we ask how loudly E_x speaks for H_j over H_k, e.g., do the facts murmur or roar in favor of one hypothesis over a rival? In acoustics, 3 dB is the minimal noticeable difference an adult with good hearing can detect. Amplification by 10 dB sounds twice as loud; 20 dB sounds four times louder; and 30 dB roughly eight times louder. Figure 1 includes suggested adjectives corresponding to different decibel ranges (i.e., weak, moderate, strong, etc.). Listening to sound files amplified by different decibel levels can help foster standardization and common calibration among scholars when quantifying.

Using Eq. (2) and working with decibels, we can conduct inference by imagining a balance (Fig. 2). Our prior log-odds sit on the side representing the hypothesis we initially favor. The weight of each piece of evidence goes on the side of the balance corresponding to the hypothesis it supports. The tilt of the scale then represents our posterior log-odds, which express how much more plausible we find one hypothesis compared to the rival given all evidence we have analyzed along with our background information. This approach allows scholars to conduct Bayesian inference without any active knowledge about logarithms. The only mathematical skills needed are addition and subtraction.

For qualitative scholars who may be wary of quantification, we emphasize that our approach does not quantify the evidence itself. The rich detail that often characterizes qualitative research should be included in the evidentiary statements to be weighed under rival hypotheses—details can matter greatly for inferential import (Appendix B.2). The log-odds approach instead quantifies our probabilistic assessments of relative likelihoods. Moreover, decibels should never stand alone; they should be accompanied by qualitative reasoning about relative likelihoods.

It is important to acknowledge that in qualitative social science, quantifying relative probabilities always involves some subjectivity and arbitrariness. The important nuances and almost infinite

https://doi.org/10.1017/psrm.2025.16 Published online by Cambridge University Press

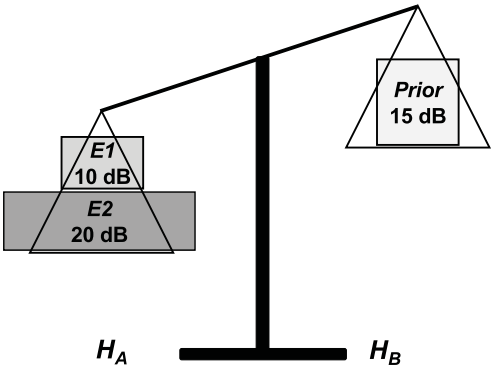


Figure 2. Bayesian balance for log-odds updating. Here the posterior log-odds favor H_A by 15 dB.

possibilities of qualitative, natural language evidence cannot be captured with some off-the-shelf likelihood function mapping every possible observation to a numerical probability, as typically assumed in traditional quantitative analysis. Some of the arbitrariness will average out when aggregating multiple weights of evidence—some will be overweighted and others underweighted. But in situations where we find it particularly difficult to quantify our reasoning, we can report a decibel range rather than a point estimate (e.g., 12–17 dB). We can also conduct sensitivity analysis and report how much the aggregate inference would change if we allocate more or fewer decibels to the pieces of evidence that seem especially challenging to analyze.[4]

Despite some inherent subjectivity and arbitrariness, quantification proffers two key advantages: it allows us to (i) formulate and communicate our judgements more precisely than can be done with qualitative adjectives and (ii) more systematically and transparently aggregate the net weight of evidence. Regarding (i), decibels allow us to identify disagreements that might otherwise be hidden behind vague language. Two scholars might both describe evidence E as "moderately" favoring H_1 over H_2, yet one of their judgements might correspond to 8 dB, whereas the other's might be 18 dB—a 10 dB difference that could be substantively important. They can then debate which line of qualitative reasoning and what quantitative weighting of the evidence is most compelling. Regarding (ii), if the available evidence is not decisive, with some observations favoring H_1 but others favoring H_2, then any effort to draw an inference at least implicitly involves quantification—e.g., we may have to make judgements about whether three observations that weakly favor H_1 counterbalance or even outweigh an observation that moderately favors H_2. Reporting decibels after explaining our reasoning about the inferential import for each piece of evidence promotes analytical transparency by forcing us to clarify our judgements and allowing readers to more easily understand and scrutinize our conclusions. While subjectivity is inevitable in qualitative research, Bayesianism brings the subjectivity into the open while promoting more rational and systematic inference.

5. Application: Oil majors and carbon pricing

We now apply the log-odds formulation to scrutinize reliability of inference for research by Vormedal *et al.* (2020) on "Big Oil and Climate Regulation." The authors adjudicate between two hypotheses regarding when and why oil majors support moderate climate mitigation policies, focusing on the time period beginning with the run-up to the Paris agreement and proceeding through Trump's first presidency:

[4]See Appendix B.3 for additional remarks on quantifying weights of evidence.

https://doi.org/10.1017/psrm.2025.16 Published online by Cambridge University Press

Strategic Accommodation (H_{SA}): *Since 2015, oil majors have endorsed moderate climate policies first and foremost when they face regulatory threats, in an effort to prevent or hedge against more radical and costly policies on the political horizon. In the absence of such regulatory threats, they are unlikely to support climate policies* (e.g., Hacker & Pierson, 2002; Meckling, 2015).

Competitive Advantage[5] (H_{CA}): *Since 2015, oil majors have endorsed moderate climate policies first and foremost when they perceive that those policies will create competitive advantages over rivals and foster new market opportunities that they are well-positioned to exploit. In the absence of such perceived advantages, they are unlikely to support climate policies* (e.g., Falkner, 2008; Vogel, 2018).

The distinction is that the climate policies oil majors support under H_{SA} impose net economic costs on them—albeit lower costs than salient policy alternatives—whereas the policies they support under H_{CA} confer net economic benefits given the current or anticipated economic and policy context. A regulation that imposes direct costs could nevertheless be economically beneficial, if it imposes even greater costs on competitors, leading to market-share growth that offsets the direct costs (e.g., Meckling, 2015). While additional hypotheses could be considered (Appendix A.3), our Bayesian re-analysis focuses on these two that the authors compare.

We begin with prior log-odds favoring H_{SA}, considering that research on oil majors and climate politics during the pre-Paris era "largely explained support or advocacy for carbon regulation"—in particular, market-based emissions trading systems—"as an opportunistic effort to offset the risk of more costly reforms" (Vormedal *et al.*, 2020:145–6). But given that Vormedal *et al.* (2020:146-7) also cite a robust literature on business and environmental regulation that falls on the competitive advantages side of the overarching theoretical debate, we take the priors to only weakly favor H_{SA}. *Prior log-odds* ≈ 6 dB for H_{SA} vs. H_{CA}.

We proceed to analyze all evidence from the article that either we or the authors deem to contribute inferential weight in favor of either hypothesis. A thorough replication exercise should examine all salient evidence from the published study—if we have overlooked empirical information with nontrivial inferential import, scholars should call the omission to attention. We can then easily incorporate it into our analysis as additional evidence that will update our log-odds.[6]

Other than the general characterization of previous literature that informed our priors, our analysis will assume minimal background knowledge. We begin with some initial information pertaining to the independent variables that the hypotheses invoke:

Regulatory pressures[7] ($E_{0a}E_{0b}E_{0c}$)

E_{0a} = *Oil majors faced regulatory threats in the US during the Obama administration, including new climate targets and a 2007 Supreme Court ruling granting the Environmental Protection Agency (EPA) authority to regulate greenhouse gas emissions under the Clean Air Act. Regulatory pressure subsided with the election of Trump and his administration's efforts to roll back climate regulations, which included a campaign pledge to withdraw the US from the Paris Agreement.*[8]

[5] Vormedal *et al.* 2020 call this the "economic interests" hypothesis, but we deem "competitive advantage" a better descriptor.

[6] If scholars think the study in question may have omitted accessible evidence that would have led to a different finding (whether due to cherry picking or inadequate data collection), then additional research to locate and analyze more evidence is merited, but that endeavor would take us from "same-data/Bayesian-procedures" to "new-data/Bayesian-procedures" scrutiny, which we would simply characterize as "continued research" (Fairfield and Charman, 2020). If evidence in the study is poorly reported or documented, or if there are concrete, well-motivated reasons to suspect that evidence has been manipulated or invented, then we would not bother with further scrutiny; we would deem the study to lack credibility.

[7] Vormedal *et al.* (2020:150–151)

[8] https://www.nbcnews.com/politics/2016-election/donald-trump-pledges-rip-paris-climate-agreement-energy-speech-n581236

https://doi.org/10.1017/psrm.2025.16 Published online by Cambridge University Press

E_{0b} = *From 2015 on new threats emerged from multiple law suits in US states aiming to hold the oil majors (Exxon in particular but also Chevron, BP, Shell, and ConocoPhillips) accountable for misleading the public and shareholders about known risks of global warming.*

E_{0c} = *In Europe, "regulatory pressures have continued to build" due to "progressively bolder climate and energy goals."* (Vormedal *et al.*, 2020:150)

Production structure and market expectations ($E_{0d}E_{0e}E_{0f}E_{0g}$)

E_{0d} = *By 2016, technological advances (fracking and liquification) helped make natural gas cheaper than coal for generating electricity in the US and led to a growing market share for gas. In continental Europe, coal remained more competitive with gas. In emerging economies (e.g., China, India), coal remained cheaper than gas.* (Vormedal *et al.*, 2020:157,158)

E_{0e} = *From 2007–2017, most of the oil majors' production profiles shifted substantially from oil toward gas. But for ConocoPhillips and BP, both gas and oil production decreased during this period.* (Vormedal *et al.*, 2020:156)

E_{0f} = *In 2014 BP noted potential for gas production to double its operating cash flow over the next 10 years and emphasized gas as a growth area.* (Vormedal *et al.*, 2020:156-7)

E_{0g} = *"Europe and the United States have the greatest opportunities for coal-to-gas switching thanks to existing infrastructure that allows... immediate displacement of around half of the coal-fired power output."* (Vormedal *et al.*, 2020:158)

This initial information $E_0 = E_{0a} \cdots E_{0g}$ is no more or less plausible under either hypothesis. If we assume no background knowledge about the oil majors' policy positions during the relevant time-period, then neither hypothesis yields expectations about intensity of regulatory pressures or production structure/market expectations. But E_0 will matter for analyzing subsequent evidence regarding the oil majors' behavior.[9]

E_1 = *Before the 2015 Paris meeting, a group of European majors—BP, Equinor, ENI, Shell, and Total—wrote to the UN Framework Convention on Climate Change: "Our companies are already taking a number of actions to help limit emissions, such as growing the share of gas in our production... We need governments across the world to provide us with clear, stable, long-term, ambitious policy frameworks. ...a price on carbon should be a key element of these frameworks."* (quoted in Vormedal *et al.*, 2020:148)

Under either hypothesis, we interpret the requested policy as a moderate carbon price (following the authors).[10] And E_1 is in line with expectations under each hypothesis, given what we know from E_0.

In the H_{SA} world, we reason that the companies were adapting to the reality of continued regulatory pressure in Europe (E_{0c}) by increasing gas production and that they viewed carbon pricing as preferable to other more economically onerous measures that the EU might propose to meet ambitious emissions targets. If we invoke background information that after failure at Copenhagen, actors anticipated that a treaty negotiated in Paris would be weak and nonbinding,[11] there would be little need to hedge against the possibility of more radical alternatives entering the agreement. Nevertheless, securing a strong endorsement of carbon pricing at Paris might help strengthen the oil

[9] In effect, E_0 becomes part of our background knowledge moving forward (Appendix B.1).
[10] They do not define precisely what "moderate" means.
[11] Robert Falkner, personal communication.

https://doi.org/10.1017/psrm.2025.16 Published online by Cambridge University Press

majors' position at the EU level by lending further legitimacy to their preferred market-based solution as opposed to more radical policies that the EU might contemplate.

In the H_{CA} world, support for carbon pricing is also the kind of thing we would expect given the companies' shifts toward gas (E_{0e}) or, for BP, express interest in gas as a strategic priority (E_{0f}). Following Vormedal *et al.*, we reason that carbon pricing gives natural gas a competitive edge over coal, which is much more carbon-intensive, and that oil majors anticipated gains from growing market opportunities for gas to outweigh direct costs incurred from a carbon tax, especially in markets beyond the USA and EU, where coal remained competitive with natural gas (E_{0d}). Even if Paris was expected to produce a weak agreement, we would reason that the EU-based oil majors viewed any progress in global carbon pricing as a step in the right direction.

In short, E_1 does not discriminate much between the two motives—hedging against less desirable policies vs. pursuing competitive advantages. The companies' explicit mention of growing gas shares when requesting carbon pricing would be perhaps slightly more expected under H_{CA}, yielding a weak inferential weight. $WoE_1 \approx 4$ for H_{CA} vs. H_{SA}.

While our reasoning about E_1 under H_{CA} accords with the authors' line of argumentation, their analysis with regard to H_{SA} diverges. On regulatory pressure in Europe, Vormedal *et al.* (2020:151) write: "it is not apparent how the majors' support for carbon pricing could have helped them offset or hedge against these developments" (increasingly ambitious emissions reduction goals), whereas we reason under H_{SA} that the companies' actions could plausibly help entrench carbon pricing as the most legitimate approach within the EU. However, the less one thinks a strong endorsement of carbon pricing in the Paris agreement would strengthen the oil majors' lobbying position against more onerous alternatives at the EU level, the more E_1 favors H_{CA}.

> E_2 = *In 2017, while President Trump was moving toward withdrawing from the Paris Agreement, US oil majors Exxon and Chevron publicly urged him to reconsider. In a widely reported letter to Trump,*[12] *Exxon's CEO argued that the US was "well positioned to compete within the framework of the Paris Agreement with abundant low-carbon resources such as natural gas" and that remaining in the accord would allow the US to retain "a seat at the negotiating table to ensure a level playing field so that all energy sources and technologies are treated equitably in an open, transparent and competitive global market."* (Vormedal *et al.*, 2020:149)[13]

Under H_{SA}, Chevron and Exxon's support for the Paris Agreement might initially seem surprising, given that Trump's election signaled reduced regulatory threats from the US government (E_{0a}). But we can make sense of E_2 in the H_{SA} world by reasoning that the extent to which regulatory threats would dissipate under Trump was not yet clear and/or that the US-based oil majors took a longer-term perspective beyond the time horizon of the Trump administration (a possibility Vormedal *et al.* also consider, albeit not in the context of Bayesian likelihood reasoning) and sought to preclude more radical climate policies that might emerge on the national agenda if Democrats regained power. Moreover, Exxon's comments regarding "a seat at the negotiating table" and "competitive global markets" seem plausible in a world where the oil companies are hedging against less market-friendly policies that might emerge in the international context absent US leadership. If we invoke background information that Paris produced a weak agreement as anticipated, and five-to-one US voters in 2017 favored participation,[14] then E_2 becomes all the more plausible in

[12] https://www.ft.com/content/fcf73abc-4202-11e7-9d56-25f963e998b2 https://www.nytimes.com/2017/05/30/us/politics/paris-climate-agreement-trump.html https://www.thestreet.com/investing/stocks/exxon-ceo-woods-urges-trump-to-remain-in-paris-agreement-14152362

[13] The authors omit "to ensure a level playing field...."

[14] https://climatecommunication.yale.edu/publications/5-1-voters-say-u-s-participate-paris-climate-agreement/

https://doi.org/10.1017/psrm.2025.16 Published online by Cambridge University Press

the hedging world—oil majors would be expected to support a weak but popular treaty to preclude pressure from below for more ambitious climate policies that might spur future government action.

E_2 seems similarly plausible under H_{CA} given the companies' shifts toward gas production (E_{0e}), following the logic discussed when reasoning about E_1. Exxon's remarks on being "well positioned to compete" thanks to natural gas fit well with expectations in this world, while we would interpret comments regarding "a seat at the negotiating table" as motivated by an interest in further promoting carbon pricing at the international level. We find the "well positioned to compete..." remark more plausible under H_{CA}, but the "seat at the negotiating table..." comment seems somewhat more plausible under H_{SA} (following our reasoning above). On balance, we judge the weight of evidence to weakly favor H_{CA}. $WoE_2 \approx 3$ dB for H_{CA} vs. H_{SA}.

It should be noted that beyond what we include in E_2, Vormedal *et al.* (2020:149) imply that Exxon's letter also mentioned that "Paris could provide a boost for gas." That detail would matter for our analysis and would shift the weight of evidence to more strongly favor H_{CA} (Appendix B.2). However, the cited sources do not contain that detail, nor does an earlier Exxon letter to the administration.[15]

The weight of E_2 also depends to some extent on how exactly one interprets H_{SA}—does support from oil majors require an explicit policy alternative currently under debate, or is a longer-term possibility of more costly alternatives gaining political traction adequate (as we have assumed above)?[16] Building consensus about the weight of evidence may require clarifying or refining theory. Structuring debate within a Bayesian framework can help scholars identify aspects of theory that merit attention to that end.

> E_3 = *In 2017, US-based oil majors Exxon and ConocoPhillips, joined by EU-based oil majors Shell and BP, created the Climate Leadership Council (CLC), "a new, Republican-backed coalition promoting a US carbon tax starting at US$ 40 per ton. ...the carbon tax would displace all federal regulations of stationary carbon sources. The initial plan included a climate liability waiver that would have shielded companies from lawsuits of alleged liability for damages caused by historical emissions." Exxon contributed US$ 1 million to the CLC's lobbying campaign; ConocoPhillips contributed US$ 2 million.* (Vormedal *et al.*, 2020:149)

This evidence seems highly plausible under H_{SA}. Proposing a moderate carbon tax to secure a climate liability waiver is very much in line with expectations, given the threat of legal action that these oil majors faced (E_{0b}). Aiming to replace extant carbon regulations is also consistent with the kind of behavior we expect in the H_{SA} world, where supporting moderate climate policy is a strategy for precluding more onerous regulations—seeking to repeal undesirable rules would be a natural extension of that logic, even if not stated explicitly in the hypothesis. Although US regulatory threats subsided with Trump's election (E_{0a}), the timing of this initiative seems plausible, because an effort to eliminate federal regulations and preclude liability (at the price of a modest carbon tax) would face better prospects under Trump's Republican-trifecta government.

Evidence E_3 might seem fairly plausible under H_{CA}, given Exxon's and Shell's shifts toward gas (E_{0e}) and BP's 2014 statements about the importance of gas (E_{0f}) (despite declining gas production from 2007–2017), alongside potential for coal-to-gas switching in the USA (E_{0g}). We would again reason that these companies thought a moderate carbon tax would create competitive advantages for their natural gas relative to coal, bringing benefits that would outstrip the direct costs. On the other hand, this story might seem a bit strained, considering that gas was already gaining ground relative to coal in US markets (E_{0d})—it is not clear that the incentives would be strong enough for

[15] https://biotech.law.lsu.edu/blog/Exxon-Letter-Supporting-Paris-Agreement-March-2017.pdf

[16] Another complication emerges: different companies might have different time horizons for evaluating regulatory threats.

https://doi.org/10.1017/psrm.2025.16 Published online by Cambridge University Press

oil majors to actively lobby for a federal carbon tax. Furthermore, ConocoPhillips' support seems a bit unexpected given that its gas production decreased more than its oil production (E_{0e})—here we do not have countervailing information regarding the company's views on the role of gas in its future production portfolio, in contrast to our state of information about BP. (Regarding the timing of the initiative, the same political considerations discussed under H_{SA} apply under H_{CA}, so this component of E_3 does not discriminate between the hypotheses.)

Overall, we judge E_3 to moderately favor H_{SA}. Beyond the question of whether incentives would be substantial enough for oil majors to actively lobby for a US carbon tax in the H_{CA} world, H_{SA} makes a stronger prediction that the tax should be tied to compensations like the liability waiver and elimination of federal regulations. It is not especially surprising for oil majors and Republican allies to include these riders in the H_{CA} world—they are also in the companies' economic interest and could serve as bargaining chips in political negotiations—but riders are not central for the oil majors' support under H_{CA}, whereas they are central under H_{SA}. ConocoPhillips' lobbying despite falling gas production in turn nudges the weight of evidence a bit above what it would otherwise be if we were to only consider Exxon, Shell, and BP. $WoE_3 \approx 12$ dB for H_{SA} vs. H_{CA}.

Our analysis of E_3 appears to differ substantially from Vormedal *et al.* (2020:144), who report finding "no compelling evidence" of oil majors supporting carbon pricing to hedge against more radical policies. While a moderate weight of 12 dB may not be "compelling," it is the strongest evidence from the article that we have analyzed.[17]

> E_4 = *As per publicly available presentations to investors: In 2018, Exxon anticipated an emerging "golden age" for natural gas, linked both to "policy-aided" coal-to-gas switching and a potentially steep rise in global LNG demand. In 2019, Total also foresaw potential for coal-to-gas switching—more than gas to renewables switching. In 2019, Equinor foresaw significant coal to gas switching in Europe if the carbon price were to exceed 30EU per ton.* (Vormedal *et al.*, 2020:157)

These projections are the kind of thing we would expect if the oil majors' support for moderate carbon pricing (as observed in E_1, E_2, and E_3) results from anticipation of competitive advantages relative to coal, as per our reasoning under H_{CA}. Enthusiasm over coal-to-gas switching is also plausible under H_{SA}—here we would interpret these market anticipations as independent from the strategic accommodation logic that motivated the companies to support moderate carbon pricing in previous years. Augmenting the shift from coal to gas (if the companies had already envisioned that trend in 2015 or 2017) is simply an added side benefit of a moderate carbon price in the H_{SA} world. But in contrast to the statements from Exxon and Total that discuss gas markets without reference to carbon pricing, Equinor's seeming interest in a higher EU carbon price is noticeably more expected under H_{CA}. This is the first evidence that sheds more direct light on a company's cost-benefit calculations regarding carbon pricing. Under H_{SA}, although regulatory pressure remained salient in Europe, we would be somewhat surprised by Equinor linking demand for gas to a higher carbon price. We keep the weight of evidence modest, however, considering that the document cited provides little context or elaboration on Equinor's strategic calculations or policy preferences. $WoE_4 \approx 7$ dB for H_{CA} vs. H_{SA}

We now aggregate to obtain the net weight of evidence (Figure 3). E_1, E_2, and E_4 together give $4 + 3 + 7 = 14$ dB in favor of H_{CA} while E_3 contributes 12 dB in favor of H_{SA}, yielding a total of just 2 dB for H_{CA}. Adding our prior log odds (6 dB for H_{SA}), we arrive at marginally weaker posterior log-odds of 4 dB in favor of H_{SA} vs. H_{CA}, and the inference is dominated by the priors. In sum, the evidence Vormedal *et al.* report affords almost no updating.[18]

[17] Again, however, the weight of evidence depends on how exactly one interprets H_{SA}. A narrower understanding of the hypothesis as pertaining strictly to potential regulatory threats (not existing regulations) would somewhat reduce the inferential weight, although in our view E_3 would still moderately favor H_{SA} for the other reasons discussed above.

[18] Appendix B.4 conducts sensitivity analysis.

https://doi.org/10.1017/psrm.2025.16 Published online by Cambridge University Press

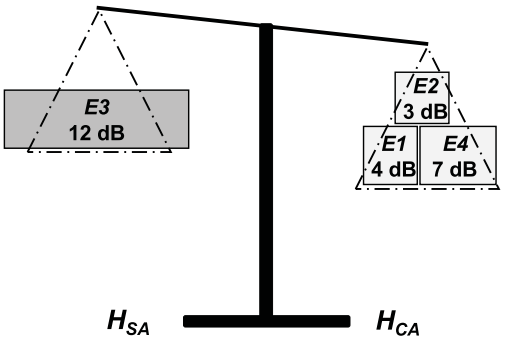


Figure 3. Net weight of evidence: 2 dB for the competitive advantage hypothesis.

In contrast to our analysis, Vormedal *et al.* (2020:161) conclude that "the majors' advocacy for the Paris Agreement and a US carbon tax is unlikely to represent... hedging." We interpret this statement as a judgment that their posterior odds favor H_{CA}, whereas using (modest) prior odds that we believe reflect their review of previous literature, our posterior odds instead favor H_{SA}. And while we would agree that the evidence they present on balance favors H_{CA}, the net weight of evidence does not differ meaningfully from zero. Vormedal *et al.* (2020:161) subsequently assert: "We have found limited grounds for interpreting Big Oil's advocacy as a case of ...hedging," which is a more ambiguous statement. If taken to mean that the aggregate evidence weakly favors H_{SA}, then it is incompatible with their previous inference in favor of H_{CA}, whereas if interpreted as referring to individual pieces of evidence, we would agree only if the adjective "limited" can be viewed as compatible with the "moderate" weight of evidence that we deem E_3 lends to H_{SA} (although in our reading, the authors' largely implicit analysis of E_3 differs ours).

We reiterate that Bayesian scrutiny would ideally be a starting point for further discussion and debate about the weight of evidence. But while we do not intend the analysis conducted here to be taken as definitive, in this instance we would prioritize collecting more evidence, with the goal of accumulating stronger inferential weight.

6. Conclusion

We have argued that Bayesianism provides a cogent framework for scrutinizing the reliability of findings in qualitative research—an analog of replication analysis that involves independently reevaluating to what extent the evidence presented in a study justifies the author's inference. The key Bayesian tenet is that the relative likelihood of the evidence determines its inferential import. Authors often overstate how strongly evidence supports their hypothesis, because they do not consider whether the evidence might be equally likely, or even more likely, under rival hypotheses. This pitfall is a common form of confirmation bias—focusing on a single favored hypothesis (Nickerson, 1998). Explicitly applying Bayesian reasoning averts this pitfall, because the key inferential step always involves evaluating which of two or more rivals makes the evidence more expected.

While Bayesian reasoning can improve inferential judgements, it cannot eliminate individual judgement and subjectivity from qualitative research. No inferential approach could do so when facing inherently qualitative evidence and informal, plain-language hypotheses. Dialog and debate among scholars is therefore especially important for building consensus and promoting knowledge accumulation. To this end, a Bayesian framework makes our reasoning more transparent for others to scrutinize and, when using decibels, helps communicate our judgements about the weight of evidence more clearly. Our approach cannot automatically produce consensus, but even if disagreements persist, we can attain a better understanding of where and why scholars disagree.

https://doi.org/10.1017/psrm.2025.16 Published online by Cambridge University Press

We end with some thoughts on the roles that qualitative Bayesian re-analysis could play in social science. One role lies in teaching. Doctoral training commonly includes quantitative replication exercises to develop data analysis skills and encourage in-depth scrutiny of published research.[19] Qualitative replication exercises are equally important for training social scientists to become better users and consumers of qualitative evidence. Peer review is another context where Bayesian scrutiny would be valuable. And while publishing replication studies may be difficult, we view this endeavor as central to knowledge accumulation and transparency. Replication work, whether qualitative or quantitative, merits publication when the original study is prominent and speaks to important questions but has not drawn inferences that stand up to scrutiny. The *APSR*'s present policy on replication shares a similar perspective.[20]

Bayesian re-analysis could also be undertaken on a larger scale, involving multiple scholars. For example, a panel of experts could be assembled to debate evidentiary weight and work toward a consensus view. Scholars with different background knowledge can rationally assign different probabilities, so it would be sensible to start by delineating a common body of background information upon which to condition all probabilities. Participants should then analyze the evidence independently before comparing their reasoning about relative likelihoods and then discuss until reaching (rough) agreement. If using the log-odds formulation, participants should revisit their decibel assignments after agreeing on the qualitative likelihood ratio reasoning. A consensus weight of evidence could then emerge through further deliberation. Alternatively, individual weights of evidence could simply be averaged, while also reporting the spread of decibels within the group.

Bayesian re-analysis can also be deployed in the context of meta-analysis. Bayesian replication and meta-analysis both involve examining evidence from published research and re-evaluating its inferential weight. But a Bayesian meta-analysis would be much broader in scope, with the goal of combining evidence from multiple studies to characterize the state of knowledge in a field, rather than focusing on evidence from a single study. We plan to pursue this agenda in future research by bringing evidence from other studies to bear on the global collective action vs. distributive politics debate discussed in Section 3 and by extending our analysis to different models of distributive politics.

Finally, Bayesian reanalysis could encourage the qualitative methods community to reconsider the proliferation of approaches that propose very different inferential logics. This state of affairs encourages fragmentation, downplays the importance of shared standards, and hinders knowledge accumulation—as Seawright (2016) has argued, "triangulation" across methods that invoke incompatible epistemological foundations is problematic at best. Conducting Bayesian re-analysis on studies that employ non-Bayesian methods could shed light on the extent to which different methodological approaches produce divergent and incompatible inferences. In confronting the question of which inference to trust, scholars might then feel more compelled to ask which approach is grounded on the most defensible foundations.

In our view, logical Bayesianism is the optimal choice. A fundamental set of desiderata for rational reasoning—individuals with identical information should hold the same degrees of belief; learning or analyzing information in different but logically equivalent ways must produce identical conclusions; and the rules for reasoning should reduce to Boolean logic in the limit of complete information—is sufficient to derive Bayes' rule and all other rules of probability (e.g., Jaynes, 2003). In short, probability theory is the unique extension of deductive logic to situations with imperfect information and uncertainty—it provides a rigorous and compelling framework for inference.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/psrm.2025.16.

[19]See for example Simon Hix's course: https://www.eui.eu/Documents/DepartmentsCentres/SPS/Seminars/2024-25-Seminars/Hix-Replicating-Research-in-Poli-Sci-outline-2024-25.pdf

[20]https://www.cambridge.org/core/journals/american-political-science-review/article/notes-from-the-editors/51EECD4AFBA415AEEB241B6E3FA48ED1

https://doi.org/10.1017/psrm.2025.16 Published online by Cambridge University Press

Acknowledgements. We wish to thank Robert Falkner for helpful conversations on the Big Oil case, as well as participants in workshops we taught in 2024 at EUI SPS and GSERM St. Gallen for comments on early drafts. We are especially grateful to two anonymous *PSRM* reviewers for their valuable suggestions and remarkably high-level engagement with our manuscript.

Competing interests. The authors declare none.

References

Aklin M and Mildenberger M (2020) Prisoners of the wrong dilemma: Why distributive conflict, not collective action, characterizes the politics of climate change. *Global Environmental Politics* **20**, 4–23.

Aklin M and Mildenberger M (2022) Reply: The persistent absence of empirical evidence for free-riding in global climate politics. *Global Environmental Politics* **23**, 145–151.

Bennett A (2015) *Disciplining our conjectures: Systematizing process tracing with Bayesian analysis*. In Andrew B and Jeffrey C (eds), *Process Tracing in the Social Sciences*, New York: Cambridge University Press, pp.276–298.

Fairfield T and Charman A (2020) *Reliability of inference*. In C. Elman, J. Gerring & J Mahoney (eds), *The Production of Knowledge* (Chapter 12), Cambridge: Cambridge University Press.

Fairfield T and Charman A (2022) *Social Inquiry and Bayesian Inference*, New York: Cambridge University Press.

Falkner R (2008) *Business Power and Conflict in International Environmental Politics*, New York: Palgrave Macmillan.

Freese J and Peterson D (2020) *Replication for quantitative research*. In Elman C. C., J. Gerring & J. Mahoney (eds), *The Production of Knowledge* (Chapter 10), Cambridge: Cambridge University Press.

Good IJ (1983) *Good Thinking*, Minneapolis: University of Minnesota Press.

Gregory P (2005) *Bayesian Logical Data Analysis for the Physical Sciences*, Cambridge: Cambridge University Press.

Hacker J and Pierson P (2010) Winner-take-all-politics. *Politics & Society* **38**, 152–204.

Hunter D (1984) *Political/Military Applications of Bayesian Analysis*, Boulder: Westview Press.

Jaynes ET (2003) *Probability Theory: The Logic of Science*, Cambridge: Cambridge University Press.

Jeffreys H (1939) *Theory of Probability*, Oxford: Clarendon Press.

Kelsey N (2021) International ozone negotiations and the green spiral. *Global Environmental Politics* **21**, 64–87.

King G, Keohane R and Verba S (1994) *Designing Social Inquiry*, Princeton: Princeton University Press.

Meckling J (2015) Oppose, support or hedge? *Global Environmental Politics* **15**, 19–37.

Nickerson R (1998) Confirmation bias. *Review of General Psychology* **2**, 175–220.

Seawright J (2016) *Multi-Method Social Science*, New York: Cambridge University Press.

Sun Y and Browen Y (2023) Greening China's Belt and Road Initiative. *Global Environmental Politics* **23**, 91–116.

Vogel D (2018) *California Greenin'*, Princeton: Princeton University Press.

Vormedal I, Gulbrandsen L and Skjaeresth JB (2020) Big oil and climate regulation. *Global Environmental Politics* **20**, 143–166.

https://doi.org/10.1017/psrm.2025.16 Published online by Cambridge University Press

Cite this article: Fairfield T and Charman A (2026) Bayesian reasoning for qualitative replication analysis: Examples from climate politics. *Political Science Research and Methods* **14**, 414–429. https://doi.org/10.1017/psrm.2025.16