Qualitative Robust Bayesianism and the Likelihood Principle

Conor Mayo-Wilson and Aditya Saraf*

September 8, 2020

Abstract

We argue that the likelihood principle (LP) and weak law of likelihood (LL) generalize naturally to settings in which experimenters are justified only in making comparative, non-numerical judgments of the form "A given B is more likely than C given D." To do so, we first *formulate* qualitative analogs of those theses. Then, using a framework for qualitative conditional probability, just as the characterizes when all Bayesians (regardless of prior) agree that two pieces of evidence are equivalent, so a qualitative/non-numerical version of LP provides sufficient conditions for agreement among experimenters' whose degrees of belief satisfy only very weak "coherence" constraints. We prove a similar result for LL. We conclude by discussing the relevance of results to stopping rules.

The central question of this paper is, "Which controversial statistical principles (e.g., likelihood, conditionality, etc.) are special cases of principles for good scientific reasoning?" We focus on the likelihood principle (Birnbaum, 1962; Berger and Wolpert, 1988), the weak law of likelihood (Sober, 2008), and the strong law of likelihood (Edwards, 1984; Hacking, 1965; Royall, 1997), as these principles are critical for foundational debates in statistics.

Our central question is motivated by the following observation. Despite the ubiquity of statistics in contemporary science, many successful instances of scientific reasoning require only comparative, non-numerical judgments of evidential strength. For instance, when Lavoisier judged that his data provided good evidence against the phlogiston theory of combustion and for the existence of what we now call "oxygen", no statistics was invoked. Statistical reasoning, therefore, is just one type of scientific reasoning, and so norms for statistical reasoning should be special cases of norms for good scientific reasoning.

Our thesis is that likelihood principle (LP) and *weak* law of likelihood (LL) generalize naturally to settings in which experimenters are justified only in making comparative, non-numerical judgments of the form "A given B is more likely than C given D." Specifically, our main results show that, just as LP characterizes when all Bayesians (regardless of prior) agree that two pieces of evidence are

*Contact: conormw@uw.edu and sarafa@cs.washington.edu

equivalent, a qualitative/non-numerical version of LP provides sufficient conditions for agreement among experimenters' whose degrees of belief satisfy only very weak "coherence" constraints (i.e., ones that do not entail the probability axioms). We prove a similar result for LL. In contrast, the *strong* law of likelihood (LL$^+$) – which asserts that the likelihood ratio is a measure of evidential strength – has no plausible qualitative analog in our framework.

Our results are important for three related reasons. First, several purported counterexamples to LP are, we think, more accurately interpreted as objections to LL$^+$. We discuss one such example in §1. Because our results show that LP generalizes to qualitative settings in ways LL$^+$ may not, there is good reason to distinguish the two theses carefully.

Second, our results provide further reason to endorse Berger and Wolpert (1988, p. 141)'s conclusion that "the only satisfactory method of [statistical] analysis based on LP seems to be robust Bayesian analysis." In standard/quantitative statistical settings, only LL$^+$, we think, provides a foundation for non-Bayesian, likelihoodist techniques; the other two likelihoodist theses (i.e., LP and LL) provide little reason to use maximum-likelihood estimation or related techniques. Unfortunately, the plausibility of LL$^+$ might be a modeling artifact: when likelihood functions are assumed to be real-valued, LL$^+$ is equivalent to another plausible principle for evidential reasoning, which we call the "robust Bayesian support principle." That equivalence, however, breaks down in more general, qualitative settings. In such settings, the robust Bayesian favoring principle remains plausible, but LL$^+$ cannot even be articulated.

Finally, LP *sometimes* recommends against the use of p-values, confidence intervals, and other classical/frequentist summaries of data. Therefore, if LP is a special case of a more general, qualitative principle of scientific reasoning, then *some* frequentist techniques will conflict with more general principles of scientific reasoning. In the final section of the paper, we discuss one example of potentially problematic frequentist reasoning; we show that one argument for the irrelevance of stopping rules extends naturally to non-statistical data.

1 Likelihoodism

At least three distinct theses are called "the likelihood principle" and are used to motivate likelihood-based methods (e.g., maximum likelihood estimation) and Bayesian tools. For clarity, we distinguish the three:

- Likelihood principle (LP): Let E and F are outcomes of two experiments $\mathbb{E}$ and $\mathbb{F}$ respectively. If there is some $c > 0$ such that $P_\theta^{\mathbb{E}}(E) = c \cdot P_\theta^{\mathbb{F}}(F)$ for all $\theta \in \Theta$, then E and F are evidentially equivalent.

- Weak Law of likelihood (LL): $P_{\theta_1}^{\mathbb{E}}(E) > P_{\theta_2}^{\mathbb{E}}(E)$ if and only if the data E favors θ_1 over θ_2.

- Strong Law of likelihood (LL$^+$): The likelihood ratio $P_{\theta_1}^{\mathbb{E}}(E)/P_{\theta_2}^{\mathbb{E}}(E)$ quantifies the degree to which the data E favors θ_1 over θ_2.

For likelihoodists (who endorse the above principles but do not embrace Bayesianism), the notions of "evidential equivalence" and "favoring" are undefined primitives that are axiomatized by principles like LP and LL. Likelihoodists, therefore, defend LP, LL, and LL$^+$ by arguing the three principles accord with (i) our informed reflections about evidential strength, (ii) accepted statistical methods that have been successfully applied in science, and (iii) other plausible and less-controversial "axioms" for statistical inference (e.g., the sufficiency principle).

Notice that LL and LL$^+$ concern only *simple/point* hypotheses (i.e., elements of Θ); the extension of those theses to *composite* hypotheses (i.e., subsets of Θ) is controversial, especially when nuisance parameters are present (see e.g., Royall (1997) [§1.7 and Chapter 7] and Bickel (2012) for different strategies). LL$^+$ is stronger than LL in two related ways: it allows one to (1) compare the *strength* of two pieces of evidence E and F by assessing whether $P^{\mathbb{E}}_{\theta_1}(E)/P^{\mathbb{E}}_{\theta_2}(E) \geq P^{\mathbb{F}}_{\theta_1}(F)/P^{\mathbb{F}}_{\theta_2}(F)$ or vice versa, and (2) compare pieces of evidence drawn from different experiments. Two examples will clarify and distinguish the theses.

Example 1: Suppose two experiments are designed to distinguish $\Theta = \{\theta_1, \theta_2\}$. The possible outcomes of the experiments (A, B, C, A', etc.) are the column headers in the tables below, and the likelihood functions are the column vectors. We intentionally list only two possible outcomes of Experiment $\mathbb{F}$.

Experiment $\mathbb{E}$

	A	B	C
θ_1	.423	.564	.011
θ_2	.039	.052	.909

Experiment $\mathbb{F}$

	A'	B'	...
θ_1	.846	.12	...
θ_2	.078	.01	...

According to LP, the outcomes A and B from $\mathbb{E}$ are evidentially equivalent to one another because $P^{\mathbb{E}}_{\theta_1}(A) = 3/4 \cdot P^{\mathbb{E}}_{\theta_1}(B)$ for all $\theta \in \{\theta_1, \theta_2\}$. Similarly, both A and B are evidentially equivalent to observing A' in $\mathbb{F}$ as $P^{\mathbb{E}}_{\theta_1}(A) = 1/2 \cdot P^{\mathbb{F}}_{\theta_1}(A')$ for all θ. However, LP says nothing about which of the hypotheses in Θ are favored by A, B, and B'. In general, LP tells one only when two samples should yield identical estimates, but it says nothing about what those estimates should be.

According to LL, the outcome A "favors" θ_1 over θ_2 because $P^{\mathbb{E}}_{\theta_1}(A) > P^{\mathbb{E}}_{\theta_2}(A)$. Similarly, LL entails that B favors θ_1 over θ_2. However, LL tells one nothing about how strong those pieces of evidence are, and in particular, whether A and B provide evidence of equal or different strength.

In contrast, LL$^+$ entails that A and B favor θ_1 to θ_2 to the same degree, and both A and B are weaker pieces of evidence than than B' because $P^{\mathbb{E}}_{\theta_1}(A)/P^{\mathbb{E}}_{\theta_2}(A) < 12 = P^{\mathbb{F}}_{\theta_1}(B')/P^{\mathbb{F}}_{\theta_2}(B')$. Notice that both LL$^+$ and LP allow one to draw conclusions about evidential value *without knowing the sample space of the second experiment.* It is for this reason that authors often stress that likelihoodist principles entail the "irrelevancy of the sample space" (Royall, 1997, §1.11).

(End Example)

Failure to distinguish the above theses carefully, we think, has caused confusion in debates about LP. For instance, (Evans et al., 1986) offer the following example as a challenge to LP; the example is due to (Fraser et al., 1984).

Example 2: Suppose $\Theta = \Omega = \mathbb{N}^+ = \{1, 2, \dots\}$ and that P_θ assigns equal probability of $1/3$ to each of the values in $\{\lfloor \theta/2 \rfloor, 2\theta, 2\theta + 1\}$. Here, $\lfloor x \rfloor$ rounds x down to the nearest integer except when $x = 1/2$, in which case $\lfloor 1/2 \rfloor = 1$. Fraser et al. (1984) note that, surprisingly, although the likelihood function is in some sense "flat", the estimator $\hat{\theta}_L(n) = \{\lfloor n/2 \rfloor\}$ gives a $2/3$ confidence set, whereas the estimators $\hat{\theta}_M(n) = \{2n\}$ and $\hat{\theta}_H(\omega) = \{2n+1\}$ each have coverage only $1/3$. Evans et al. (1986) conclude, "Examples such as this would seem to make the likelihood principle questionable for statistical inference."

However, in this example, LP implies only the unobjectionable claim that each $\omega \in \Omega$ is evidentially equivalent to itself. To see why, suppose $n, m \in \Omega$ are distinct, and so we may assume $n > m$ without loss of generality. Recall that LP says that n and m are evidentially equivalent if there is some constant c such that $P_\theta(n) = c \cdot P_\theta(m)$ for *all* θ. To see that LP fails to entail that n and m are evidentially equivalent, let $\theta_0 = 2n$. Then, $P_{\theta_0}(n) = 1/3$, while $P_{\theta_0}(m) = 0$ since $m < \lfloor \theta_0/2 \rfloor = n$. Thus, there is no $c > 0$ such that $P_\theta(n) = c \cdot P_\theta(m)$ for all θ. Hence, LP says nothing about the relationship between distinct outcomes.

One might think the example challenges LL and LL$^+$. Because only $\hat{\theta}_L(n) = \{\lfloor n/2 \rfloor\}$ gives a $2/3$ confidence set, one might infer that observing n favors $\lfloor n/2 \rfloor$ over $2n$ and $2n + 1$. But LL and LL$^+$ entail that n does not favor any of the parameters $\lfloor n/2 \rfloor, 2n$ or $2n+1$ over the others. Although we think the example is not a problem for LL or LL$^+$, we will not argue so here. For now, we stress that the example shows the importance of distinguishing LP from LL$^+$ and LL.

(End Example)

The failure to distinguish LL from LP, we conjecture, is a result of assuming that the mathematical formulations of LP (like Birnbaum (1962)'s above) capture certain informal statements of the principle. Many practitioners quote Berger and Wolpert (1988, p. 1)'s informal gloss of LP; they say that LP "essentially states that all evidence, which is obtained from an experiment, about an unknown quantity θ, is contained in the likelihood function of θ for the given data." Yet other informal glosses of LP differ significantly. For example, Birnbaum (1962, p. 271) says LP asserts the "irrelevance of [experimental] outcomes not actually observed." Although these informal statements capture important insights about likelihoodist theses, we think care should be taken to stick to the technical statements when evaluating particular examples, as shown above.

Distinguishing the three theses is also important for understanding which statistical tests and techniques are justified by each. For instance, as we noted above, the formal statement of LP says nothing about which estimates are favored by a sample. LL entails that the MLE is always favored over rivals, but it does not say by how much. Thus, LL is of little use when the fit of the MLE needs to be weighed against considerations of simplicity and prior plausibility.

Only $\textsc{ll}^+$, we think, can be used to justify likelihoodist estimation procedures, but sadly, it is also the only thesis that has no obvious qualitative analog in our framework below.

2 Baysianism and Likelihoodism

$\textsc{ll}$, $\textsc{lp}$, and $\textsc{ll}^+$ are often said to be "compatible" with Bayes rule (Edwards, 1984, p. 28). Here is how that "compatibility" is often explained for $\textsc{ll}^+$. Think of an experiment $\mathbb{E}$ as a pair $\langle \Omega^{\mathbb{E}}, \{P^{\mathbb{E}}_\theta\}_{\theta \in \Theta}\rangle$, where $\Omega^{\mathbb{E}}$ represents the possible outcomes of the experiment and $P^{\mathbb{E}}_\theta(\cdot)$ is a probability distribution over $\Omega^{\mathbb{E}}$ that specifies how likely each outcome is if θ is the true value of the parameter. Suppose Q is a prior probability distribution over Θ, which we will assume is finite for the remainder of the paper to avoid measurability assumptions. Define the posterior in the standard way:

$$Q^{\mathbb{E}}(H|E) = \frac{Q^{\mathbb{E}}(E \cap H)}{Q^{\mathbb{E}}(E)} := \frac{\sum_{\theta \in H} P^{\mathbb{E}}_\theta(E) \cdot Q(\theta)}{\sum_{\theta \in \Theta} P^{\mathbb{E}}_\theta(E) \cdot Q(\theta)}$$

for any $H \subseteq \Theta$. Then Bayes' Rule entails

$$\frac{Q^{\mathbb{E}}(\theta_1|E)}{Q^{\mathbb{E}}(\theta_2|E)} = \frac{P^{\mathbb{E}}_{\theta_1}(E)}{P^{\mathbb{E}}_{\theta_2}(E)} \cdot \frac{Q(\theta_1)}{Q(\theta_2)}$$

Although this calculation is standard in introductory remarks about Bayes rules (e.g., see (Gelman et al., 2013, p. 8)), we note that it holds *for any prior* Q. So the likelihood ratio is a measure of the degree to which *all* Bayesians' posterior degrees of belief in θ_1 increase (or decrease) upon learning E. We introduce another way of capturing that same idea below.

We would like to suggest one new argument for $\textsc{ll}^+$, as it will be important below. $\textsc{ll}^+$ *unifies* $\textsc{ll}$ and $\textsc{lp}$, in the sense that it entails both theses. Why? Assume that, for every piece of evidence E and any two hypotheses θ_1 and θ_2, there is a numerical degree $\mathsf{deg}(E, \theta_1, \theta_2)$ to which E favors θ_1 over θ_2 (here, the degree might be negative). To show $\textsc{ll}^+$ entails $\textsc{ll}$ under plausible assumptions, say that E *favors* θ_1 over θ_2 if $\mathsf{deg}(E, \theta_1, \theta_2) > \mathsf{deg}(\Omega, \theta_1, \theta_2)$, in other words, if E provides better evidence for θ_1 (over θ_2) than the sure event. If $\textsc{ll}^+$ holds, then $\mathsf{deg}(E, \theta_1, \theta_2) = P_{\theta_1}(E)/P_{\theta_2}(E)$, and so E favors θ_1 over θ_2 precisely if $P_{\theta_1}(E)/P_{\theta_2}(E) = \mathsf{deg}(E, \theta_1, \theta_2) > \mathsf{deg}(\Omega, \theta_1, \theta_2) = P_{\theta_1}(\Omega)/P_{\theta_2}(\Omega) = 1$. In other words, $\textsc{ll}^+$ plus the above assumptions entails that E favors θ_1 over θ_2 precisely if $P_{\theta_1}(E) > P_{\theta_2}(E)$, exactly as $\textsc{ll}$ asserts.[1]

To show $\textsc{ll}^+$ entails $\textsc{lp}$ under plausible assumptions, say that E and F are *evidentially equivalent* if $\mathsf{deg}(E, \theta_1, \theta_2) = \mathsf{deg}(F, \theta_1, \theta_2)$ for all θ_1 and θ_2. In other words, E and F favor all hypotheses by equal amounts. Again, if $\textsc{ll}^+$ holds, then $\mathsf{deg}(E, \theta_1, \theta_2) = P_{\theta_1}(E)/P_{\theta_2}(E)$, and so E and F are evidentially equivalent precisely if

$$\frac{P_{\theta_1}(E)}{P_{\theta_2}(E)} = \frac{P_{\theta_1}(F)}{P_{\theta_2}(F)} \tag{1}$$

[1]Notice, we drop the superscript $\mathbb{E}$ when it's clear from context.

for all θ_1 and θ_2. If there is $c > 0$ such that $P_\theta(E) = c \cdot P_\theta(F)$ for all θ, then holds. So LL$^+$ entails LP.

In sum, (i) LL$^+$ is "compatible" with Bayes rule in the sense that the likelihood ratio is also the ratio of *all* Bayesians' posterior to prior probabilities, and (ii) LL$^+$ entails both LL and LP, which are also thought to be intuitively plausible (and are "compatible" with Bayesianism in still other ways). In the next section, we first prove that LL$^+$ is compatible with Bayes rule in another important way: it characterizes when *all* Bayesians agree that E is better evidence than F. In the second half of the paper, we show that the qualitative analog of such Bayesian agreement unifies qualitative analogs of LL and LP in the same way LL$^+$ unifies quantitative versions of those principles.

2.1 Robust Bayesianism and Likelihoodism

Robust Bayesianism is roughly the thesis that statistical decisions should be stable under a variety of different prior distributions (Berger, 1990; Kadane, 1984). In this section, we discuss three elementary propositions that show likelihoodist theses characterize when all Bayesians agree about how evidence ought to change one's posterior. The propositions, we think, clarify why advocates of robust Bayesian analysis have found likelihoodist theses so attractive. The main results of our paper – in the next section – are the qualitative analogs of the second and third claim below.

Before stating our main definition, we introduce some notation. Given an experiment $\mathbb{E}$, we let $\Delta^{\mathbb{E}} = \Theta \times \Omega^{\mathbb{E}}$. We use H_1, H_2 etc. to denote subsets of Θ, and E, F, etc. to denote subsets of Ω. Again, we drop the superscript $\mathbb{E}$ when it is clear from context. We will write $Q(\cdot|H)$ and $Q(\cdot|E)$ instead of $Q(\cdot|H \times \Omega)$ and $Q(\cdot|\Theta \times E)$, and similarly for events to the left of the conditioning bar.

Definition 1. Suppose $H_1, H_2 \subseteq \Theta$ are disjoint. Let E and F be outcomes of experiments $\mathbb{E}$ and $\mathbb{F}$ respectively. Say E *Bayesian supports* H_1 over H_2 *at least as much as* F if $Q^{\mathbb{E}}(H_1|E \cap (H_1 \cup H_2)) \geq Q^{\mathbb{F}}(H_1|F \cap (H_1 \cup H_2))$ for all priors Q for which $Q^{\mathbb{F}}(\cdot|F \cap (H_1 \cup H_2))$ is well-defined. If the inequality is strict for all such Q, then we say E *Bayesian supports* H_1 over H_2 *strictly more than* F. In the former case, we write $E \,{}_{H_1}\!\overset{\mathcal{B}}{\trianglerighteq}_{H_2} F$, and in the latter, we write $E \,{}_{H_1}\!\overset{\mathcal{B}}{\gg}_{H_2} F$.

Our first goal is to find a necessary and sufficient condition for Bayesian support that involves only likelihoods. The following claim does exactly that.

Claim 1. Suppose H_1 and H_2 are finite and disjoint. Let E and F be outcomes of experiments $\mathbb{E}$ and $\mathbb{F}$ respectively. Then $E \,{}_{H_1}\!\overset{\mathcal{B}}{\trianglerighteq}_{H_2} F$ if and only if (1) for all $\theta \in H_1 \cup H_2$, if $P^{\mathbb{F}}_\theta(F) > 0$, then $P^{\mathbb{E}}_\theta(E) > 0$, and (2) for all $\theta_1 \in H_1$ and $\theta_2 \in H_2$:

$$P^{\mathbb{E}}_{\theta_1}(E) \cdot P^{\mathbb{F}}_{\theta_2}(F) \geq P^{\mathbb{E}}_{\theta_2}(E) \cdot P^{\mathbb{F}}_{\theta_1}(F)$$

Similarly, $E \,{}_{H_1}\!\overset{\mathcal{B}}{\gg}_{H_2} F$ if and only if 1 holds and the inequality in the above equation is always strict.

We omit the proof, as it uses only basic probability theory. However, details of all omitted proofs (including the elementary ones) in this paper can be found in the appendix. The claim asserts that $E \,{}_{H_1}\!\overset{\mathcal{B}}{\trianglerighteq}_{H_2} F$ exactly when LL$^+$ says that E provides stronger evidence for θ_1 over θ_2 than F does. The Bayesian support relation, therefore, is one way of clarifying how a robust Bayesian can interpret the notion of "favoring" in likelihoodist theses like LL$^+$. This is important because, as we show below, the Bayesian support relation has a direct qualitative analog, unlike LL$^+$, which requires numerical degrees of favoring.

The Bayesian support relation is also important because it can be used to define the following notions of "Bayesian favoring" and "Bayesian favoring equivalence" that are respectively equivalent to the notions of "favoring" in LL and "evidential equivalence" in LP. This should be unsurprising given our argument above that LL$^+$ can be used to derive LL and LP under plausible assumptions.

Definition 2. *E Bayesian favors H_1 to H_2 if $E \,{}_{H_1}\!\overset{\mathcal{B}}{\trianglerighteq}_{H_2} \Omega$. Say it strictly does if $E \,{}_{H_1}\!\overset{\mathcal{B}}{\gg}_{H_2} \Omega$.*

So Bayesian favoring is a special case of Bayesian support. In essence, the support relation *compares* (a) how much E Bayesian favors H_1 to H_2 to (b) how much F Bayesian favors H_1 to H_2. Thus, the next claim, which shows that Bayesian favoring is equivalent to LL, follows directly from Claim 1.

Claim 2. E Bayesian favors H_1 to H_2 if and only if (1) $P_\theta(E) > 0$ for all $\theta \in H_1 \cup H_2$ and (2) $P_{\theta_1}(E) \geq P_{\theta_2}(E)$ for all $\theta_1 \in H_1$ and $\theta_2 \in H_2$. If $H_1 = \{\theta_1\}$ and $H_2 = \{\theta_2\}$ are simple, E strictly Bayesian favors H_1 to H_2 if and only if LL entails E favors H_1 to H_2.

Next, we study two robust Bayesian notions of evidential equivalence that are easy to translate to the qualitative setting.

Definition 3. Say E and F are *Bayesian posterior equivalent* if all priors Q (1) $Q^{\mathbb{E}}(\cdot|E)$ is well-defined if and only if $Q^{\mathbb{F}}(\cdot|F)$ is, and (2) $Q^{\mathbb{E}}(H|E) = Q^{\mathbb{F}}(H|F)$ for all hypotheses H and for which those conditional probabilities are well-defined.

Definition 4. *E and F are Bayesian favoring equivalent if $F \,{}_{H_1}\!\overset{\mathcal{B}}{\trianglerighteq}_{H_2} E$ and $E \,{}_{H_1}\!\overset{\mathcal{B}}{\trianglerighteq}_{H_2} F$ for all disjoint hypotheses H_1 and H_2. In other words, E and F are Bayesian favoring equivalent if for any prior Q and any disjoint hypotheses H_1 and H_2 we have (1) $Q^{\mathbb{E}}(\cdot|E \cap (H_1 \cup H_2))$ is well-defined if and only if $Q^{\mathbb{F}}(\cdot|F \cap (H_1 \cup H_2))$ is well-defined and (2) $Q^{\mathbb{E}}(H_1|E \cap (H_1 \cup H_2)) = Q^{\mathbb{F}}(H_1|E \cap (H_1 \cup H_2))$ whenever those conditional probabilities are well-defined.*

It is well-known that, if LP entails E and F are evidentially equivalent, then they are posterior equivalent (Edwards et al., 1984, p. 56); that fact makes precise in what sense LP is compatible with Bayesianism. This is one reason we conjecture that some researchers (e.g., see Example 3 of Wechsler et al. (2008)) have isolated posterior equivalence as a relation worthy of study. It is easy to show that posterior and favoring equivalence are identical relations.

Claim 3. The following are equivalent: (1) LP entails E and F are evidentially equivalent, (2) E and F are Bayesian posterior equivalent, (3) E and F are Bayesian favoring equivalent.

The three claims above show that there is an alternative way of explaining why (1) LL$^+$ and LP are intuitively plausible in many cases and (2) LL$^+$ "unifies" LL and LP. To see why, consider one last philosophical thesis about statistical evidence.

Robust Bayesian Support Principle: E provides at least as good statistical evidence for H_1 over H_2 as F if $E \; {}_{H_1}\!\overset{\mathcal{B}}{\trianglerighteq}_{H_2} F$.

By Claim 1, the robust Bayesian support principle is equivalent to LL$^+$. And by Claim 2 and Claim 3, it entails LL and LP with some plausible, additional assumptions, thereby "unifying" LL and LP in roughly the same way LL$^+$ does. Thus, LL$^+$ may be plausible only because it is equivalent to the robust Bayesian support principle in common statistical settings. And the robust Bayesian support principle is plausible, we conjecture, precisely because it captures the idea that evidence persuades *all* rational parties to change their beliefs in particular ways. This is important because, as we now show, the robust Bayesian principle generalizes to qualitative settings in which LL$^+$ cannot be formulated.

3 Qualitative Likelihoodism

3.1 Key Concepts

To move from quantitative to qualitative probability, we replace probability functions with two *orderings*. As before, let Θ be the set of simple hypotheses, and for any experiment $\mathbb{E}$, we let $\Omega^{\mathbb{E}}$ the set of experimental outcomes. The first relation, $\sqsubseteq^{\mathbb{E}}$, is the qualitative analog of the set of likelihood functions. As before, we drop $\mathbb{E}$ when it's clear from context. Informally, $A|\theta \sqsubseteq B|\eta$ represents the claim that "experimental outcome B is at least as likely under supposition η as outcome A is under supposition θ"; it is the qualitative analog of $P_\theta(A) \leq P_\eta(B)$. We write $A|\theta \equiv B|\theta$ if $A|\theta \sqsubseteq B|\theta$ and vice versa.

Although we typically consider expressions of the form $A|\theta \sqsubseteq B|\eta$, the $\sqsubseteq$ relation is also defined when experimental outcomes $E \subseteq \Omega$ appear to the right of the conditioning bar, i.e., $A|\theta \cap E \sqsubseteq A|\theta \cap F$ is a well-defined expression if $E, F \subseteq \Omega$. However, it is not-well defined when composite hypotheses H appear

Quantitative Probabilistic Notions	**Qualitative Analog**
$P^{\mathbb{E}}_\theta(E) \leq P^{\mathbb{F}}_\upsilon(F)$	$E\|\theta \sqsubseteq F\|\upsilon$
$Q^{\mathbb{E}}(H_1\|E) \leq Q^{\mathbb{F}}(H_2\|F)$	$H_1\|E \preceq H_2\|F$
LL $\Leftrightarrow E$ Bayesian favors H_1 to H_2	QLL $\Leftrightarrow E$ qualitatively favors H_1 to H_2.
LP $\Leftrightarrow E$ and F are Bayesian posterior and favoring equivalent	QLP$^-$ $\Rightarrow E$ and F are qualitatively posterior and favoring equivalent.

to the right of the conditioning bar, just as there are no likelihood functions $P_H(\cdot)$ for composite hypotheses in the quantitative case.

Bayesians assume that beliefs are representable by a probability function. We weaken that assumption and assume beliefs are representable by an ordering $\preceq$ on $\mathcal{P}(\Delta) \times \mathcal{P}(\Delta)$, where $\mathcal{P}(\Delta)$ is the power set of Δ. As before, if $E \subseteq \Omega$ and $H \subseteq \Theta$ we write $E|H$ instead of $\Theta \times E|H \times \Omega$. In the special case in which $H = \{\theta\}$ is a singleton, we omit the curly brackets and write $E|\theta$ instead of $E|\{\theta\}$. We write $A|B \sim C|D$ is $A|B \preceq C|D$ and vice versa.

The notation is suggestive; $A|B \preceq C|D$ if the experimenter regards C as at least as likely under supposition D as A would be under supposition B. Just as Bayesians condition on both composite hypotheses and experimental outcomes, there are no restrictions on what can appear to the right of the conditioning bar in expressions involving $\preceq$.

Clearly, to prove anything, we must assume that $\sqsubseteq$ and $\preceq$ satisfy certain constraints; those constraints are specified in Section 3.3. But before stating those constraints, we state our main results.

3.2 Main Results

It is now easy to state a qualitative analog of LL:

- Qualitative law of likelihood (QLL): E favors θ_1 to θ_2 if $E|\theta_2 \sqsubset E|\theta_1$.

Claim 2 asserts that LL characterizes precisely when E Bayesian favors θ_1 over θ_2, i.e., when all Bayesian agents agree E favors one hypothesis over another. So by analogy to the definition of. "Bayesian support" and "Bayesian favoring", define:

Definition 5. E *qualitatively favors* H_1 to H_2 at least as much as F if $H_1|E \cap (H_1 \cup H_2) \succeq H_1|F \cap (H_1 \cup H_2)$ for all orderings $\preceq$ satisfying the axioms below and for which the expression $\cdot|F \cap (H_1 \cup H_2)$ is well-defined.

Definition 6. E *qualitatively favors* H_1 to H_2 if E supports H_1 to H_2 at least as much as Ω.

Our first major result is the qualitative analog of Claim 2. Just as Claim 2 entails that E Bayesian favors θ_1 to θ_2 when LL entails so, our first major result shows that E qualitatively favors θ_1 to θ_2 if and only if QLL entails so. Under mild assumptions, this equivalence can be extended to finite composite hypotheses.

Theorem 1. Suppose H_1 and H_2 are finite. Then E qualitatively favors H_1 over H_2 if (1) $\emptyset|\theta \sqsubset E|\theta$ for all $\theta \in H_1 \cup H_2$ and (2) $E|\theta_2 \sqsubseteq E|\theta_1$ for all $\theta_1 \in H_1$ and $\theta_2 \in H_2$. Under Assumption 1 (below), the converse holds as well. In both directions, the favoring inequality is strict exactly when the likelihood inequality is strict. If $H_1 = \{\theta_1\}$ and $H_2 = \{\theta_2\}$ are simple, then E qualitatively favors H_1 over H_2 if and only if QLL entails so. No additional assumptions are required in this case.

Assumption 1 below more-or-less says that one's "prior" ordering over the hypotheses is unconstrained by the "likelihood" relation $\subseteq$. This is exactly analogous to the quantitative probability theory. In the quantitative case, a joint distribution $Q^{\mathbb{E}}$ on $\Delta = \Theta \times \Omega^{\mathbb{E}}$ is determined (i) the measures $\langle P_\theta(\cdot)\rangle_{\theta\in\Theta}$ over experimental outcomes $\Omega^{\mathbb{E}}$ and (ii) one's prior Q over the hypotheses Θ. Although the joint distribution $Q^{\mathbb{E}}$ is constrained by $\langle P_\theta(\cdot)\rangle_{\theta\in\Theta}$, one's prior Q on Θ is not, and so for any non-empty hypothesis $H \subseteq \Theta$, one can define some Q such that $Q(H) = 1$ and $Q(\theta) > 0$ for all $\theta \in H$. That's what the following assumption says in the qualitative case.

Assumption 1. For all orderings $\sqsubseteq$ satisfying the axioms above and for all non-empty $H \subseteq \Theta$, there exists an ordering $\preceq$ satisfying the axioms such that (A) $H|\Delta \sim \Delta|\Delta$ and (B) $\theta|\Delta \succ \emptyset|\Delta$ for all $\theta \in H$.

We believe Assumption 1 is provable, but we have not yet produced a proof.

Our second main result is a qualitative analog of Claim 3, which says that LP characterizes when two pieces of evidence are posterior and Bayesian-favoring equivalent. The qualitative analogs of posterior equivalence and favoring equivalence are obvious, but we state them for full clarity.

Definition 7. E and F are *qualitative favoring equivalent* if E supports H_1 over H_2 at least as much as F and vice versa, for any two disjoint hypotheses H_1 and H_2.

Definition 8. E and F are *qualitative posterior equivalent* if all orderings $\preceq$ that satisfy the axioms (1) the expression $\cdot|E$ is well-defined if and only if $\cdot|E$ is, and (2) $H|E \sim H|F$ for all hypotheses $H \subseteq \Theta$.

As before, these two definitions are equivalent.

Claim 4. E and F are qualitative posterior equivalent if and only if they are qualitative favoring equivalent.

Although posterior and favoring equivalence are easy to make qualitative, what about LP? Quantitatively, the statement of LP mentions both multiplication and a global constant $c > 0$, which seem difficult to translate to the qualitative setting.

Notice, however, that we can multiply probabilities when two events are independent. So suppose E and F are outcomes of the same experiment and

1. $P_\theta(E) = P_\theta(F \cap C_\theta)$ for all θ,
2. For all $\theta \in \Theta$, the events F and C_θ are conditionally independent given θ, and
3. $P_\theta(C_\theta) = P_\upsilon(C_\upsilon) > 0$ for all $\theta, \upsilon \in \Theta$.

Roughly, the event C_θ acts as a witness to the equality $P_\theta(E) = c \cdot P_\theta(F)$. Specifically, assumptions 1 and 2 encode the equality, and assumption 3 asserts this constant is invariant with respect to the parameter θ.

By LP, the three conditions entail that E and F are evidentially equivalent. The proof is simple. Let $c = P_{\theta_0}(C_{\theta_0})$ for any $\theta_0 \in \Theta$. Then for all θ:

$$\begin{aligned} P_\theta(E) &= P_\theta(F \cap C_\theta) \text{ by Assumption 1} \\ &= P_\theta(F) \cdot P_\theta(C_\theta) \text{ by Assumption 2} \\ &= c \cdot P_\theta(F) \text{ by Assumption 3} \end{aligned}$$

Since $P_\theta(E) = c \cdot P_\theta(F)$ for all θ, then E and F are evidentially equivalent by LP.

For an example, suppose we are trying to discern the type of an unmarked urn, with colored balls in frequencies according to the table below. Let B and W respectively denote the events that one draws a blue and a white ball on the first draw. By LP, these events are equivalent, as $P_\theta(B) = 1/2 \cdot P_\theta(W)$ for all $\theta \in \{\theta_1, \theta_2\}$, which denote the urn type. We can see this equivalence in another way. Let C_1/C_2 respectively denote the events that, after replacing the first draw, one draws a cyan/cobalt ball respectively on the second draw. Then, $P_{\theta_1}(W \cap C_1) = P_{\theta_1}(W) \cdot P_{\theta_1}(C_1) = 1/2 \cdot P_{\theta_1}(W) = P_{\theta_1}(B)$ by independence of the draws, if the urn is Type 1. Similarly, $P_{\theta_2}(W \cap C_2) = P_{\theta_2}(W) \cdot P_{\theta_2}(C_2) = 1/2 \cdot P_{\theta_2}(W) = P_{\theta_2}(B)$, if the urn is Type 2. By LP, because $P_{\theta_1}(C_1) = P_{\theta_2}(C_2) > 0$ and $P_{\theta_i}(B) = P_{\theta_i}(W \cap C_i) = P_{\theta_i}(C_i) \cdot P_{\theta_i}(W)$ for all i, we know B and W are evidentially equivalent.

	Number of balls				
	Blue	White	Cyan	Cobalt	Total
Urn Type 1	15	30	50	5	100
Urn Type 2	10	20	20	50	100

So define LP$^-$ to be the thesis that, if conditions 1-3 hold, then E and F are evidentially equivalent. We just showed that, if LP$^-$ entails E and F are evidentially equivalent, then so does LP. By Claim 3, it follows that E and F are also Bayesian posterior and favoring equivalent. Our second major result is the qualitative analog of that fact.

Because conditions 1-3 do not contain any arithmetic operations, each has a direct qualitative analog. The only slightly tricky condition is the second. To define a qualitative analog of conditional independence, note that two events A and B are conditionally independent given C if and only if $P(A|B \cap C) = P(A|C)$ or $P(B \cap C) = 0$. Analogously, we will say events A and B are qualitatively conditionally independent given C if $A|B \cap C \sim A|C$ or $B \cap C \in \mathcal{N}$. In this case, we write $A \perp\!\!\!\perp_C B$.[2]

Our second main result is the following:

Theorem 2. Let $\{C_\theta\}_{\theta \in \Theta}$ be events such that

1. $E|\theta \equiv F \cap C_\theta|\theta$ for all $\theta \in \Theta$,

[2] The lemmas below show this definition of (qualitative) independence has the desired properties. For example, Lemma 8 entails $A \perp\!\!\!\perp_C B$ if and only if $B \perp\!\!\!\perp_C A$.

2. $F \perp\!\!\!\perp_\theta C_\theta$ for all $\theta \in \Theta$, and

3. $\emptyset|\theta \sqsubset C_\theta|\theta \equiv C_\eta|\eta$ for all $\theta, \eta \in \Theta$.

If Θ is finite, then E and F are qualitatively posterior and favoring equivalent.

The following corollary of Theorem 2 is the qualitative analog of the fact that when $P_\theta(E) = P_\theta(F)$ (i.e., when the constant c in LP equals one), then E and F are posterior and favoring equivalent. In the quantitative case, the corollary is trivial because one can let $C_\theta = \Delta$ for all θ. The three conditions of Theorem 2 are satisfied then since (1) $P_\theta(E) = P_\theta(F) = P_\theta(F \cap \Delta)$ for all θ by assumption, (2) F and Δ are conditionally independent given every θ, as $P_\theta(F \cap \Delta) = P_\theta(F) = P_\theta(F) \cdot 1 = P_\theta(F) \cdot P_\theta(\Delta)$, and (3) $P_\theta(\Delta) = 1 > 0$ for all $\theta \in \Theta$. Analogous reasoning works in the qualitative case using the axioms below.

Corollary 1. If Θ is finite and $E|\theta \equiv F|\theta$, then E and F are qualitatively posterior and favoring equivalent.

3.3 Axioms for Qualitative Probability

We assume that both $\sqsubseteq$ and $\preceq$ satisfy the first set of axioms below. To state the axioms, therefore, we let $\underline{\blacktriangleleft}$ be either $\sqsubseteq$ or $\preceq$, and let $\blacktriangleleft$ be the corresponding "strict" inequality defined by $x \blacktriangleleft y$ if $x \underline{\blacktriangleleft} y$ and $y \not\underline{\blacktriangleleft} x$. Define $A|B \overset{\blacktriangle}{=} C|D$ if and only if $A|B \underline{\blacktriangleleft} C|D$ and vice versa.

> Axiom 1: $\underline{\blacktriangleleft}$ is a weak order (i.e., it is linear/total, reflexive, and transitive).
>
> Axiom 3: $A|A \overset{\blacktriangle}{=} B|B$ and $A|B \underline{\blacktriangleleft} \Delta|C$ for all $C \notin \mathcal{N}$.
>
> Axiom 4: $A \cap B|B \overset{\blacktriangle}{=} A|B$.
>
> Axiom 5: Suppose $A \cap B = A' \cap B' = \emptyset$. If $A|C \underline{\blacktriangleleft} A'|C'$ and $B|C \underline{\blacktriangleleft} B'|C'$, then $A \cup B|C \underline{\blacktriangleleft} A' \cup B'|C'$; moreover, if either hypothesis is $\blacktriangleleft$, then the conclusion is $\blacktriangleleft$.
>
> Axiom 6: Suppose $C \subseteq B \subseteq A$ and $C' \subseteq B' \subseteq A'$.
>
> > Axiom 6a: If $B|A \underline{\blacktriangleleft} C'|B'$ and $C|B \underline{\blacktriangleleft} B'|A'$, then $C|A \underline{\blacktriangleleft} C'|A'$; moreover, if either hypothesis is $\blacktriangleleft$, the conclusion is $\blacktriangleleft$.
> >
> > Axiom 6b: If $B|A \underline{\blacktriangleleft} B'|A'$ and $C|B \underline{\blacktriangleleft} C'|B'$, then $C|A \underline{\blacktriangleleft} C'|A'$; moreover, if either hypothesis is $\blacktriangleleft$ and $C \notin \mathcal{N}$, the conclusion is $\blacktriangleleft$.

We discuss Axiom 2 below, as it is different for $\sqsubseteq$ and $\preceq$. The above axioms are due to (Krantz et al., 2006b, p. 222) and enumerated in the same order as in that text. What we call Axiom 6b is what they call Axiom 6' (p. 227). For several reasons (e.g., there is no Archimedean condition), our axioms are not sufficient for $\preceq$ to be representable as either a (conditional) probability measure or a set of probability measures. Thus the axioms represent weaker "coherence" constraints on belief than are typically assumed by Bayesian decision theorists. See Alon and Lehrer (2014) for a recent representation theorem for sets of probability measures, and see (Kraft et al., 1959) and (Krantz et al., 2006a, p. 205-206) for examples illustrating that the axioms are not sufficient for representability as a conditional probability measure.

Axioms 1, 3, 4, and 5, are fairly analogous to facts of quantitative probability. Axiom 6 is useful because it allows us to "multiply" in a qualitative setting. To see its motivation, consider Axiom 6a and note that if $C \subseteq B \subseteq A$ and $C' \subseteq B' \subseteq A'$, then

$$P(C|B) = \frac{P(C)}{P(B)} \text{ and } P(B|A) = \frac{P(B)}{P(A)},$$

and similarly for the A', B', and C'. So if $P(B|A) \geq P(C'|B')$ and $P(C|B) \geq P(B'|A')$, then

$$\frac{P(B)}{P(A)} \geq \frac{P(C')}{P(B')} \text{ and } \frac{P(C)}{P(B)} \geq \frac{P(B')}{P(A')}.$$

When we multiply the left and right-hand sides of those inequalties, we obtain $P(C)/P(A) \geq P(C')/P(A')$, which is equivalent to $P(C|A) \geq P(C'|A')$ given our assumption about the nesting of the sets. Axiom 6b can be motivated similarly.

In addition to these axioms, in statistical contexts, one typically assumes that experimenters agree upon the likelihood functions of the data, which means that, in discrete contexts, $Q^{\mathbb{E}}(\cdot|\theta) = P^{\mathbb{E}}_{\theta}(\cdot)$ whenever $Q(\theta) > 0$. Similarly, we assume that $\preceq$ extends $\sqsubseteq$ in the following sense:

> Axiom 0: If $B, D \notin \mathcal{N}_{\preceq}$ and $B, D \in \Theta \times \mathcal{P}(\Omega)$, then $A|B \preceq C|D$ if and only if $A|B \sqsubseteq C|D$.

We now return to Axiom 2, which concerns probability zero events. For any given fixed θ, the conditional probability $P_\theta(\cdot|E)$ is undefined if and only if $P_\theta(E) = 0$. Similarly, we define $\mathcal{N}_{\sqsubseteq}$ to be all and only sets of that form $\{\theta\} \times E$ such that $E|\theta \sqsubseteq \emptyset|\theta$. We call such events $\sqsubseteq$-*null*. Notice that an experimenter may have a prior that assigns a parameter $\theta \in \Theta$ zero probability, even if $P_\theta(E) > 0$ for all events E. So the set of null events, in the experimenter's joint distribution, will contain all of the null events for each P_θ plus many others. Accordingly, we assume the following about $\preceq$-null events.

> Axiom 2 (for $\sqsubseteq$): $\theta \times \Omega \notin \mathcal{N}_{\sqsubseteq}$ for all θ, and $A \in \mathcal{N}_{\sqsubseteq}$ if and only if $A = \theta \times E$ and $E|\theta \sqsubseteq \emptyset|\theta$.

> Axiom 2 (for $\preceq$): $\Delta \notin \mathcal{N}$, and $A \in \mathcal{N}$ if and only if $A|\Delta \preceq \emptyset|\Delta$.

We say an expression of the form $\cdot|A$ is *undefined* with respect to $\preceq$ / $\sqsubseteq$ if A is null with respect to the appropriate relation. Because it is typically clear from context, we do not specify with respect to which ordering an expression is undefined.

4 Comparing Experiments: Stopping Rules and Mixtures

One of the most controversial consequences of LP is that it entails the "irrelevance of stopping rules" (e.g., see Savage's contribution to Savage et al. (1962)). However, the qualitative version of LP we have stated, one might argue, has no such implication. Why? The careful reader will have noticed that the relations $\sqsubseteq$ and $\preceq$ are always implicitly indexed to a fixed experiment. For example, we defined $\preceq$ to be an ordering on $\mathcal{P}(\Delta) \times \mathcal{P}(\Delta)$, where $\Delta = \Theta \times \Omega^{\mathbb{E}}$ for some fixed experiment $\mathbb{E}$. Thus, on first glance, our qualitative "analog" of LP does not allow one to compare the likelihoods of results drawn from different experiments, and such a comparison is precisely what is at issue in debates about stopping rules.

In this section, we argue that our results do not have this limitation, at least for an important class of "non-informative" stopping rules (Raiffa and Schlaifer, 1961). To do so, let us first clarify what it means for a stopping rule to be "irrelevant." From the robust Bayesian perspective, the choice between two stopping rules $s_{\mathbb{E}}$ and $s_{\mathbb{F}}$ is typically called "irrelevant" if, whenever an outcome ω can be obtained from two experiments $\mathbb{E}$ and $\mathbb{F}$ that *differ only* in their respective stopping rules $s_{\mathbb{E}}$ and $s_{\mathbb{F}}$ respectively, it follows that $Q^{\mathbb{E}}(H|\omega) = Q^{\mathbb{F}}(H|\omega)$ for any hypothesis $H \subseteq \Theta$ and any prior Q. In other words, the stopping rule is irrelevant if, for any outcome ω that can be obtained in both experiments, observing ω in $\mathbb{E}$ is posterior equivalent to observing ω in $\mathbb{F}$. Below, we say what it means for two experiments to "differ only" in their stopping rules.

To generalize this idea to the qualitative setting, we characterize irrelevance in a second, equivalent way. Imagine a *mixed experiment* is conducted in which a fair coin is flipped to decide whether to conduct experiment $\mathbb{E}$ or $\mathbb{F}$. Let $\langle \mathbb{E}, \omega \rangle$ denote the outcome of $\mathbb{M}$ in which the coin lands heads, $\mathbb{E}$ is conducted, and ω is observed. Clearly the following two conditions are equivalent:

1. $Q^{\mathbb{E}}(H|\omega) = Q^{\mathbb{F}}(H|\omega)$ for any prior Q and hypothesis H.
2. $Q^{\mathbb{M}}(H|\langle \mathbb{E}, \omega \rangle) = Q^{\mathbb{M}}(H|\langle \mathbb{F}, \omega \rangle)$ for any prior Q and hypothesis H.

Notice the second condition requires only comparing posterior probabilities of outcomes of a single experiment, $\mathbb{M}$. Therefore, we use the second condition in our definition of *irrelevance* of stopping rules, as it generalizes to the qualitative setting in which one cannot directly compare posterior probabilities from different experiments.

Definition 9. Given two experiments $\mathbb{E}$ and $\mathbb{F}$ that differ only in stopping rule, we say knowledge of the stopping rule is $\mathbb{E}/\mathbb{F}$-*irrelevant* if, for any ω that can be obtained in both experiments, the outcomes $\langle \mathbb{E}, \omega \rangle$ and $\langle \mathbb{F}, \omega \rangle$ of $\mathbb{M}$ are posterior equivalent, where $\mathbb{M}$ is the mixed experiment in which a fair coin is flipped to decide whether to conduct $\mathbb{E}$ or $\mathbb{F}$.

Below, we argue that the choice between certain pairs of stopping rules is irrelevant in this sense. Our argument is a straightforward application of

Corollary 1. We begin by identifying the stopping rules we will focus on and what it means for two experiments to "differ only" in stopping rule.

4.1 Non-Informative Stopping Rules

Often, the data-generating process in an experiment $\mathbb{E}$ is decomposable into two parts (Raiffa and Schlaifer, 1961, §2.3). The first part

$$h^{\mathbb{E}}(\vec{x};\theta_1) = f^{\mathbb{E}}(x_1;\theta_1)\cdot f^{\mathbb{E}}(x_2|x_1;\theta_1)\cdots f^{\mathbb{E}}(x_n|x_1,\ldots x_{n-1};\theta_1)$$

determines the chances that a measurement device takes a sequence of values $\vec{x} = \langle x_1,\ldots x_n\rangle$. Those probabilities are a function exclusively of the parameter of interest θ_1. The second component $\phi(n;\theta_1,\theta_2)$, which is called the *stopping rule*, determines the probability that an additional measurement is made at all, and its values might be affected by innumerable factors in addition to the value of the parameter of interest (e.g., budget, time, etc.); we can represent those factors by the nuisance parameter θ_2. The probability of stopping after n many points is

$$s^{\mathbb{E}}(n|\vec{x};\theta_1,\theta_2) = \phi(x_1;\theta_1,\theta_2)\cdots\phi(n|x_1,\ldots,x_{n-1};\theta_1,\theta_2)\cdot(1-\phi(n+1|\vec{x};\theta_1,\theta_2)),$$

and so letting $\theta = \langle\theta_1,\theta_2\rangle$, one can factor the likelihood function as follows:

$$P^{\mathbb{E}}_{\theta}(\vec{x}) = h^{\mathbb{E}}(\vec{x};\theta_1)\cdot s^{\mathbb{E}}(n|\vec{x};\theta_1,\theta_2)$$

A very special - but still controversial - case is when $s^{\mathbb{E}}$ depends on neither parameter and is either zero or or one, depending on $\vec{x}$. Such stopping rules form a special subset of what Raiffa and Schlaifer (1961) call "non-informative" stopping rules, and they can easily be seen to be irrelevant in the sense above. Why? Suppose two experiments $\mathbb{E}$ and $\mathbb{F}$ have the same data-generating mechanism $h(\vec{x};\theta)$; this is what it means for the two experiments to "differ only" in stopping rule. Further, suppose both $s^{\mathbb{E}}$ and $s^{\mathbb{F}}$ depend on neither parameter, and assume that $s^{\mathbb{E}}(\vec{x})$ and $s^{\mathbb{E}}(\vec{y})$ are always either zero or or one. If a sample $\vec{x}$ can be obtained with positive probability in both experiments, then

$$P^{\mathbb{E}}_{\theta}(\vec{x}) = h(\vec{x};\theta)\cdot s^{\mathbb{E}}(n|\vec{x}) = h(\vec{x};\theta) = h(\vec{x};\theta)\cdot s^{\mathbb{F}}(n|\vec{x}) = P^{\mathbb{F}}_{\theta}(\vec{x}).$$

Thus, the likelihood function of $\vec{x}$ is the *exactly the same* (not just proportional!) in both experiments, and so observing $\vec{x}$ in $\mathbb{E}$ is posterior equivalent to observing $\vec{x}$ in $\mathbb{F}$, as claimed. Notice that a likelihoodist who rejects a fully Bayesian viewpoint could also argue that the choice between such stopping rules is irrelevant because, by the likelihood principle, outcomes obtained in experiments that differ only by such stopping rules are evidentially equivalent.

Example 3: Suppose you are interested whether the fraction θ of people who survive three months of a new cancer treatment exceeds the survival rate of the conventional treatment, which is known to be 94%. Two experimental designs are considered: $\mathbb{E}$, which consists in treating 100 patients, and $\mathbb{F}$, which consists in treating patients until two deaths are recorded. In both experiments, it is possible to treat 100 patients and record precisely two deaths, the second of which is the 100th patient. Let $\vec{x}$ be a binary sequence of length 100 representing such a data set, and note that $P^{\mathbb{E}}_{\theta}(\vec{x}) = P^{\mathbb{F}}_{\theta}(\vec{x}) = \theta^2 \cdot (1-\theta)^{98}$ no matter the survival rate θ. Here, the stopping rules of the respective experiments are given by:

$$\phi^{\mathbb{E}}(n|x_1, \ldots x_{n-1}; \theta) = \begin{cases} 1 \text{ if } n < 100 \\ 0 \text{ otherwise} \end{cases}$$

and

$$\phi^{\mathbb{F}}(n|x_1, \ldots x_{n-1}; \theta) = \begin{cases} 1 \text{ if } \sum_{k \leq n-1} x_k < 2 \\ 0 \text{ otherwise} \end{cases}$$

By the arguments above, the choice between the two stopping rules are irrelevant from both a robust Bayesian and likelihoodist perspective. However, a quick calculation shows that under the null hypothesis – that the new drug is no more effective than the conventional treatment – the chance of two or fewer deaths in $\mathbb{E}$ is .0566, which is not significant at the .05 level. In contrast, in $\mathbb{F}$, the chance of treating 100 or more patients under the null is .014, which is significant at the .05 level. Advocates of classical testing sometimes conclude, therefore, that the latter experiment provides better evidence against the null than the former. This indicates a way in which LP is at odds with some classical testing procedures.

4.2 Qualitative Irrelevance of Stopping Rules

We now argue that, in our qualitative framework, the special class of stopping rules discussed in the previous section are irrelevant. Of course, it is unclear what the analog of "factoring" likelihood functions is in our qualitative framework, and so one might wonder what exactly a qualitative stopping rule is. That question can be avoided entirely. Notice that our argument above showed that, for the special class of stopping rules identified, if $\mathbb{E}$ and $\mathbb{F}$ "differ only" in stopping rule, then $P^{\mathbb{E}}_{\theta}(\omega) = P^{\mathbb{F}}_{\theta}(\omega)$ for all θ and all ω that can be obtained in both experiments. Then, it will also be the case that $P^{\mathbb{M}}_{\theta}(\langle \mathbb{E}, \omega \rangle) = P^{\mathbb{M}}_{\theta}(\langle \mathbb{F}, \omega \rangle)$ for all θ and all ω, where (1) $\mathbb{M}$ is the mixed experiment in which one flips a fair coin to decide whether to conduct $\mathbb{E}$ or $\mathbb{F}$ and (2) the outcomes $\langle \mathbb{E}, \omega \rangle$ and $\langle \mathbb{F}, \omega \rangle$ of $\mathbb{M}$ are as defined above.

Thus, whatever a "stopping rule" is in the qualitative setting, we can define two experiments to *differ only* in stopping rule (for the special class of stopping rules we've identified) if $\langle \mathbb{E}, \omega \rangle | \theta \equiv^{\mathbb{M}} \langle \mathbb{F}, \omega \rangle | \theta$ for all θ and any outcome ω obtainable in both experiments. By Corollary 1, if $\langle \mathbb{E}, \omega \rangle | \theta \equiv^{\mathbb{M}} \langle \mathbb{F}, \omega \rangle | \theta$ for all θ, then $\langle \mathbb{E}, \omega \rangle$ and $\langle \mathbb{F}, \omega \rangle$ are posterior and favoring equivalent. Thus, by the definition of "irrelevance" of stopping rule, knowledge of whether $\mathbb{E}$ or $\mathbb{F}$ was conducted is irrelevant. It should be noted that, just as in the quantitative case, a likelihoodist could also argue for the "irrelevance" of stopping rules by appealing to the qualitative analog of LP$^-$.

5 Conclusions and Future Work

We have shown that, just as LL and LP characterize when all Bayesian reasoners agree how evidence should alter one's posterior probabilities, qualitative likelihoodist theses like QLL characterize agreement among agents whose beliefs satisfy weak coherence requirements. Our work should be extended in at least four ways.

First, one should characterize necessary conditions for qualitative posterior equivalence; we have proven only that our qualitative analog of LP$^-$ is sufficient.

Second, our framework assumes the totality of the relation $\preceq$, but the most plausible generalization of Bayesian reasoning models agents as having degrees of beliefs that are not totally ordered but rather, are representable only by "imprecise" probabilities (Walley, 1991). It is necessary to assess how many of our results continue to hold when $\preceq$ is only a partial order and to see what additional axioms might be appropriate in that case to secure the above results.

Third, we have assumed that the parameter space is finite. Assuming a qualitative countable additivity assumption (Krantz et al., 2006b, p. 216), we conjecture, will allow us to extend those results to countable spaces, but it is unclear how to extend our framework to uncountable spaces.

Finally, Birnbaum (1962) began the exploration of the relationship between LP and a variety of other evidential "axioms", including the sufficiency, conditionality, and invariance principles. The relationship between those additional axioms should be explored in our qualitative framework.

A Appendices

A.1 Notation

Let Θ represent the possible values of an unknown parameter. For simplicity, we assume Θ is finite. An experiment $\mathbb{E}$ will be represented by an ordered pair $\langle \Omega^{\mathbb{E}}, \langle P_\theta^{\mathbb{E}} \rangle_{\theta \in \Theta} \rangle$, where $\Omega^{\mathbb{E}}$ represents the possible outcomes of the experiment and each $P_\theta^{\mathbb{E}}$ is a probability measure over the power set algebra of $\Omega^{\mathbb{E}}$. We assume $\Omega^{\mathbb{E}}$ is finite. The results below hold if the measures $P_\theta^{\mathbb{E}}$ are all defined on some other algebra $\mathcal{A}$ over $\Omega^{\mathbb{E}}$, but for simplicity, we will always consider the power set algebra.

Given an experiment $\mathbb{E}$, we let $\Delta^{\mathbb{E}} = \Theta \times \Omega^{\mathbb{E}}$. We use H_1, H_2 etc. to denote subsets of Θ, and E, F, etc. to denote subsets of Ω. We drop the superscript $\mathbb{E}$ when it is clear from context.

We use $Q^{\mathbb{E}}$ to denote a joint distribution on $\Delta^{\mathbb{E}}$ and will use the letter Q (without a superscript) to denote the associated *prior* on Θ, i.e., $Q(H) = Q^{\mathbb{E}}(H \times \Omega^{\mathbb{E}})$. We assume an experimenter's prior does not depend on the experiment, and so given a prior Q over Θ and an experiment $\mathbb{E} = \langle \Omega^{\mathbb{E}}, \langle P_\theta^{\mathbb{E}} \rangle_{\theta \in \Theta} \rangle$, we can define a joint distribution on $\Delta^{\mathbb{E}}$ by $Q^{\mathbb{E}}(H \times E) := \sum_{\theta \in H, \omega \in E} Q(\theta) P_\theta^{\mathbb{E}}(\omega)$.

We will write $Q^{\mathbb{E}}(\cdot|H)$ and $Q^{\mathbb{E}}(\cdot|E)$ instead of $Q^{\mathbb{E}}(\cdot|H \times \Omega)$ and $Q^{\mathbb{E}}(\cdot|\Theta \times E)$ respectively, and similarly for events to the left of the conditioning bar. When the experiment $\mathbb{E}$ is clear from context, we occasionally drop the superscript $\mathbb{E}$ on the letter Q.

A.2 LL$^+$ is Equivalent to Bayesian Support

The following definition contains the central concept we use to characterize when all Bayesians agree that one piece of evidence is better than another.

Definition 1. Suppose $H_1, H_2 \subseteq \Theta$ are disjoint. Let E and F be outcomes of experiments $\mathbb{E}$ and $\mathbb{F}$ respectively. Say E *Bayesian supports* H_1 over H_2 *at least as much as* F if $Q^{\mathbb{E}}(H_1|E \cap (H_1 \cup H_2)) \geq Q^{\mathbb{F}}(H_1|F \cap (H_1 \cup H_2))$ for all priors Q for which $Q^{\mathbb{F}}(\cdot|F \cap (H_1 \cup H_2))$ is well-defined. If the inequality is strict for all such Q, then we say E *Bayesian supports* H_1 over H_2 *strictly more than* F. In the former case, we write $E \,{}_{H_1}^{\mathcal{B}}\!\trianglerighteq_{H_2} F$, and in the latter, we write $E \,{}_{H_1}^{\mathcal{B}}\!\gg_{H_2} F$.

The following claim links Bayesian support to LL$^+$ by showing that $E \,{}_{H_1}^{\mathcal{B}}\!\trianglerighteq_{H_2} F$ exactly when LL$^+$ says that E provides stronger evidence for θ_1 over θ_2 than F does.

Claim 1. Suppose H_1 and H_2 are finite and disjoint. Let E and F be outcomes of experiments $\mathbb{E}$ and $\mathbb{F}$ respectively. Then $E \,{}_{H_1}^{\mathcal{B}}\!\trianglerighteq_{H_2} F$ if and only if (1) for all $\theta \in H_1 \cup H_2$, if $P_\theta^{\mathbb{F}}(F) > 0$, then $P_\theta^{\mathbb{E}}(E) > 0$, and (2) for all $\theta_1 \in H_1$ and $\theta_2 \in H_2$:

$$P_{\theta_1}^{\mathbb{E}}(E) \cdot P_{\theta_2}^{\mathbb{F}}(F) \geq P_{\theta_2}^{\mathbb{E}}(E) \cdot P_{\theta_1}^{\mathbb{F}}(F) \tag{2}$$

Similarly, $E \,{}_{H_1}^{\mathcal{B}}\!\gg_{H_2} F$ if and only if the inequality in the above equation is always strict.

Proof. In both directions, we will need the following basic arithmetic fact.

Observation 1: Suppose $\langle a_j \rangle_{j \leq n}, \langle b_k \rangle_{k \leq m}, \langle c_j \rangle_{j \leq n}$, and $\langle d_k \rangle_{k \leq m}$ are sequences of non-negative real numbers. Then $a_j \cdot d_k \geq b_k \cdot c_j$ for all $j \leq n$ and $k \leq m$ if and only if

$$\frac{\sum_{j \leq n} r_j \cdot a_j}{\sum_{j \leq n} r_j \cdot a_j + \sum_{k \leq m} s_k \cdot b_k} \geq \frac{\sum_{j \leq n} r_j \cdot c_j}{\sum_{j \leq n} r_j \cdot c_j + \sum_{k \leq m} s_k \cdot d_k} \tag{3}$$

for all sequences $\langle r_j \rangle_{j \leq n}$ and $\langle s_k \rangle_{k \leq m}$ of non-negative real numbers such that $\sum_{j \leq n} r_j a_j + \sum_{k \leq m} s_k b_k$ and $\sum_{j \leq n} r_j c_j + \sum_{k \leq m} s_k d_k$ are both positive.

Proof: First, note that if $\langle r_j \rangle_{j \leq n}$ and $\langle s_k \rangle_{k \leq m}$ are sequences of non-negative real numbers such that $\sum_{j \leq n} r_j a_j + \sum_{k \leq m} s_k b_k$ and $\sum_{j \leq n} r_j c_j + \sum_{k \leq m} s_k d_k$

are positive, then both sides of Equation 3 are well-defined and

$$\begin{aligned}
\text{Equation 3 holds} \quad &\Leftrightarrow \quad \left(\sum_{j\le n} r_j a_j\right)\cdot\left(\sum_{j\le n} r_j c_j + \sum_{k\le m} s_k d_k\right) \ge \left(\sum_{j\le n} r_j c_j\right)\cdot\left(\sum_{j\le n} r_j a_j + \sum_{k\le m} s_k b_k\right)\\
&\qquad \text{Cross multiplying}\\
&\Leftrightarrow \quad \left(\sum_{j\le n} r_j a_j\right)\cdot\left(\sum_{k\le m} s_k d_k\right) \ge \left(\sum_{j\le n} r_j c_j\right)\cdot\left(\sum_{k\le m} s_k b_k\right)\\
&\qquad \text{Cancelling common terms}\\
&\Leftrightarrow \quad \left(\sum_{j\le n, k\le m} r_j a_j s_k d_k\right) \ge \left(\sum_{j\le n, k\le m} r_j c_j s_k b_k\right) \text{[Condition †]}\\
&\qquad \text{Rules of double sums}
\end{aligned}$$

Now we prove the right-to-left direction. Assume Equation 3 holds *for all* sequences of r_js and s_ks for which the sums specified above are positive. Pick arbitrary $j_0 \le n$ and $k_0 \le m$. Our goal is to show $a_{j_0} d_{k_0} \ge c_{j_0} b_{k_0}$. Consider the specific sequence of r_j's and s_k's such that $r_{j_0} = s_{k_0} = 1$ and $r_j = s_k = 0$ for all $j \ne j_0$ and $k \ne k_0$. There are two cases two consider.

If $\sum_{j\le n} r_j a_j + \sum_{k\le m} s_k b_k$ and $\sum_{j\le n} r_j c_j + \sum_{k\le m} s_k d_k$ are both positive, then Condition † holds and reduces to the claim $a_{j_0} d_{k_0} \ge c_{j_0} b_{k_0}$, as desired.

If $\sum_{j\le n} r_j a_j + \sum_{k\le m} s_k b_k = a_{j_0} + b_{k_0}$ and $\sum_{j\le n} r_j c_j + \sum_{k\le m} s_k d_k = c_{j_0} + d_{k_0}$ are not both positive, then because all the a's, b's, etc. are non-negative, it follows that either (i) $a_{j_0} = b_{k_0} = 0$ or (i) $c_{j_0} = d_{k_0} = 0$. In either case $a_{j_0} d_{k_0} = 0 \ge 0 = c_{j_0} b_{k_0}$.

The left-to-right direction merely reverses the direction of the above calculations. In greater detail, assume $a_j \cdot d_k \ge b_k \cdot c_j$ for all $j \le n$ and $k \le m$. Let $\langle r_j \rangle_{j\le n}$ and $\langle s_k \rangle_{k\le m}$ be sequences of non-negative real numbers such that $\sum_{j\le n} r_j a_j + \sum_{k\le m} s_k b_k$ and $\sum_{j\le n} r_j c_j + \sum_{k\le m} s_k d_k$ are both positive. Our goal is to show Equation 3 holds. By the reasoning above, it suffices to show Condition † holds. Note that because $a_j \cdot d_k \ge b_k \cdot c_j$ for all $j \le n$ and $k \le m$ by assumption, every term in the sum on the left-hand-side of Condition † is greater than the corresponding term on the right hand side. So † holds as desired.

(End Proof of Observation).

The central claim follows from Observation 1 by appropriate substitutions. Specifically, enumerate $H_1 = \{\theta_{1,1}, \theta_{1,2}, \dots \theta_{1,n}\}$ and $H_2 = \{\theta_{2,1}, \theta_{2,2}, \dots \theta_{2,m}\}$. To apply the observation, let $a_j = P^{\mathbb{E}}_{\theta_{1,j}}(E)$, $c_j = P^{\mathbb{F}}_{\theta_{1,j}}(F)$, $b_k = P^{\mathbb{E}}_{\theta_{2,k}}(E)$, and $d_k = P^{\mathbb{F}}_{\theta_{2,k}}(F)$. Those terms are clearly all non-negative. So Observation 1 entails that

$$P^{\mathbb{E}}_{\theta_{1,j}}(E) \cdot P^{\mathbb{F}}_{\theta_{2,k}}(F) = a_j d_k \ge c_j b_k \ge P^{\mathbb{E}}_{\theta_{2,k}}(E) \cdot P^{\mathbb{F}}_{\theta_{1,j}}(F) \text{ for all } j, k \tag{4}$$

if and only if

$$\frac{\sum_{j\leq n} r_j \cdot P^{\mathbb{E}}_{\theta_{1,j}}(E)}{\sum_{j\leq n} r_j \cdot P^{\mathbb{E}}_{\theta_{1,j}}(E) + \sum_{k\leq m} s_k \cdot P^{\mathbb{E}}_{\theta_{2,k}}(E)} \geq \frac{\sum_{j\leq n} r_j \cdot P^{\mathbb{F}}_{\theta_{1,j}}(F)}{\sum_{j\leq n} r_j \cdot P^{\mathbb{F}}_{\theta_{1,j}}(F) + \sum_{k\leq m} s_k \cdot P^{\mathbb{F}}_{\theta_{2,k}}(F)} \tag{5}$$

for all sequences $\langle r_j \rangle_{j\leq n}$ and $\langle s_k \rangle_{k\leq m}$ of non-negative real numbers such that $\sum_{j\leq n} r_j \cdot P^{\mathbb{E}}_{\theta_{1,j}}(E) + \sum_{k\leq m} s_k \cdot P^{\mathbb{E}}_{\theta_{2,k}}(E)$ and $\sum_{j\leq n} r_j \cdot P^{\mathbb{F}}_{\theta_{1,j}}(F) + \sum_{k\leq m} s_k \cdot P^{\mathbb{F}}_{\theta_{2,k}}(F)$ are both positive.

Now, we prove the right-to-left direction of Claim 1. Suppose (1) for all $\theta \in H_1 \cup H_2$, if $P^{\mathbb{F}}_\theta(F) > 0$, then $P^{\mathbb{E}}_\theta(E) > 0$, and (2) Equation 4 holds. Let Q be some prior such that $Q^{\mathbb{F}}(\cdot|F \cap (H_1 \cup H_2))$ is well-defined. We claim $Q^{\mathbb{E}}(\cdot|E \cap (H_1 \cup H_2))$ is also well-defined. Why? Since $Q^{\mathbb{F}}(\cdot|F \cap (H_1 \cup H_2))$, we have by definition that:

$$Q^{\mathbb{F}}(H_1|F \cap (H_1 \cup H_2)) = \frac{\sum_{j\leq n} Q(\theta_{1,j}) \cdot P_{\theta_{1,j}}(F)}{\sum_{j\leq n} Q(\theta_{1,j}) \cdot P^{\mathbb{F}}_{\theta_{1,j}}(F) + \sum_{k\leq m} Q(\theta_{2,k}) \cdot P^{\mathbb{F}}_{\theta_{2,k}}(F)} \tag{6}$$

That fraction is well-defined only if the denominator is positive, i.e.,

$$\sum_{j\leq n} Q(\theta_{1,j}) \cdot P^{\mathbb{F}}_{\theta_{1,j}}(F) + \sum_{k\leq m} Q(\theta_{2,k}) \cdot P^{\mathbb{F}}_{\theta_{2,k}}(F) > 0 \tag{7}$$

Therefore, there is some $\theta \in H_1 \cup H_2$ such that $Q(\theta) \cdot P^{\mathbb{F}}_\theta(F) > 0$. By assumption (1), it follows that $Q(\theta) \cdot P^{\mathbb{E}}_\theta(E) > 0$. Hence:

$$\sum_{j\leq n} Q(\theta_{1,j}) \cdot P^{\mathbb{E}}_{\theta_{1,j}}(E) + \sum_{k\leq m} Q(\theta_{2,k}) \cdot P^{\mathbb{E}}_{\theta_{2,k}}(E) > 0 \tag{8}$$

from which it follows that $Q^{\mathbb{E}}(\cdot|E \cap (H_1 \cup H_2))$ is also well-defined. It thus remains to be shown that $Q^{\mathbb{E}}(H_1|E \cap (H_1 \cup H_2)) \geq Q^{\mathbb{F}}(H_1|F \cap (H_1 \cup H_2))$.

Letting $r_j = Q(\theta_{1,j})$ and $s_k = Q(\theta_{2,k})$, the inequalities in Equation 7 and Equation 8 entail that $\sum_{j\leq n} r_j a_j + \sum_{k\leq m} s_k b_k$ and $\sum_{j\leq n} r_j c_j + \sum_{k\leq m} s_k d_k$ are both positive. Since we've assumed (2) that Equation 4 holds, the left-to-right direction of Observation 1 entails that Equation 5 must hold as desired.

In the left-to-right direction, suppose that $Q^{\mathbb{E}}(H_1|E \cap (H_1 \cup H_2)) \geq Q^{\mathbb{F}}(H_1|F \cap (H_1 \cup H_2))$ for all priors Q for which $Q^{\mathbb{F}}(H_1|F \cap (H_1 \cup H_2))$ is well-defined. We must show that (1) for all $\theta \in H_1 \cup H_2$, if $P^{\mathbb{F}}_\theta(F) > 0$, then $P^{\mathbb{E}}_\theta(E) > 0$, and (2) Equation 4 holds.

To show (1), suppose $P^{\mathbb{F}}_\theta(F) > 0$ for some $\upsilon \in H_1 \cup H_2$. Define Q so that $Q(\upsilon) = 1$. Then

$$Q^{\mathbb{F}}(F \cap (H_1 \cup H_2)) = \sum_{\theta \in H_1 \cup H_2} Q(\theta) \cdot P^{\mathbb{F}}_\theta(F) = Q(\upsilon) \cdot P^{\mathbb{F}}_\upsilon(F) = P^{\mathbb{F}}_\upsilon(F) > 0 \tag{9}$$

and so $Q^{\mathbb{F}}(\cdot|F \cap (H_1 \cup H_2))$ is well-defined. Thus, $Q^{\mathbb{E}}(H_1|E \cap (H_1 \cup H_2)) \geq Q^{\mathbb{F}}(H_1|F \cap (H_1 \cup H_2))$ by our assumption. Hence, $Q^{\mathbb{E}}(\cdot|E \cap (H_1 \cup H_2))$ is also well-defined, which entails $Q^{\mathbb{E}}(E \cap (H_1 \cup H_2)) > 0$. But using the same reasoning that lead to Equation 9, we obtain $P^{\mathbb{E}}_{\upsilon}(E) = Q^{\mathbb{E}}(E \cap (H_1 \cup H_2))$, and so $P^{\mathbb{E}}_{\upsilon}(E) > 0$ as desired.

Next, we must show that Equation 4 holds. Since $Q^{\mathbb{E}}(H_1|E \cap (H_1 \cup H_2)) \geq Q^{\mathbb{F}}(H_1|F \cap (H_1 \cup H_2))$, we make the same substitutions as in the left-to-right direction and infer that Equation 5 holds. We then apply Observation 1 to obtain that Equation 4 holds, as desired. □

From this claim, we can derive an equivalence between what the notion of "favoring" in the weak law of likelihood and what we call "Bayesian favoring." Recall, the weak law of likelihood (LL) says that $P^{\mathbb{E}}_{\theta_1}(E) \geq P^{\mathbb{E}}_{\theta_2}(E)$ if and only if the data E (weakly) favors θ_1 over θ_2. Bayesian favoring is defined below.

Definition 2. Say E *Bayesian favors* H_1 to H_2 if $E \; {}_{H_1}^{\mathcal{B}}\!\trianglerighteq_{H_2} \Omega$, where Ω is the sure event. E strictly Bayesian favors H_1 to H_2 if $E \; {}_{H_1}^{\mathcal{B}}\!\gg_{H_2} \Omega$.

Claim 2. E Bayesian favors H_1 to H_2 if and only if (1) $P_\theta(E) > 0$ for all $\theta \in H_1 \cup H_2$ and (2) $P_{\theta_1}(E) \geq P_{\theta_2}(E)$ for all $\theta_1 \in H_1$ and $\theta_2 \in H_2$. If $H_1 = \{\theta_1\}$ and $H_2 = \{\theta_2\}$ are simple, E strictly Bayesian favors H_1 to H_2 if and only if LL entails E favors H_1 to H_2.

Proof. This follows immediately from Claim 1, since the likelihood of $P_\theta(\Omega) = 1$ for all θ. □

A.3 LP is Equivalent to Bayesian Posterior and Favoring Equivalence

We now characterize LP in terms of Bayesian agreement.

Definition 3. E and F are *Bayesian favoring equivalent* if $F \; {}_{H_1}^{\mathcal{B}}\!\trianglerighteq_{H_2} E$ and $E \; {}_{H_1}^{\mathcal{B}}\!\trianglerighteq_{H_2} F$ for all disjoint hypotheses H_1 and H_2.

Definition 4. Say E and F are *Bayesian posterior equivalent* if for all priors Q (i) $Q(\cdot|E)$ is well-defined if and only if $Q(\cdot|F)$ is well-defined, and (ii) $Q(H|E) = Q(H|F)$ for all hypotheses H, whenever those conditional probabilities are well-defined.

Claim 3. The following are equivalent: (1) LP entails E and F are evidentially equivalent, (2) E and F are Bayesian posterior equivalent, (3) E and F are Bayesian favoring equivalent.

Proof. First, we prove $1 \Rightarrow 2$. Suppose there is a $c > 0$ such that $P^{\mathbb{E}}_\theta(E) = c \cdot P^{\mathbb{F}}_\theta(F)$. We must show E and F are posterior equivalent. To do so, we must first show $Q^{\mathbb{E}}(\cdot|E)$ is well-defined if and only if $Q^{\mathbb{F}}(\cdot|F)$ is. To do that, we must show $Q^{\mathbb{E}}(E)$ is positive if and only if $Q^{\mathbb{F}}(F)$ is.

By definition of $Q^{\mathbb{E}}$, we have

$$Q^{\mathbb{E}}(E) = \sum_{\theta \in \Theta} Q(\theta) \cdot P^{\mathbb{E}}_\theta(E). \tag{10}$$

Similarly,

$$Q^{\mathbb{F}}(F) = \sum_{\theta \in \Theta} Q(\theta) \cdot P_\theta^{\mathbb{F}}(F) \tag{11}$$

Since $P_\theta^{\mathbb{E}}(E) = c \cdot P_\theta^{\mathbb{F}}(F)$ for some $c > 0$, each term in the sum in Equation 10 is positive if and only if the corresponding term in the sum Equation 11 is positive. Hence, $Q^{\mathbb{E}}(E) > 0$ if and only if $Q^{\mathbb{F}}(F) > 0$, as desired.

The second thing we must show is that $Q^{\mathbb{E}}(H|E) = Q^{\mathbb{F}}(H|F)$ for all H, whenever those conditional probabilities are well-defined. To that end, let $H \subseteq \Theta$ be any (simple or composite) hypothesis. Then:

$$\begin{aligned}
Q^{\mathbb{E}}(H|E) &= \frac{Q^{\mathbb{E}}(E \cap H)}{Q^{\mathbb{E}}(E)} && \text{definition of conditional probability} \\
&= \frac{\sum_{\theta \in H \cap \Theta} Q^{\mathbb{E}}(E \cap \{\theta\})}{\sum_{\theta \in \Theta} Q^{\mathbb{E}}(E \cap \{\theta\})} && \text{law of total probability} \\
&= \frac{\sum_{\theta \in H \cap \Theta} P_\theta^{\mathbb{E}}(E) \cdot Q(\theta)}{\sum_{\theta \in \Theta} P_\theta^{\mathbb{E}}(E) \cdot Q(\theta)} && \text{because } Q(\cdot|\theta) = P_\theta^{\mathbb{E}}(\cdot) \text{ for all } \theta \in \Theta \text{ and } Q(\theta) \text{ does not depend upon } \mathbb{E} \\
&= \frac{\sum_{\theta \in H \cap \Theta} c \cdot P_\theta^{\mathbb{F}}(F) \cdot Q(\theta)}{\sum_{\theta \in \Theta} c \cdot P_\theta^{\mathbb{F}}(F) \cdot Q(\theta)} && \text{as } P_\theta^{\mathbb{E}}(E) = c \cdot P_\theta^{\mathbb{F}}(F) \\
&= \frac{\sum_{\theta \in H \cap \Theta} P_\theta^{\mathbb{F}}(F) \cdot Q(\theta)}{\sum_{\theta \in \Theta} P_\theta^{\mathbb{F}}(F) \cdot Q(\theta)} && \text{canceling } c \\
&= Q^{\mathbb{F}}(H|F) && \text{reversing the above steps}
\end{aligned}$$

Next, we prove 2 $\Rightarrow$ 1. Assume that, for all Q, we have (i) $Q^{\mathbb{E}}(\cdot|E)$ is well-defined if and only if $Q^{\mathbb{F}}(\cdot|F)$ is and (ii) $Q^{\mathbb{E}}(H|E) = Q^{\mathbb{F}}(H|F)$ for all non-empty hypotheses $H \subseteq \Theta$ and all Q for which those conditional probabilities are well-defined. We must find some $c > 0$ such that $P_\theta^{\mathbb{E}}(E) = c \cdot P_\theta^{\mathbb{F}}(F)$ for all θ.

We first assume there is some $\theta \in \Theta$ such that $P_\theta^{\mathbb{E}}(E) > 0$ or $P_\theta^{\mathbb{F}}(F) > 0$. Otherwise, $P_\theta^{\mathbb{E}}(E) = P_\theta^{\mathbb{F}}(F) = 0$ for all $\theta \in \Theta$, and so $P_\theta^{\mathbb{E}}(E) = c \cdot P_\theta^{\mathbb{F}}(F)$ for all $c > 0$ trivially.

So pick any $\upsilon \in \Theta$ such that either $P_\upsilon^{\mathbb{E}}(E) > 0$ or $P_\upsilon^{\mathbb{F}}(F) > 0$. Without loss of generality, assume that $P_\upsilon^{\mathbb{F}}(F) > 0$. We claim $P_\upsilon^{\mathbb{E}}(E) > 0$ as well. Why? Let Q be the prior such that $Q(\upsilon) = 1$. Then $Q^{\mathbb{F}}(F) = Q(\upsilon) \cdot P_\upsilon^{\mathbb{F}}(F) = P_\upsilon^{\mathbb{F}}(F) > 0$. Hence, $Q^{\mathbb{F}}(\cdot|F)$ is well-defined. Since E and F are posterior equivalent, $Q^{\mathbb{E}}(\cdot|E)$ is also well-defined. Because $Q^{\mathbb{F}}(E) = Q(\upsilon) \cdot P_\upsilon^{\mathbb{E}}(E) = P_\upsilon^{\mathbb{E}}(E)$, it follows that $P_\upsilon^{\mathbb{E}}(E) > 0$ as desired.

Thus, if we let $c = P_\upsilon^{\mathbb{E}}(E)/P_\upsilon^{\mathbb{F}}(F)$, it follows that $c > 0$. We want to show that $P_\theta^{\mathbb{E}}(E) = c \cdot P_\theta^{\mathbb{F}}(F)$ for all θ. It suffices, therefore, to show $P_\upsilon^{\mathbb{E}}(E)/P_\upsilon^{\mathbb{F}}(F) = P_\theta^{\mathbb{E}}(E)/P_\theta^{\mathbb{F}}(F)$ for all θ.

To do so, let θ be arbitrary and now define Q to be any prior such that $Q(\upsilon)$ and $Q(\theta)$ are both positive. Again, it's easy to check that $Q^{\mathbb{E}}(\cdot|E)$ and $Q^{\mathbb{F}}(\cdot|F)$

are well-defined. So Bayes rule entails:

$$\frac{Q^{\mathbb{E}}(\theta|E)}{Q^{\mathbb{E}}(\upsilon|E)} = \frac{P^{\mathbb{E}}_{\theta}(E)}{P^{\mathbb{E}}_{\upsilon}(E)} \cdot \frac{Q(\theta)}{Q(\upsilon)}$$

And similarly for F. By assumption, $Q^{\mathbb{E}}(\{\theta\}|E) = Q^{\mathbb{E}}(\{\theta\}|F)$, and hence:

$$\frac{P^{\mathbb{E}}_{\theta}(E)}{P^{\mathbb{E}}_{\upsilon}(E)} \cdot \frac{Q(\theta)}{Q(\upsilon)} = \frac{P^{\mathbb{F}}_{\theta}(F)}{P^{\mathbb{F}}_{\upsilon}(F)} \cdot \frac{Q(\theta)}{Q(\upsilon)}$$

It immediately follows that $P^{\mathbb{E}}_{\upsilon}(E)/P^{\mathbb{F}}_{\upsilon}(F) = P^{\mathbb{E}}_{\theta}(E)/P^{\mathbb{F}}_{\theta}(F)$ as desired.

Next we prove $2 \Rightarrow 3$. Suppose that E and F are posterior equivalent. We must show that $Q^{\mathbb{E}}(H_1|E \cap (H_1 \cup H_2)) = Q^{\mathbb{F}}(H_1|F \cap (H_1 \cup H_2))$ for all disjoint hypotheses $H_1, H_2 \subseteq \Theta$ and all Q for which those conditional probabilities are well-defined.

To do so, let H_1 and H_2 be disjoint hypotheses, and let Q be a prior such that $Q^{\mathbb{E}}(\cdot|E \cap (H_1 \cup H_2))$ and $Q^{\mathbb{F}}(\cdot|F \cap (H_1 \cup H_2))$ are both well-defined. It follows that $Q^{\mathbb{E}}(E)$ and $Q^{\mathbb{F}}(F)$ are both positive (since $Q^{\mathbb{E}}(E \cap (H_1 \cup H_2))$ and $Q^{\mathbb{F}}(F \cap (H_1 \cup H_2))$ are). Since E and F are posterior equivalent, we can infer that $Q(H_1|E) = Q(H_1|F)$ and $Q(H_2|E) = Q(H_2|F)$. Thus:

$$(Q(H_1|E) \cdot Q(E)) \cdot (Q(H_2|F) \cdot Q(F)) = (Q(H_1|F) \cdot Q(E)) \cdot (Q(H_2|E) \cdot Q(F)) \quad (12)$$

We obtain the desired result via the following sequences of arithmetic operations:

$$\begin{aligned}
\text{Equation 12 holds} \quad & \Leftrightarrow \\
Q(E \cap H_1) \cdot Q(F \cap H_2) \quad & = \quad Q(H_1 \cap F) \cdot Q(E \cap H_2) \\
& \Leftrightarrow \\
Q(E \cap H_1) \cdot Q(F \cap H_2) + Q(E \cap H_1) \cdot Q(F \cap H_1) \quad & = \quad Q(F \cap H_1) \cdot Q(E \cap H_2) + Q(E \cap H_1) \cdot Q(F \cap H_1) \\
& \Leftrightarrow \\
Q(E \cap H_1) \cdot [Q(F \cap H_1) + Q(F \cap H_2)] \quad & = \quad Q(F \cap H_1)\,[Q(E \cap H_1) + Q(E \cap H_2)] \\
& \Leftrightarrow \\
Q(E \cap H_1)/\,[Q(E \cap H_1) + Q(E \cap H_2)] \quad & = \quad Q(F \cap H_1)/\,[Q(F \cap H_1) + Q(F \cap H_2)] \\
& \Leftrightarrow \\
Q(H_1|E \cap (H_1 \cup H_2)) \quad & = \quad Q(H_1|F \cap (H_1 \cup H_2)) \text{ as } H_1 \cap H_2 = \emptyset
\end{aligned}$$

Finally, we prove $3 \Rightarrow 2$. Suppose E and F are Bayesian support equivalent. We must show that E and F are Bayesian posterior equivalent, i.e., that (1) $Q(\cdot|E)$ is well-defined if and only if $Q(\cdot|F)$ is and (2) $Q(H|E) = Q(H|F)$ for all hypotheses H and all Q for which those conditional probabilities are well-defined.

To show (1), pick any H_1 and H_2 such that $H_1 \cup H_2 = \Theta$. If $Q(\cdot|E)$ is well-defined, then $Q(E) > 0$, and note that $Q(E) = Q(E \cap (H_1 \cup H_2))$ because $H_1 \cup H_2 = \Theta$. Since E and F are Bayesian support equivalent, it follows that

$Q(F) = Q(F \cap (H_1 \cup H_2)) \geq Q(E \cap (H_1 \cup H_2)) > 0$, and hence, $Q(\cdot|F)$ is all well-defined. Thus, we've shown that if $Q(\cdot|E)$ is well-defined, so is $Q(\cdot|F)$. The converse is proven in the exact same manner.

To show (2), let Q be any prior for which $Q(\cdot|E)$ and $Q(\cdot|F)$ are well-defined. Let $H \subseteq \Theta$ by some arbitrary hypothesis. We must show $Q(H|E) = Q(H|F)$. Then define $H_1 = H$ and $H_2 = \Theta \setminus H$. Since E and F are Bayesian support equivalent, we know that

$$Q(H_1|E \cap (H_1 \cup H_2)) = Q(H_1|F \cap (H_1 \cup H_2))$$

from which it follows that

$$Q(H|E) = Q(H_1|E) = Q(H_1|E \cap (H_1 \cup H_2)) = Q(H_1|F \cap (H_1 \cup H_2)) = Q(H|F)$$

as desired. □

B Main Results

In this section, we prove generalizations of Claim 2 and Claim 3. Our results apply in contexts in which an experimenter can justifiably order events in terms of probability but may not be able to assign events precise numerical probabilities. To do so, we represent experiment $\mathbb{E}$ by a pair $\langle \Omega^{\mathbb{E}}, \sqsubseteq^{\mathbb{E}} \rangle$ where, as before, $\Omega^{\mathbb{E}}$ represents the set of possible experimental outcomes. The relation $\sqsubseteq^{\mathbb{E}}$ is the qualitative analog of the set of likelihood functions in an experiment. The idea is that $A|\theta \sqsubseteq^{\mathbb{E}} B|\eta$ represents the claim that "experimental outcome B is at least as likely under supposition η as outcome A is under supposition θ." We write $A|B \equiv^{\mathbb{E}} C|D$ if $A|B \sqsubseteq^{\mathbb{E}} C|D$ and vice versa. We drop the superscript $\mathbb{E}$ when it is clear from context.

In standard statistical contexts, one can use the measure P_θ to compare *conditional* probabilities. For instance, if E, F, and G respectively represent the events that at least six, seven, and eight heads are observed in a sequence of ten tosses, then $P_\theta(E|F) = 1 > P_\eta(G|F)$ for any $\theta, \eta \in (0, 1)$. Similarly, we assume that $\sqsubseteq$ is a binary relation on $\mathcal{P}(\Delta) \times ((\Theta \times \mathcal{P}(\Omega)) \setminus \mathcal{N}_{\sqsubseteq})$ (where $\mathcal{P}$ is the power set operation) so that one can make comparisons of the form $A|\theta \cap E \sqsubseteq B|\eta \cap F$, where $E, F \subseteq \Omega$ are observable events, $\theta, \eta \in \Theta$ are possible values of the unknown parameter, and $A, B \subseteq \Delta$ are arbitrary events. Here, members of $\mathcal{N}_{\sqsubseteq}$ are the analog of probability zero events, which we cannot condition on. We call members of $\mathcal{N}_{\sqsubseteq}$ *null sets*.

Similarly, Bayesians assume that an experimenter's beliefs are representable by a probability function. We weaken that assumption and assume beliefs are representable by an ordering $\preceq$ on $\mathcal{P}(\Delta) \times (\mathcal{P}(\Delta) \setminus \mathcal{N}_{\preceq})$. Here, $\mathcal{N}_{\preceq}$ is the set of events the experimenter regards as having zero probability. We write $A|B \sim C|D$ when $A|B \preceq C|D$ and vice versa.

B.1 Important Lemmas

The following lemmas are copied from (Krantz et al., 2006a, p. 229), where their proofs can be found.

Lemma 1. $A|B \succeq \emptyset|\Delta$ whenever $B \notin \mathcal{N}$.

Lemma 2. $\emptyset|A \sim \emptyset|B$.

Lemma 3. If $A \supseteq B$, then $A|C \succeq B|C$.

Lemma 4.

1. $\emptyset \in \mathcal{N}$,
2. If $A \in \mathcal{N}$, then $\Delta \setminus A \notin \mathcal{N}$.
3. If $A \in \mathcal{N}$ and $B \subseteq A$, then $B \in \mathcal{N}$.
4. If $A, B \in \mathcal{N}$, then $A \cup B \in \mathcal{N}$.

In standard probability theory, calls an event "P-almost sure" if $P(A) = 1$. By analogy, define:

Definition 5. An event A is called $\preceq$-*almost sure* if $\Delta \setminus A \in \mathcal{N}_{\preceq}$. Let $\mathcal{S}_{\preceq}$ denote the set of all $\preceq$-almost sure events.

As before, we drop the reference to the ordering $\preceq$ when it's clear from context.

Lemma 5. Let $A \in \mathcal{S}$ be an almost sure event. Then:

1. $A \notin \mathcal{N}$
2. $A \cap B \notin \mathcal{N}$ for any $B \notin \mathcal{N}$.

Proof. For brevity, we let $\neg X$ abbreviate $\Delta \setminus X$ below.

[Part 1:] This is just a restatement of Part 2 of Lemma 4.

[Part 2:] Suppose for the sake of contradiction that $A \cap B \in \mathcal{N}$. By Part 2 of Lemma 4, since $\neg A \cap B \subseteq \neg A$, it follows that $\neg A \cap B \in \mathcal{N}$. Since both $A \cap B$ and $\neg A \cap B$ are members of $\mathcal{N}$, it follows from Part 4 of Lemma 4 that $(A \cap B) \cup (\neg A \cap B) \in \mathcal{N}$. However, $(A \cap B) \cup (\neg A \cap B) = B$, and so it follows that $B \in \mathcal{N}$, contradicting assumption. □

In addition to Lemma 4 and Lemma 5, we will frequently need to refer to additional facts about null and almost-sure events. The first three parts of the following lemma slightly strengthen (Krantz et al., 2006b, p. 230)'s Lemma 9. The remaining parts of the are the lemma are analogous to facts from standard probability theory that characterize when measure zero sets can be ignored, and hence, when one can focus on the almost sure events. More precisely, in standard probability theory, if $P(A) = 1$, then (1) $P(C|B) = P(C \cap A|B)$ and (2) $P(C|B) = P(C|B \cap A)$ whenever $P(B) > 0$. The last parts of the lemma are the analog of those two facts.

We separate the following lemma from the previous ones since its proof uses Axiom 6b. As we have said, (Krantz et al., 2006b, p. 230), prove the first parts of the lemma, but in doing so, they use a "structural" axiom (specifically, what they call Axiom 8) that requires the algebra of events to be sufficiently rich. What we show is that similar facts can be derived from Axiom 6b, which poses no constraint on the algebra of events itself.

Lemma 6.

1. If $A \in \mathcal{N}$, then $A|B \sim \emptyset|\Delta$ for all $B \notin \mathcal{N}$.
2. If $A \in \mathcal{N}$, then $A|B \sim \emptyset|C$ for all $B, C \notin \mathcal{N}$.
3. If $A \subseteq B$ and $A|B \preceq \emptyset|\Delta$, then $A \in \mathcal{N}$.
4. If $A \in \mathcal{S}$, then $A|A \sim A|B$ for all $B \notin \mathcal{N}$.
5. If $A \in \mathcal{S}$, then $C|B \sim (A \cap C)|B$ for all $B \notin \mathcal{N}$. Similarly, if $A \in \mathcal{N}$, then $C|B \sim C \cup A|B$ for all $B \notin \mathcal{N}$.
6. If $A \in \mathcal{S}$, then $C|B \sim C|(B \cap A)$ for all $B \notin \mathcal{N}$. Similarly, if $A \in \mathcal{N}$, then $C|B \sim C|B \cup A$ for all $B \notin \mathcal{N}$.

Proof. As before, for brevity, we let $\neg X$ abbreviate $\Delta \setminus X$ below.

[Part 1:] Suppose for the sake of contradiction that $A|B \not\sim \emptyset|\Delta$. By Lemma 1, $A|B \succeq \emptyset|\Delta$, and thus, it follows from our supposition that $A|B \succ \emptyset|\Delta$. By Axiom 4, we obtain that $A \cap B|B \sim A|B$, and so by Axiom 1 (specifically, transitivity), we get $A \cap B|B \succ \emptyset|\Delta$. By Lemma 2, $\emptyset|B \sim \emptyset|\Delta$, and so by Axiom 1 (specifically, transitivity), we get $A \cap B|B \succ \emptyset|B$. Now define:

$$\begin{aligned} X &= \Delta & X' &= \Delta \\ Y &= B & Y' &= B \\ Z &= A \cap B & Z' &= \emptyset \end{aligned}$$

Clearly, $Z \subseteq Y \subseteq X$ and $Z' \subseteq Y' \subseteq X'$. We've just shown that $Z|Y \succ Z'|Y'$, and since $Y|X = Y'|X'$, we have by Axiom 1 (specifically, reflexivity) that $Y|X \succeq Y'|X'$. Thus, Axiom 6b entails $Z|X \succ Z'|X'$, i.e., that $A|\Delta \succ \emptyset|\Delta$. But because $A \in \mathcal{N}$, it follows from Axiom 2 that $A|\Delta \preceq \emptyset|\Delta$, which is a contradiction.

[Part 2:] Follows immediately from Part 1 and Lemma 2.

[Part 3:] Define:

$$\begin{aligned} X &= \Delta & X' &= \Delta \\ Y &= B & Y' &= \Delta \\ Z &= A & Z' &= \emptyset \end{aligned}$$

Clearly, $Z' \subseteq Y' \subseteq X'$. Because $A \subseteq B$, we also know $Z \subseteq Y \subseteq X$. By assumption

$$Z|Y = A|B \preceq \emptyset|\Delta = Z'|Y'.$$

Further, since $B \subseteq \Delta$, Lemma 3 entails $Y|X = B|\Delta \preceq \Delta|\Delta = Y'|X'$. Thus, Axiom 6b entails $Z|Y \preceq Z'|X'$, i.e., that $A|\Delta \preceq \emptyset|\Delta$. So by Axiom 2, it follows that $A \in \mathcal{N}$, as desired.

[Part 4:]. Suppose for the sake of contradiction that $A|B \not\sim B|B$. Since $B \notin \mathcal{N}$, it follows from Axiom 1 (specifically, the totality of $\preceq$) that either $A|B \prec B|B$ or $A|B \succ B|B$. By Axiom 3 and Axiom 4, we have $A|B \preceq \Delta|B \sim B|B$. Thus, $A|B \prec B|B$. Since $A \in \mathcal{S}$, by definition of $\mathcal{S}$, we know $\neg A \in \mathcal{N}$. Hence, $\neg A|B \preceq \emptyset|B$ by Part 2 of this lemma. Since (1) $A|B \prec B|B$ and (2) $\neg A|B \preceq \emptyset|B$, it follows from Axiom 5 that $A \cup \neg A|B \prec B|B$. But $A \cup \neg A = \Delta$, and so we've shown $\Delta|B \prec B|B$. Since $\Delta|B \sim B|B$ by Axiom 4, it follows that $B|B \prec B|B$, contradicting the Axiom 1 (specifically, the reflexivity of $\preceq$).

[Part 5:] Since $A \in \mathcal{S}$, we have $\neg A \in \mathcal{N}$, and thus, by Part 3 of Lemma 4, we have $\neg A \cap C \in \mathcal{N}$. Thus, $\neg A \cap C|B \sim \emptyset|B$ by Part 2 of this Lemma. Since (1) $A \cap C|B \sim A \cap C|B$ by Axiom 1 (specifically, the reflexivity of $\preceq$) and (2) $\neg A \cap C|B \sim \emptyset|B$ (as just shown), Axiom 5 entails:

$$C|B = (A \cap C) \cup (\neg A \cap C)|B \sim (A \cap C) \cup \emptyset|B = A \cap C|B$$

as desired.

To show the second claim, suppose $A \in \mathcal{N}$ so that $\neg A \in \mathcal{S}$ by definition. Because $\neg A \in \mathcal{S}$, what we've just shown entails both that $C|B \sim C \cap \neg A|B$ and that $(C \cup A)|B \sim (C \cup A) \cap \neg A|B$. Notice that $C \cap \neg A = (C \cup A) \cap \neg A$, and so by transitivity we get $C|B \sim C \cup A|B$ as desired.

[Part 6:] Suppose for the sake of contradiction that $C|B \not\sim C|(B \cap A)$. By assumption $B \notin \mathcal{N}$, and so the left-hand-side is well-defined. The right-hand side is also well-defined. Why? By assumption $A \in \mathcal{S}$ and $B \notin \mathcal{N}$, and so $A \cap B \notin \mathcal{N}$ by Lemma 5.

Since $C|B \not\sim C|(B \cap A)$ and both sides are well-defined, by Axiom 1, either $C|B \succ C|(B \cap A)$ or vice versa. Without loss of generality, assume the former. Define:

$$\begin{aligned} X &= \Delta & X' &= \Delta \\ Y &= B & Y' &= B \cap A \\ Z &= C \cap B & Z' &= C \cap B \cap A \end{aligned}$$

By Axiom 4 and assumption, $Z|Y \succ Z'|Y'$. Because $A \in \mathcal{S}$, Part 5 of this lemma entails $Y|X = B|\Delta \sim Y'|X' = (B \cap A)|\Delta$. So by Axiom 6b, it follows that

$$Z|X \succ Z'|X' = (C \cap B \cap A)|\Delta.$$

But again, by Part 4 of this lemma, $Z'|X' = (C \cap B \cap A)|\Delta \sim (C \cap B)|\Delta = Z|X$. So we've shown $Z|X \succ Z'|X'$ and $Z|X \sim Z'|X'$, which is a contradiction.

The second claim is proven in the same way as the second claim in Part 5. $\square$

The following lemma is the analog of the fact that, if $P(C) > 0$, then $P(A|C) \leq P(B|C)$ if and only if $P(A \cap C) \leq P(B \cap C)$. For brevity, we sometimes write X instead of $X|\Delta$ below.

Lemma 7. If $X|Y \preceq W|Y$, then $X \cap Y \preceq W \cap Y$. The converse holds if $Y \notin \mathcal{N}$.

Proof. To prove the left to right direction, we apply Axiom 6b. Let $A = A' = \Delta$; $B = B' = Y$; $C = X \cap Y$ and $C' = W \cap Y$. Notice that $C \subseteq B \subseteq A$ and $C' \subseteq B' \subseteq A'$. Because $A = A'$ and $B = B'$, it immediately follows from the reflexivity of $\preceq$ that $B|A \preceq B'|A'$. Further:

$$\begin{aligned} C|B &= X \cap Y|Y \text{ by definition of } C\&B \\ &\sim X|Y \text{ by Axiom 4} \\ &\preceq W|Y \text{ by assumption} \\ &\sim W \cap Y|Y \text{ by Axiom 4} \\ &= C'|B' \text{ by definition of } C' \text{ and } B' \end{aligned}$$

So by Axiom 6b, it follows that $C|A \preceq C'|A'$, i.e., that $X \cap Y|\Delta \preceq W \cap Y|\Delta$, i.e., that $X \cap Y \preceq W \cap Y$, as desired.

In the right to left direction, we prove the contrapositive. That is, suppose that $X|Y \npreceq W|Y$. Because $Y \notin \mathcal{N}$ and the relation $\preceq$ is total (Axiom 1), it follows that $W|Y \prec X|Y$. From this, we immediately get that $X \cap Y \notin \mathcal{N}$. One can then apply Axiom 6b in the same way we just did to show that $W \cap Y \prec X \cap Y$ (since $X \cap Y \notin \mathcal{N}$, we get the strict conclusion). By the definition of the $\prec$, it follows that $X \cap Y \npreceq W \cap Y$, as desired. $\square$

The final important lemma is special case of Axiom 6a that we use frequently; it's proof relies on Axiom 6b in cases in which null sets are involved.

Lemma 8. Suppose $A \supseteq B \supseteq C$ and $A \supseteq B' \supseteq C$. If $B|A \succeq C|B'$ and $B \notin \mathcal{N}$, then $B'|A \succeq C|B$. Further, if the antecedent is $\succ$, the consequent is $\succ$.

Proof. Assume $A \supseteq B \supseteq C$, $A \supseteq B' \supseteq C$, and $B|A \succeq C|B'$.

First note that if $C \in \mathcal{N}$, then it follows from Lemma 6 that $C|B \sim \emptyset|\Delta$. Hence, by Lemma 1, we get $B'|A \succeq C|B$. In this case, we also get the stronger result that $B'|A \succ C|B$ without any further assumption. Why? Since by hypothesis $B|A \succeq C|B'$, the expression $C|B'$ is defined and $B' \notin \mathcal{N}$. Since $B' \subseteq A$, it follows from Part (3) of Lemma 6 that $B'|A \npreceq \emptyset|\Delta$. By Axiom 1 (specifically, totality), we obtain that $B'|A \succ \emptyset|\Delta \sim C|B$ as desired.

So suppose $C \notin \mathcal{N}$. Suppose, for the sake of contradiction, that $B'|A \nsucceq C|B$. Since $B|A \succeq C|B'$, we know $A \notin \mathcal{N}$ and so the expression $B|A$ is well-defined. Further, $B \notin \mathcal{N}$ by assumption. Thus, if $B'|A \nsucceq C|B$, it follows from by Axiom

1 (specifically, totality of the ordering) that $C|B \succ B'|A$. Now we apply Axiom 6a with $A' = A$, $C' = C$. We thus get $C|A \succ C'|A' = C|A$, which contradicts the reflexivity of the weak ordering (i.e., Axiom 1). Thus, $B'|A \succeq C|B$.

Finally, we must show that if the antecedent is strict (i.e., $B|A \succ C|B'$), then so is the conclusion (i.e., $B'|A \succ C|B$). Again, we suppose for the sake of contradiction that $B'|A \not\succ C|B$. Because $A, B \notin \mathcal{N}$ by the above reasoning, it follows from by Axiom 1 (specifically, totality of the ordering), $C|B \succeq B'|A$. Again, we apply Axiom 6a with $A' = A$ and $C' = C$ to obtain $C|A \succ C|A$, which again contradicts the reflexivity of the weak ordering. $\square$

B.2 QLL and Qualitative Favoring

In this section, we aim to prove a qualitative analog of Claim 2. To do so, we first state the definition of qualitative favoring and an analog of LL.

Definition 6. Given disjoint hypotheses H_1 and H_2, we say E *qualitatively supports* H_1 over H_2 at least as much as F if $H_1|E \cap (H_1 \cup H_2) \succeq H_1|F \cap (H_1 \cup H_2)$ for all orderings $\succeq$ satisfying our axioms of qualitative conditional probability and for which $F \cap (H_1 \cup H_2) \notin \mathcal{N}_{\preceq}$. The support is said to be strict when $\succeq$ is replaced with $\succ$.

By analogy to the quantitative case, we define "qualitative favoring" to be a special case of qualitative support.

Definition 7. Given disjoint hypotheses H_1 and H_2, we say E *qualitatively favors* H_1 to H_2 if E supports H_1 over H_2 at least as much as much as Ω. The favoring is called strict when $\succeq$ is replaced with $\succ$.

We'll show that qualitative favoring is equivalent to the notion axiomatized by a qualitative form of the law of likelihood.

- Qualitative law of likelihood (QLL): E favors θ_1 to θ_2 if $E|\theta_2 \sqsubset E|\theta_1$.

Here is our first major result.

Theorem 1. Suppose H_1 and H_2 are finite. Then E qualitatively favors H_1 over H_2 if (1) $\emptyset|\theta \sqsubset E|\theta$ for all $\theta \in H_1 \cup H_2$ and (2) $E|\theta_2 \sqsubseteq E|\theta_1$ for all $\theta_1 \in H_1$ and $\theta_2 \in H_2$. Under Assumption 1 (below), the converse holds as well. In both directions, the favoring inequality is strict exactly when the likelihood inequality is strict.

If $H_1 = \{\theta_1\}$ and $H_2 = \{\theta_2\}$ are simple, then E qualitatively favors H_1 over H_2 if and only if QLL entails so. No additional assumptions are required in this case.

Assumption 1. For all orderings $\sqsubseteq$ satisfying the axioms above and for all non-empty $H \subseteq \Theta$, there exists an ordering $\preceq$ satisfying the axioms such that (A) $H \in \mathcal{S}_{\preceq}$ and (B) $\theta|\Delta \succ \emptyset|\Delta$ for all $\theta \in H$.

We believe Assumption 1 is provable, but we have not yet produced a proof. The assumption more-or-less says that one's "prior" ordering over the hypotheses is unconstrained by the "likelihood" relation $\subseteq$. This is exactly analogous to the quantitative case. In that case, a joint distribution $Q^{\mathbb{E}}$ on $\Delta = \Theta \times \Omega^{\mathbb{E}}$ is determined (i) the measures $\langle P_\theta(\cdot)\rangle_{\theta\in\Theta}$ over $\Omega^{\mathbb{E}}$ and (ii) one's prior Q over Θ. Although the joint distribution $Q^{\mathbb{E}}$ is constrained by $\langle P_\theta(\cdot)\rangle_{\theta\in\Theta}$, one's prior Q on Θ is not, and so for any non-empty hypothesis $H \subseteq \Theta$, one can define some Q such that $Q(H) = 1$ and $Q(\theta) > 0$ for all $\theta \in H$. That's what the assumption says.

The theorem follows from the following three propositions, which we prove in the ensuing sections.

Proposition 1. Suppose $E \cap (H_1 \cup H_2) \notin \mathcal{N}$ for all orderings satisfying the axioms above and for which $H_1 \cup H_2 \notin \mathcal{N}_{\preceq}$. If Assumption 1 holds, then $\emptyset|\theta \sqsubset E|\theta$ for all $\theta \in H_1 \cup H_2$.

Proposition 2. Suppose $H_1 \cap H_2 = \emptyset$.

1. Suppose $E \cap (H_1 \cup H_2) \notin \mathcal{N}$. If $E|H_1 \succeq E|H_2$, then $H_1|E \cap (H_1 \cup H_2) \succeq H_1|H_1 \cup H_2$.
2. Suppose $H_1, H_2 \notin \mathcal{N}$. If $H_1|E \cap (H_1 \cup H_2) \succeq H_1|H_1 \cup H_2$, then $E|H_1 \succeq E|H_2$.

Further, if the antecedent of either conditional contains a strict inequality $\succ$, then so does the consequent.

Proposition 3. Suppose H_1 and H_2 are finite and that $H_1 \cap H_2 = \emptyset$. If $E|\theta_2 \sqsubseteq E|\theta_1$ for all $\theta_1 \in H_1$ and $\theta_2 \in H_2$, then $E|H_2 \preceq E|H_1$ for all orderings $\preceq$ (satisfying the axioms above) for which $H_1, H_2 \notin \mathcal{N}$. The converse holds under Assumption 1. In both directions, $\preceq$ can be replaced by the strict inequality $\prec$ if and only if the $\sqsubseteq$ can be replaced by the strict relation $\sqsubset$.

Notice that the second proposition holds in both directions when H_1 and H_2 are simple hypotheses by Axiom 0. The additional assumption is necessary only if H_1 and H_2 are composite.

Proof of Theorem 1: We consider the right-to-left direction first. Suppose (1) $\emptyset|\theta \sqsubset E|\theta$ for all $\theta \in H_1 \cup H_2$ and (2) $E|\theta_2 \sqsubseteq E|\theta_1$ for all $\theta_1 \in H_1$ and $\theta_2 \in H_2$. We must show E qualitatively favors H_1 over H_2. To do so, let $\preceq$ be any ordering satisfying the axioms such that $H_1 \cup H_2 \notin \mathcal{N}$. We must show $H_1|E \cap (H_1 \cup H_2) \succeq H_1|H_1 \cup H_2$. To do so, we must first show that $E \cap (H_1 \cup H_2) \notin \mathcal{N}$, so that the left-hand side of the last inequality is well-defined. Since $H_1 \cup H_2$ is finite and not a member of $\mathcal{N}$, it follows from Part 4 of Lemma 4 that there is some $\eta \in H_1 \cup H_2$ such that $\{\eta\} \notin \mathcal{N}$. By assumption (1) that $\emptyset|\theta \sqsubset E|\theta$ for all $\theta \in H_1 \cup H_2$, we know $\emptyset|\eta \sqsubset E|\eta$ in particular. By Axiom 0 and the fact that $\{\eta\} \notin \mathcal{N}$, we can infer $\emptyset|\eta \prec E|\eta$. Lemma 7 then entails that $\emptyset \cap \{\eta\}|\Delta \prec E \cap \{\eta\}|\Delta$. So by Axiom 2, $E \cap \{\eta\} \notin \mathcal{N}$. Because $E \cap \{\eta\} \subseteq E \cap (H_1 \cup H_2) \notin \mathcal{N}$ as desired.

Now that we know $E \cap (H_1 \cup H_2) \notin \mathcal{N}$, we argue that $H_1|E \cap (H_1 \cup H_2) \succeq H_1|H_1 \cup H_2$ via proof by cases. Since $H_1 \cup H_2 \notin \mathcal{N}$, it follows from Lemma 4 that at least one of H_1 and H_2 is not a member of $\mathcal{N}$. Thus, there are three cases to consider: (A) Neither is in $\mathcal{N}$, (B) $H_1 \in \mathcal{N}$ but $H_2 \notin \mathcal{N}$, and (C) $H_2 \in \mathcal{N}$ but $H_1 \notin \mathcal{N}$.

Case A: Suppose $H_1, H_2 \notin \mathcal{N}$. By Proposition 3 and our first assumption (that $\emptyset|\theta \sqsubset E|\theta$ for all $\theta \in H_1 \cup H_2$), it follows that $E|H_2 \preceq E|H_1$. Since $E \cap (H_1 \cup H_2) \notin \mathcal{N}$ and $E|H_2 \preceq E|H_1$, Part 1 of Proposition 2 allows us to conclude that $H_1|E \cap (H_1 \cup H_2) \succeq H_1|H_1 \cup H_2$, as desired.

Case B: Suppose $H_1 \in \mathcal{N}$ but $H_2 \notin \mathcal{N}$. Since $H_1 \in \mathcal{N}$, Part 1 of Lemma 6 entails both that $H_1|E \cap (H_1 \cup H_2) \sim \emptyset|\Delta$ and $H_1|H_1 \cup H_2 \sim \emptyset|\Delta$. By transitivity of $\sim$ (which follows from Axiom 1), we know that $H_1|E \cap (H_1 \cup H_2) \sim H_1|H_1 \cup H_2$, and hence, $H_1|E \cap (H_1 \cup H_2) \succeq H_1|H_1 \cup H_2$ as desired.

Case C: Suppose $H_2 \in \mathcal{N}$ but $H_1 \notin \mathcal{N}$. First note that because $H_2 \in \mathcal{N}$, we obtain that $\neg H_2 := \Delta \setminus H_2 \in \mathcal{S}$ by definition of $\mathcal{S}$. Further, because $H_1 \cap H_2 = \emptyset$, we know $E \cap H_1 \subseteq H_1 \subseteq \neg H_2$, and hence $(E \cap (H_1 \cup H_2)) \cap \neg H_2 = E \cap H_1$. Thus:

$$
\begin{aligned}
H_1|E \cap (H_1 \cup H_2) \quad &\sim \quad H_1|(E \cap (H_1 \cup H_2)) \cap \neg H_2 \text{ by Part 6 of Lemma 6 since } \neg H_2 \in \mathcal{S} \\
&= \quad H_1|E \cap H_1 \text{ as just shown} \\
&\sim \quad E \cap H_1|E \cap H_1 \text{ by Axiom 4} \\
&\sim \quad H_1|H_1 \text{ by Axiom 3} \\
&\sim \quad H_1|H_1 \cup H_2 \text{ by Part 6 of Lemma 6 since } H_2 \in \mathcal{N}
\end{aligned}
$$

In the left to right direction, suppose E qualitatively favors H_1 over H_2, or in other words, $H_1|E \cap (H_1 \cup H_2) \succeq H_1|H_1 \cup H_2$ for all orderings $\preceq$ for which $H_1 \cup H_2 \notin \mathcal{N}_{\preceq}$. Further, suppose Assumption 1 holds. We must show (1) $\emptyset|\theta \sqsubset E|\theta$ for all $\theta \in H_1 \cup H_2$ and (2) $E|\theta_2 \sqsubseteq E|\theta_1$ for all $\theta_1 \in H_1$ and $\theta_2 \in H_2$.

To show (1), recall we've assumed $H_1|E \cap (H_1 \cup H_2) \succeq H_1|H_1 \cup H_2$ for all orderings $\preceq$ satisfying the axioms and for which $H_1 \cup H_2 \notin \mathcal{N}$. Thus, $E \cap (H_1 \cup H_2) \notin \mathcal{N}$ for all orderings $\preceq$ satisfying the axioms and for which $H_1 \cup H_2 \notin \mathcal{N}$. By Proposition 1 and Assumption 1, we're done.

To show (2), recall we've assumed $H_1|E \cap (H_1 \cup H_2) \succeq H_1|H_1 \cup H_2$ for all orderings $\preceq$ for which $H_1 \cup H_2 \notin \mathcal{N}$. Thus, it is also the case theat $H_1|E \cap (H_1 \cup H_2) \succeq H_1|H_1 \cup H_2$ for all orderings $\preceq$ for which *both* $H_1 \notin \mathcal{N}$ and $H_2 \notin \mathcal{N}$. It follows from Part 2 of Proposition 2 that $E|H_2 \preceq E|H_1$ for all orderings $\preceq$ for which *both* $H_1 \notin \mathcal{N}$ and $H_2 \notin \mathcal{N}$. Proposition 3 then entails our desired conclusion that $E|\theta_2 \sqsubseteq E|\theta_1$ for all $\theta_1 \in H_1$ and $\theta_2 \in H_2$.

$\square$

B.3 Proving Proposition 1

Proof of Proposition 1: By contraposition. Assume that it's not the case that $\emptyset|\theta \sqsubset E|\theta$ for all θ. We show there is an ordering $\preceq$ satisfying the axioms such that $H_1 \cup H_2 \notin \mathcal{N}_{\preceq}$ and $E \cap (H_1 \cup H_2) \in \mathcal{N}_{\preceq}$.

Because it's not the case that $\emptyset|\theta \sqsubset E|\theta$ for all $\theta \in H_1 \cup H_2$, there exists $\theta \in H_1 \cup H_2$ such that $\emptyset|\theta \not\sqsubset E|\theta$. By Axiom 2 (for $\sqsubseteq$), $\theta \notin \mathcal{N}_{\sqsubseteq}$ and so the expression $\cdot|\theta$ is well-defined (with respect to $\sqsubseteq$). Since $\emptyset|\theta \not\sqsubset E|\theta$ and $\cdot|\theta$ is well-defined, it follows from Axiom 1 (specifically totality of $\sqsubseteq$) that $E|\theta \sqsubseteq \emptyset|\theta$.

Using Assumption 1, pick any ordering $\preceq$ satisfying the axioms such that (1) $H = \{\theta\} \in \mathcal{S}_{\preceq}$ and (2) $\theta|\Delta \succ \emptyset|\Delta$. By 2 and Axiom 2, we know $\{\theta\} \notin \mathcal{N}_{\preceq}$, and since $\{\theta\} \subseteq H_1 \cup H_2$, it follows from Part 3 of Lemma 4 that $H_1 \cup H_2 \notin \mathcal{N}_{\preceq}$.

Because $E|\theta \sqsubseteq \emptyset|\theta$, Axiom 0 entails that $E|\theta \preceq \emptyset|\theta$. By Lemma 7, it follows that $E \cap \{\theta\}|\Delta \preceq \emptyset|\Delta$, and hence, $E \cap \{\theta\} \in \mathcal{N}_{\preceq}$ by Axiom 2. By assumption, $\{\theta\} \in \mathcal{S}_{\preceq}$, and so $\neg\{\theta\} \in {}_{\preceq}$. Because $E \cap \neg\{\theta\} \subseteq \neg\{\theta\}$, Part 3 of Lemma 4 allows us to conclude $E \cap \neg\{\theta\} \in \mathcal{N}$. Because both $E \cap \{\theta\}$ and $E \cap \neg\{\theta\}$ are null sets, Part 4 of Lemma 4 entails $E = (E \cap \neg\{\theta\}) \cup (E \cap \{\theta\})$ is in $\mathcal{N}$. Again, because $E \cap (H_1 \cup H_2) \subseteq E$ and $E \in \mathcal{N}$, we know from Part 3 of Lemma 4 that $E \cap (H_1 \cup H_2) \in \mathcal{N}$, and we're done.

□

B.3.1 Proof of Proposition 2

The proof of Proposition 2 requires the following lemmas.

Lemma 9. Suppose $H_1 \cap H_2 = \emptyset$. If $E|H_1 \succeq E|H_2$, then $E|H_1 \succeq E|(H_1 \cup H_2)$. The converse holds if $H_2 \notin \mathcal{N}$. Further, if either side of the biconditional is $\succ$, the other side is $\succ$ too.

Proof. In the left-to-right direction, assume that $E|H_1 \succeq E|H_2$, so that $H_1, H_2 \notin \mathcal{N}$. Suppose, for the sake of contradiction that $E|H_1 \not\succeq E|(H_1 \cup H_2)$. Since $H_1 \notin \mathcal{N}$, it follows from Lemma 4 (part 3) that $H_1 \cup H_2 \notin \mathcal{N}$, and hence, the expression $\cdot|(H_1 \cup H_2)$ is well-defined. Since $E|H_1 \not\succeq E|(H_1 \cup H_2)$, by totality of $\preceq$ (part of Axiom 1), it follows that $E|H_1 \prec E|(H_1 \cup H_2)$. By Axiom 4, we get $E \cap (H_1 \cup H_2)|H_1 \cup H_2 \succ E \cap H_1|H_1$. We now apply Lemma 8 with:

$$\begin{aligned} A &= H_1 \cup H_2 \\ B &= E \cap (H_1 \cup H_2) \qquad B' = H_1 \\ C &= E \cap H_1 \end{aligned}$$

Clearly, the necessary containment relations hold and we just showed $B \notin \mathcal{N}$. Since $B|A \succ C|B'$, we get that $B'|A \succ C|B$. So:

$$H_1|H_1 \cup H_2 \succ E \cap H_1|E \cap (H_1 \cup H_2) \tag{13}$$

Further, since $E|H_2 \preceq E|H_1$ and $E|H_1 \prec E|(H_1 \cup H_2)$, we know $E|H_2 \prec E|H_1 \cup H_2$. By similar reasoning as before (swapping H_1, H_2 in our application of Lemma 8), we can thus conclude that:

$$H_2|H_1 \cup H_2 \succ E \cap H_2|E \cap (H_1 \cup H_2) \tag{14}$$

We now apply Axiom 5 to equations (13) and (14). To do so, let

$$\begin{aligned} A &= H_1, & A' &= E \cap H_1 \\ B &= H_2, & B' &= E \cap H_2 \\ C &= H_1 \cup H_2, & C' &= E \cap (H_1 \cup H_2) \end{aligned}$$

Note that $A \cap B = \emptyset$ by our initial assumption that $H_1 \cap H_2 = \emptyset$. From this, it follows from $A' \cap B' = \emptyset$. Now, Equation 13 says that $A|C \succ A'|C'$ and Equation 14 says that $B|C \succ B'|C'$. Thus, we get that $A \cup B|C \succ A' \cup B'|C'$ by Axiom 5. In other words:

$$H_1 \cup H_2|H_1 \cup H_2 \succ (E \cap H_1) \cup (E \cap H_2)|E \cap (H_1 \cup H_2) = E \cap (H_1 \cup H_2)|E \cap (H_1 \cup H_2) \tag{15}$$

This equation says that $X|X \succ Y|Y$ for two events X, Y. This contradicts Axiom 3, so we get that $E|H_1 \succeq E|(H_1 \cup H_2)$, as desired.

Note that if $E|H_1 \succ E|H_2$, Equation 13 would be $\succeq$, since we would get $E|H_1 \preceq E|(H_1 \cup H_2)$ from our initial steps. However, Equation 14 would still be $\succ$, since $E|H_2 \prec E|H_1$. And for Axiom 5, as long as either premise is $\succ$, the conclusion is $\succ$. Thus, we would still reach a contradiction, and thus get $E|H_1 \succ E|(H_1 \cup H_2)$.

Now, for the right-to-left direction assume $E|H_1 \succeq E|(H_1 \cup H_2)$ and that $H_2 \notin \mathcal{N}$. We want to show $E|H_1 \succeq E|H_2$.

Note that if $E \cap (H_1 \cup H_2) \in \mathcal{N}$, then since $E \cap H_i \subseteq E \cap (H_1 \cup H_2)$ for $i = 1, 2$, we know from Part 3 of Lemma 4 that $E \cap H_i \in \mathcal{N}$. It follows that:

$$\begin{aligned} E|H_1 &\sim E \cap H_1|H_1 \text{ by Axiom 4} \\ &\sim \emptyset|H_2 \text{ by Part 2 of Lemma 6, since } E \cap H_1 \in \mathcal{N} \\ &\sim E \cap H_2|H_2 \text{ by Part 2 of Lemma 6} E \cap H_2 \in \mathcal{N} \\ &\sim E \cap H_2|H_2 \text{ by Axiom 4} \end{aligned}$$

and $E|H_1 \succeq E|H_2$ as desired.

Thus, assume that $E \cap (H_1 \cup H_2) \notin \mathcal{N}$, and define:

$$\begin{aligned} A &= H_1 \cup H_2 & & \\ B &= E \cap (H_1 \cup H_2) & B' &= H_1 \\ C &= E \cap H_1 & & \end{aligned}$$

We use Lemma 8. Because $C|B' \succeq B|A$ (by Axiom 4) and we have assumed that $B = E \cap (H_1 \cup H_2) \notin \mathcal{N}$, we get that $C|B \succeq B'|A$. This means that:

$$E \cap H_1|E \cap (H_1 \cup H_2) \succeq H_1|H_1 \cup H_2 \tag{16}$$

Suppose, for the sake of contradiction, that $E|H_1 \not\succeq E|H_2$. We know $H_1 \notin \mathcal{N}$ (because $E|H_1 \succeq E|(H_1 \cup H_2)$) and we've simply assumed that $H_2 \notin \mathcal{N}$ for this direction of the proof. Thus, both $\cdot|H_1$ and $\cdot|H_2$ are well-defined, and so by the totality of $\preceq$ (part of Axiom 1), it follows that $E|H_2 \succ E|H_1$. Because $E|H_1 \succeq E|(H_1 \cup H_2)$, by transitivity we get $E|H_2 \succ E|(H_1 \cup H_2)$. So if we apply the same reasoning we did to derive Equation 16 (switching H_1 and H_2),we get that:

$$E \cap H_2|E \cap (H_1 \cup H_2) \succ H_2|H_1 \cup H_2 \tag{17}$$

Note that this comparison is strict as the premise was strict. From here, as in the other direction of the proof, we use Axiom 5 to derive $E \cap (H_1 \cup H_2)|E \cap (H_1 \cup H_2) \succ H_1 \cup H_2|H_1 \cup H_2$, contradicting Axiom 3 (the conclusion is strict because at least one of the premises is strict). So we've shown $E|H_1 \succeq E|H_2$.

If our assumption contained a strict inequality (i.e., $E|H_1 \succ E|(H_1 \cup H_2)$), then the only changes in the preceding argument would be that Equation 16 would be strict whereas Equation 17 would not. This would still lead to a contradiction, and so in this case we would get $E|H_1 \succ E|H_2$. □

Proof of Proposition 2. Suppose $H_1 \cap H_2 = \emptyset$.

[Part 1:] Suppose $E|H_1 \succeq E|H_2$ and that $E \cap (H_1 \cup H_2) \notin \mathcal{N}$. By Lemma 9, it follows that $E|H_1 \succeq E|(H_1 \cup H_2)$. Applying Axiom 4, we get

$$E \cap H_1|H_1 \succeq E \cap (H_1 \cup H_2)|H_1 \cup H_2 \tag{18}$$

Define:

$$\begin{aligned} &A = H_1 \cup H_2 \\ &B = E \cap (H_1 \cup H_2) \qquad B' = H_1 \\ &C = E \cap H_1 \end{aligned}$$

Note that Equation 18 says that $C|B' \succeq B|A$ and we've assumed that $B = E \cap (H_1 \cup H_2) \notin \mathcal{N}$. So Lemma 8 tells us that $C|B \succeq B'|A$, i.e.,

$$E \cap H_1|E \cap (H_1 \cup H_2) \succeq H_1|H_1 \cup H_2$$

Applying Axiom 4 to the left-hand side of that inequality yields the desired result. Note that if we had assumed $E|H_1 \succ E|H_2$, then our conclusion would contain $\succ$ because both Lemma 8 and Lemma 9 yield strict comparisons.

In the right to left direction, suppose $H_1|E \cap (H_1 \cup H_2) \succeq H_1|H_1 \cup H_2$ and that $H_1, H_2 \notin \mathcal{N}$. By Axiom 4:

$$\begin{aligned} &H_1|E \cap (H_1 \cup H_2) \\ &\quad \sim H_1 \cap E \cap (H_1 \cup H_2)|E \cap (H_1 \cup H_2) \\ &\quad = E \cap H_1|E \cap (H_1 \cup H_2) \end{aligned}$$

So, we know $E \cap H_1 | E \cap (H_1 \cup H_2) \succeq H_1 | H_1 \cup H_2$. Now, we apply Lemma 8 again with:

$$\begin{aligned} & A = H_1 \cup H_2 \\ & B = H_1 \qquad\qquad B' = E \cap (H_1 \cup H_2) \\ & C = E \cap H_1 \end{aligned}$$

The containment relations are satisfied; $B = H_1 \notin \mathcal{N}$ by assumption, and we just showed $C|B' \succeq B|A$. Thus, we get $C|B \succeq B'|A$, i.e.,

$$E \cap H_1 | H_1 \succeq E \cap (H_1 \cup H_2) | (H_1 \cup H_2).$$

Applying Axiom 4 to both sides of the last inequality yields $E|H_1 \succeq E|H_1 \cup H_2$. So by Lemma 9 and the assumption that $H_2 \notin \mathcal{N}$, we get $E|H_1 \succeq E|H_2$, as desired. As before, if the premise were $\succ$, the conclusion would also be $\succ$ because the necessary lemmas would yield strict comparisons. $\square$

B.3.2 Proving Proposition 3

The proof of Proposition 3 requires a few lemmata.

Lemma 10. Suppose $H_1 \cap H_2 = \emptyset$. If $E|H_1 \sim E|H_2$ then $E|(H_1 \cup H_2) \sim E|H_1 \sim E|H_2$.

Proof. First note that if $E|H_1 \sim E|H_2$, then $H_1, H_2 \notin \mathcal{N}$. Hence, $H_1 \cup H_2 \notin \mathcal{N}$ by Part 3 of Lemma 4. So expressions of the form $\cdot | H_1 \cup H_2$ are well-defined.

Suppose for the sake of contradiction that $E|(H_1 \cup H_2) \not\sim E|H_1 \sim E|H_2$. Since expressions of the form $\cdot | H_1 \cup H_2$ are well-defined, Axiom 1 (specifically totality of the ordering), either $E|(H_1 \cup H_2) \prec E|H_1 \sim E|H_2$ or vice versa. We thus consider two cases below.

In both cases, we use the fact that $E \cap (H_1 \cup H_2) \notin \mathcal{N}$. We can prove that by contradiction as follows. Suppose that $E \cap (H_1 \cup H_2) \in \mathcal{N}$. Then by Part 3 of Lemma 4, we know that $E \cap H_1 \in \mathcal{N}$ since $E \cap H_1 \subseteq \cap (H_1 \cup H_2)$. Hence:

$$\begin{aligned} E|H_1 \cup H_2 \quad & \sim \quad E \cap (H_1 \cup H_2) | H_1 \cup H_2 \text{ by Axiom 4} \\ & \sim \quad \emptyset | H_1 \text{ by Part 2 of Lemma 6 since } E \cap (H_1 \cup H_2) \in \mathcal{N} \\ & \sim \quad E \cap H_1 | H_1 \text{ by Part 2 of Lemma 6 since } E \cap H_1 \in \mathcal{N} \\ & \sim \quad E|H_1 \text{ by Axiom 4} \end{aligned}$$

which contradictions our assumption that $E|(H_1 \cup H_2) \not\sim E|H_1$.

We now consider the two cases discussed above.

Case 1: Suppose $E|(H_1 \cup H_2) \prec E|H_1 \sim E|H_2$. Define:

$$\begin{aligned} & A = H_1 \cup H_2, \\ & B = E \cap (H_1 \cup H_2), \qquad B' = H_2 \\ & C = E \cap H_2 \end{aligned}$$

Then $C|B' \succ B|A$ by assumption of the case and Axiom 4. We also showed $B = E \cap (H_1 \cup H_2) \notin \mathcal{N}$ above. So by Lemma 8 we have $C|B \succeq B'|A$, i.e.,

$$H_2|H_1 \cup H_2 \prec E \cap H_2|E \cap (H_1 \cup H_2)$$

Analogous reasoning shows that

$$H_1|H_1 \cup H_2 \prec E \cap H_1|E \cap (H_1 \cup H_2)$$

Together, those two equation and Axiom 5 yields that

$$H_1 \cup H_2|H_1 \cup H_2 \prec E \cap (H_1 \cup H_2)|E \cap (H_1 \cup H_2).$$

But this contradicts Axiom 3.

Case 2: Suppose $E|H_2 \sim E|H_1 \prec E|(H_1 \cup H_2)$. We get a similar contradiction. Define:

$$\begin{aligned} &A = H_1 \cup H_2, \\ &B = E \cap (H_1 \cup H_2), \qquad && B' = H_1 \\ &C = E \cap H_1 \end{aligned}$$

Then $C|B' \succ B|A$ by Axiom 4 and assumption of the case. We showed $B = E \cap (H_1 \cup H_2) \notin \mathcal{N}$ above. By Lemma 8 we have $C|B \succeq B'|A$, i.e.,

$$H_1|H_1 \cup H_2 \prec E \cap H_1|E \cap (H_1 \cup H_2)$$

Analogous reasoning shows:

$$H_2|H_1 \cup H_2 \prec E \cap H_2|E \cap (H_1 \cup H_2)$$

Together, those two equation and Axiom 5 yields that

$$H_1 \cup H_2|H_1 \cup H_2 \prec E \cap (H_1 \cup H_2)|E \cap (H_1 \cup H_2).$$

But this contradicts Axiom 3. $\square$

Lemma 11. Suppose $H_1 \cap H_2 = \emptyset$.

a.) If $E|H_1, E|H_2 \preceq E|H_3$, then $E|(H_1 \cup H_2) \preceq E|H_3$.

b.) If $E|H_3 \preceq E|H_1, E|H_2$, then $E|H_3 \preceq E|(H_1 \cup H_2)$.

If the premise is not strict AND $E|H_1 \sim E|H_2$, then the conclusion is not strict. Otherwise, the conclusion is strict.

Proof. We first prove a). By totality of $\preceq$ (which is part of Axiom 1), one of the following three cases must hold:

1. $E|H_2 \prec E|H_1$,

2. $E|H_1 \prec E|H_2$,

3. $E|H_1 \sim E|H_2$,

If the first case holds, then Lemma 9 entails $E|(H_1 \cup H_2) \prec E|H_1$. Hence, Axiom 1 (transitivity) and the assumption that $E|H_1, E|H_2 \preceq E|H_3$ together entail that $E|(H_1 \cup H_2) \prec E|H_3$.

If the second case holds, we reason analogously to show $E|(H_1 \cup H_2) \prec E|H_2 \prec E|H_3$.

If the third case holds, then Lemma 10 entails that $E|(H_1 \cup H_2) \sim E|H_1 \preceq E|H_3$, and we're done. Note that in the first two cases the conclusion is always strict. And even in this case, if $E|H_1 \sim E|H_2 \prec E|H_3$, the conclusion would be strict.

The proof of part b) is analogous to a). □

We're now read to prove the main proposition.

Proof of Proposition 3. In the right-to-left direction, suppose $E|\theta_2 \sqsubseteq E|\theta_1$ for all $\theta_1 \in H_1$ and $\theta_2 \in H_2$. We want to show $E|H_2 \preceq E|H_1$ for all orderings $\preceq$ for which $H_1, H_2 \notin \mathcal{N}$. So let $\preceq$ be any such ordering.

We show $E|H_2 \preceq E|H_1$ by induction on the maximum of the number of elements in H_1 or H_2.

In the base case, suppose both $H_1 = \{\theta_1\}$ and $H_2 = \{\theta_2\}$ both have one element. Then by Axiom 0 and the assumption that $H_1, H_2 \notin \mathcal{N}$, it immediately follows that $E|H_2 \preceq E|H_1$.

For the inductive step, suppose the result holds for all natural numbers $m \leq n$, and assume that $H_1 = \{\theta_{1,1}, \ldots \theta_{1,k}\}$ and $H_2 = \{\theta_{2,1} \ldots, \theta_{2,l}\}$ where either k or l (or both) is equal to $n+1$. Define $H_1' = \{\theta_{1,1}\}$ and $H_1'' = H_1 \setminus H_1'$, and similarly, define $H_2' = \{\theta_{2,1}\}$ and $H_2'' = H_2 \setminus H_2'$. Then H_1', H_1'', H_2', and H_2'' all have fewer than n elements, and by assumption, $E|\theta_2 \sqsubseteq E|\theta_1$ for all $\theta_2 \in H_2', H_2''$ and all $\theta_1 \in H_1', H_1''$. Suppose, for the moment, that H_1', H_1'', H_2', and H_2'' are all non-null. Then by inductive hypothesis, it would follow that $E|H_1', E|H_1'' \preceq E|H_2', E|H_2''$. By repeated application of Lemma 11, it follows that $E|H_1 = E|(H_1' \cup H_1'') \preceq E|(H_2' \cup H_2'') = E|H_2$.

If any of H_1', H_1'', H_2', and H_2'' are members of $\mathcal{N}$, the proof involves only a small modification. Not that since $H_1 \notin \mathcal{N}$ and $H_1 = H_1' \cup H_1''$, at least one of H_1' and H_1'' must be non-null by Part 4 of Lemma 4. Suppose, without lost of generality, that $H_1' \in \mathcal{N}$ but $H_1'' \notin \mathcal{N}$. Then by Part 6 of Lemma 6, we get $E|H_1 = E|H_1' \cup H_1'' \sim E|H_1''$ and the inductive hypothesis already applies to H_1'' because it contains fewer than n elements. Similar, reasoning apply to H_2.

Note that if the premise were $\sqsubset$, Axiom 0 would prove the strict version of the base case. In the inductive step, every inequality would also be strict because Lemma 11 preserves the strictness. Thus, we would get $E|H_2 \prec E|H_1$.

In the left-to-right direction, we proceed by contraposition. Suppose that Assumption 1 holds and that it's *not* the case that $E|\theta_2 \sqsubseteq E|\theta_1$ for all $\theta_1 \in H_1$ and $\theta_2 \in H_2$. We construct an ordering $\preceq$ such that $H_1, H_2 \notin \mathcal{N}$ and $E|H_2 \not\preceq E|H_1$.

To do so, note that because it's *not* the case that $E|\theta_2 \sqsubseteq E|\theta_1$ for all $\theta_1 \in H_1$ and $\theta_2 \in H_2$, we can fix $\theta_1 \in H_1$ and $\theta_2 \in H_2$ for which $E|\theta_2 \not\sqsubseteq E|\theta_1$. By totality of $\sqsubseteq$ (which is part of Axiom 1), it immediately follows that $E|\theta_1 \sqsubset E|\theta_2$.

Using Assumption 1 with $H = \{\theta_1, \theta_2\}$, define $\preceq$ so that it satisfies all of the axioms and:

$$\begin{aligned} \theta_1|\Delta, \theta_2|\Delta &\succ \emptyset|\Delta \\ A|\Delta &\sim \emptyset|\Delta \text{ if } \theta_1, \theta_2 \notin A \end{aligned}$$

Because $\theta_1|\Delta, \theta_2|\Delta \succ \emptyset|\Delta$, Axiom 2 entails that $\{\theta_1\}, \{\theta_2\} \notin \mathcal{N}$. Since $\{\theta_1\}, \{\theta_2\} \notin \mathcal{N}$ and $E|\theta_1 \sqsubset E|\theta_2$, by Axiom 0 we obtain that $E|\theta_1 \prec E|\theta_2$.

Now consider $H_1' = H_1 \setminus \{\theta_1\}$ and $H_2' = H_2 \setminus \{\theta_2\}$. Since H_1 and H_2 are disjoint, neither H_1' nor H_2' contain either θ_1 or θ_2. Thus by construction of $\preceq$, it follows that $H_1', H_2' \in \mathcal{N}$. So by Part 6 of Lemma 6, we obtain that

$$\begin{aligned} E|H_1 &= E|H_1' \cup \{\theta_1\} \sim E|\theta_1 \text{ and} \\ E|H_2 &= E|H_2' \cup \{\theta_2\} \sim E|\theta_2 \end{aligned}$$

Since $E|\theta_1 \prec E|\theta_2$, it follows that $E|H_1 \prec E|H_2$, as desired.

If the premise were $\prec$, the conclusion would clearly be $\sqsubset$, by simply swapping all strict and weak inequalities in the preceding argument. □

B.4 A Qualitative Likelihood Principle

We start with analogs of Bayesian favoring and posterior equivalence.

Definition 8. E and F are *qualitative favoring equivalent* if E qualitatively supports H_1 to H_2 at least as much as F and vice versa, for any two disjoint hypotheses H_1 and H_2. This means that, for all disjoint H_1, H_2, and for all $\preceq$ we have (1) $E \cap (H_1 \cup H_2) \in \mathcal{N}_{\preceq}$ if and only if $F \cap (H_1 \cup H_2) \in \mathcal{N}_{\preceq}$, and (2) if $E \cap (H_1 \cup H_2), F \cap (H_1 \cup H_2) \notin \mathcal{N}_{\preceq}$, then

$$H_1|E \cap (H_1 \cup H_2) \sim H_1|F \cap (H_1 \cup H_2)$$

Definition 9. E and F are *qualitative posterior equivalent* if for any ordering $\preceq$ satisfying the axioms: (1) $E \notin \mathcal{N}$ if and only if $F \notin \mathcal{N}$ and (2) $H|E \sim H|F$ for any $H \subseteq \Theta$ whenever $E \notin \mathcal{N}$ and $F \notin \mathcal{N}$.

We start by showing that these two definitions are equivalent.

Claim 4. E and F are qualitative posterior equivalent if and only if they are qualitative favoring equivalent.

Proof. In the right to left direction, assume E and F are qualitative favoring equivalent. We must show that E and F are posterior equivalent, i.e., for all orderings satisfying the above axioms (1) $E \in \mathcal{N}_{\preceq}$ if and only if $F \in {}_{\preceq}$ and (2) $H|E \sim H|F$ for any hypothesis $H \subseteq \Theta$ whenever $E, F \notin {}_{\preceq}$. So let $\preceq$ be an arbitrary ordering and $H \subseteq \Theta$ be an arbitrary hypothesis. Define $H_1 = H$ and

$H_2 = \Theta \setminus H$. Notice that H_1 and H_2 are disjoint and $H_1 \cup H_2 = \Theta$ is the entire space. We'll now proceed to prove both 1 and 2.

To show 1, note that $E = E \cap (H_1 \cup H_2)$ and $F = F \cap (H_1 \cup H_2)$. Since E and F are qualitatively favoring equivalent, it follows that $E \cap (H_1 \cup H_2) \in \mathcal{N}$ if and only if $F \cap (H_1 \cup H_2) \in \mathcal{N}$, from which it follows that $E \in \mathcal{N}$ if and only if $F \in \mathcal{N}$ as desired.

To show the 2, again note that by qualitative favoring equivalence and transitivity, we get

$$H|E = H_1|E = H_1|E \cap (H_1 \cup H_2) \sim H_1|F \cap (H_1 \cup H_2) = H_1|F = H|F$$

In the left to right direction, assume that E and F are qualitatively posterior equivalent. Let H_1 and H_2 be disjoint hypotheses, and let $\preceq$ be an ordering satisfying the axioms above. We must show that (1) $E \cap (H_1 \cup H_2) \in \mathcal{N}_{\preceq}$ if and only if $F \cap (H_1 \cup H_2)$ and (2) $H_1|E \cap (H_1 \cup H_2) \sim_1 |F \cap (H_1 \cup H_2)$ whenever both sides of that expression are well defined.

To show 1, notice that if either E or F is in $\mathcal{N}$, then because E and F are posterior equivalent, so is the other. Because $E \cap (H_1 \cup H_2) \subseteq E$ and $F \cap (H_1 \cup H_2) \subseteq F$, it would then follow from Part 3 of Lemma 4 entails that $E \cap (H_1 \cup H_2), F \cap (H_1 \cup H_2) \in \mathcal{N}$. Thus, if either E or F is in $\mathcal{N}$, then $E \cap (H_1 \cup H_2)$ if and only if $F \cap (H_1 \cup H_2) \in \mathcal{N}$.

So consider the case in which neither E nor F is a member of $\mathcal{N}$, and thus, the expressions $\cdot|E$ and $\cdot|F$ are well-defined. We'll show that if $E \cap (H_1 \cup H_2) \in \mathcal{N}$, then $F \cap (H_1 \cup H_2)$; the converse is proved identically. If $E \cap (H_1 \cup H_2) \in \mathcal{N}$, then both $E \cap H_1 \in \mathcal{N}$ and $E \cap H_2 \in \mathcal{N}$ by Part 3 of Lemma 4 since $E \cap H_1, E \cap H_2 \subseteq E \cap (H_1 \cup H_2)$. Therefore:

$$\begin{aligned}
\emptyset|E \sim & E \cap H_1|E \text{ by Part 2 of Lemma 6 and as } E \cap H_1 \in \mathcal{N} \\
\sim & H_1|E \text{ by Axiom 4} \\
\sim & H_1|F \text{ since } E\&F \text{ are posterior equivalent and } E, F \notin \mathcal{N} \\
\sim & F \cap H_1|F \text{ by Axiom 4}
\end{aligned}$$

Thus, $\emptyset|E \sim F \cap H_1|F$ and so $F \cap H_1 \in \mathcal{N}$ by Part 2 of Lemma 6. By identical reasoning, $F \cap H_2 \in \mathcal{N}$. By Part 4 of Lemma 4, it follows that $F \cap (H_1 \cup H_2) \in \mathcal{N}$ as desired.

Now we show 2, i.e., if $E \cap (H_1 \cup H_2), F \cap (H_1 \cup H_2) \notin \mathcal{N}$, then $H_1|E \cap (H_1 \cup H_2) \sim H_1|F \cap (H_1 \cup H_2)$. Suppose, for the sake of contradiction, that $H_1|E \cap (H_1 \cup H_2) \not\sim H_1|F \cap (H_1 \cup H_2)$. By Axiom 1, we obtain that either $H_1|E \cap (H_1 \cup H_2) \prec H_1|F \cap (H_1 \cup H_2)$ or vice versa. Without loss of generality, assume

$$H_1|E \cap (H_1 \cup H_2) \prec H_1|F \cap (H_1 \cup H_2) \tag{19}$$

Since $E \cap (H_1 \cup H_2), F \cap (H_1 \cup H_2) \notin \mathcal{N}$, it follows from Part 3 of Lemma 4 that $E, F \notin \mathcal{N}$. Because E and F are posterior equivalent, we know that

$$H_1 \cup H_2|E \sim H_1 \cup H_2|F. \tag{20}$$

Define:

$$\begin{aligned} &A = F && A' = E \\ &B = F \cap (H_1 \cup H_2) && B' = E \cap (H_1 \cup H_2) \\ &C = F \cap H_1 && C' = E \cap H_1 \end{aligned}$$

Thus:

$$\begin{aligned} B|A &= F \cap (H_1 \cup H_2)|F \\ &\sim H_1 \cup H_2|F && \text{by Axiom 4} \\ &\sim H_1 \cup H_2|E && \text{by Equation 20} \\ &\sim E \cap (H_1 \cup H_2)|E && \text{by Axiom 4} \\ &= B'|A' \end{aligned}$$

Further:

$$\begin{aligned} C|B &= F \cap H_1|F \cap (H_1 \cup H_2) \\ &= H_1 \cap (F \cap (H_1 \cup H_2))|F \cap (H_1 \cup H_2) \\ &\sim H_1|F \cap (H_1 \cup H_2) && \text{by Axiom 4} \\ &\succ H_1|E \cap (H_1 \cup H_2) && \text{by Equation 19} \\ &\sim C'|B' && \text{reversing the above steps} \end{aligned}$$

Note that since $C|B \succ C'|B'$, we know that $C \notin \mathcal{N}$. Further, since $B|A \sim B'|A'$, Axiom 6b entails $C|A \succ C'|A'$, i.e., that $F \cap H_1|F \succ E \cap H_1|E$. But then Axiom 4 entails $H_1|F \succ H_1|E$, and that contradicts the assumption that E and F are qualitative posterior equivalent. Thus, by contradiction, E and F are qualitative favoring equivalent. □

Now, we prove that a qualitative analog of LP is a sufficient condition for qualitative favoring/posterior equivalence. To do so, we first introduce the definition of conditional independence.

Definition 10. For any $C \notin \mathcal{N}$, the events A and B are said to be *conditionally independent* given C if $A|B \cap C \sim A|C$ or $B \cap C$ is null. In this case, we write $A \perp\!\!\!\perp_C B$.

Lemma 12. If $A \perp\!\!\!\perp_C B$, then $B \perp\!\!\!\perp_C A$.

Proof. Suppose $A \perp\!\!\!\perp_C B$. There are two cases to consider: 1. $A|B \cap C \sim A|C$ or 2. $B \cap C \in \mathcal{N}$. If $A \cap C \in \mathcal{N}$, then $B \perp\!\!\!\perp_C A$ by definition. So for the remainder of the proof we assume $A \cap C \notin \mathcal{N}$.

Case 1: Suppose $A|B \cap C \sim A|C$ (so that $B \cap C \notin \mathcal{N}$). Define:

$$\begin{aligned} &X = C \\ &Y = A \cap C && Y' = B \cap C \\ &Z = A \cap B \cap C \end{aligned}$$

Since $B \cap C \notin \mathcal{N}$, we can condition on it and reason as follows:

$$\begin{aligned} Z|Y' &= A \cap B \cap C | B \cap C \\ &\sim A | B \cap C \text{ by Axiom 4} \\ &\sim A|C \text{ by assumption of case} \\ &\sim A \cap |C \text{ by Axiom 4} \\ &= Y|X \end{aligned}$$

Lemma 8 then immediately entails that $Y'|X \sim Z|Y$ since $Y = A \cap C \notin \mathcal{N}$. But notice that Axiom 4 entails both that $Y'|X = B \cap C|C \sim B|C$ and that $Z|Y = A \cap B \cap C|A \cap C \sim B|A \cap C$. Thus, we obtain $B|C \sim B|A \cap C$, and $B \perp\!\!\!\perp_C A$ as desired.

Case 2: Suppose $B \cap C \in \mathcal{N}$. We show that (1) $B|C \sim \emptyset|C$ and (2) $B|A \cap C \sim \emptyset|C$. Thus, we obtain $B|C \sim B|A \cap C$, and $B \perp\!\!\!\perp_C A$ as desired. To show 1, note:

$$\begin{aligned} B|C &\sim B \cap C|C \text{ by Axiom 4} \\ &\sim \emptyset|C \text{ by Part 2 of Lemma 6 as } B \cap C \in \mathcal{N} \end{aligned}$$

To show 2, note that $A \cap C \notin \mathcal{N}$, and so we can condition on it. Thus:

$$\begin{aligned} B|A \cap C &\sim A \cap B \cap C|A \cap C \text{ by Axiom 4} \\ &\preceq \emptyset|A \cap C \text{ by Part 3 of Lemma 4 as } A \cap B \cap C \subseteq B \cap C \in \mathcal{N} \\ &\sim \emptyset|C \text{ by Part 2 of Lemma 6} \end{aligned}$$

□

Theorem 2. Let $\{C_\theta\}_{\theta \in \Theta}$ be events such that

1. $E|\theta \equiv F \cap C_\theta|\theta$ for all $\theta \in \Theta$,
2. $F \perp\!\!\!\perp_\theta C_\theta$ with respect to $\sqsubseteq$ for all $\theta \in \Theta$,
3. $C_\theta|\theta \equiv C_\eta|\eta$ for all $\theta, \eta \in \Theta$.
4. $\emptyset|\theta \sqsubset C_\theta|\theta$ for all θ.

If Θ is finite, then E and F are qualitatively posterior and favoring equivalent.

This theorem follows from the following three propositions.

Proposition 4. If conditions 1, 2, and 4 of Theorem 2 hold, then $E \in \mathcal{N}_{\preceq}$ if and only if $F \in \mathcal{N}_{\preceq}$.

Proposition 5. Suppose $E|\theta \sim F \cap C_\theta|\theta$ for all $\theta \in \Theta_s \subseteq \Theta$. If Θ_s is finite and $E, F \notin \mathcal{N}$, then for all $\theta \in \Theta_s$:

$$\theta|E \sim F \cap C_\theta \cap \theta| \bigcup_{\eta \in \Theta_s} F \cap C_\eta \cap \eta.$$

Proposition 6. Suppose that $F \perp\!\!\!\perp_\theta C_\theta$ with respect to $\preceq$, that $C_\theta|\theta \sim C_\eta|\eta$, and that $F \cap \theta \notin \mathcal{N}$ for all $\theta, \eta \in \Theta_F \subseteq \Theta$. If Θ_F is finite, then for all $\theta \in \Theta_F$:

$$\theta|F \sim F \cap C_\theta \cap \theta| \bigcup_{\eta \in \Theta_F} F \cap C_\eta \cap \eta.$$

Proof of Theorem 2. The first proposition shows that $\cdot|E$ is well-defined if and only if $\cdot|F$ is, so the first condition of qualitative posterior equivalence is satisfied. Let Θ_s be the set of $\theta \notin \mathcal{N}_\preceq$, and let Θ_F be the set of all $F \cap \theta \notin \mathcal{N}_\preceq$ (notice that $\Theta_F \subseteq \Theta_s$). We claim that $\theta|E \sim \theta|F$, for all $\theta \in \Theta$. We go by cases.

Case 1: $\theta \notin \Theta_s$. Then, $\theta \in \mathcal{N}$ by definition. This means that $\theta|E \sim \emptyset|E \sim \emptyset|F \sim \theta|F$.

Case 2: $\theta \in \Theta_s \setminus \Theta_F$. Then, $F \cap \theta \in \mathcal{N}$, so $\theta|F \sim \emptyset$. We claim that $\theta|E \sim \emptyset$ as well. To see why, note that, for all $\eta \in \Theta_s$, we get that $E|\eta \sim F \cap C_\eta|\eta$, from condition 1 of the theorem and Axiom 0. Thus, we can apply Proposition 5 to θ and Θ_s, yielding:

$$\theta|E \sim F \cap C_\theta \cap \theta| \bigcup_{\eta \in \Theta_s} F \cap C_\eta \cap \eta \tag{21}$$

Now, since $F \cap \theta \in \mathcal{N}$, it follows from Lemma 4 that $F \cap C_\theta \cap \theta \in \mathcal{N}$. So, $\theta|E \sim \emptyset$, completing this case.

Case 3: $\theta \in \Theta_F$. This implies that $\theta \in \Theta_s$ as well; thus, Equation 21 remains true. Now, for all $\theta_F, \eta_F \in \Theta_F$, we get that $F \cap \theta_F \notin \mathcal{N}$, that $F \perp\!\!\!\perp_{\theta_F} C_{\theta_F}$ with respect to $\preceq$, and that $C_{\theta_F}|\theta_F \sim C_{\eta_F}|\eta_F$. We thus apply Proposition 6 to θ and Θ_F, yielding:

$$\theta|F \sim F \cap C_\theta \cap \theta| \bigcup_{\eta \in \Theta_F} F \cap C_\eta \cap \eta \tag{22}$$

Now, for any $\theta_s \in \Theta_s \setminus \Theta_F$, we know that $F \cap C_{\theta_s} \cap \theta \in \mathcal{N}$, and thus applying Part 6 of Lemma 6 to Equation 22, we get $\theta|F \sim F \cap C_\theta \cap \theta| \bigcup_{\eta \in \Theta_s} F \cap C_\eta \cap \eta$. Along with Equation 21, this means that $\theta|E \sim \theta|F$.

Now, one can repeatedly apply Axiom 5 over the constituent simple hypotheses in any H to get $H|E \sim H|F$, showing that E and F are qualitative posterior equivalent. With Claim 4, we also get qualitative favoring equivalence. ☐

Before getting to the proofs of the propositions, we first prove the following corollary to Theorem 2.

Corollary 2. If Θ is finite and $E|\theta \equiv F|\theta$ for all $\theta \in \Theta$, then E and F are qualitatively posterior and favoring equivalent.

Proof. We simply apply Theorem 2 with $C_\theta = \Delta$ for all θ. To see that the conditions holds:

1. Note that, for any θ, $E|\theta \equiv F|\theta$ implies that $E|\theta \equiv F \cap \Delta|\theta$, since $F \cap \Delta = F$.

2. Similarly, $F \perp\!\!\!\perp_\theta \Delta$, since $F|\Delta \cap \theta \equiv F|\theta$, since $\Delta \cap \theta = \theta$.

3. We have to show $\Delta|\theta \equiv \Delta|\eta$ for all $\theta, \eta \in \Theta$. Note that by Axiom 4, $\Delta|\theta \sim \Delta \cap \theta|\theta = \theta|\theta$. By Axiom 3, $\theta|\theta \sim \eta|\eta$. Applying Axiom 4 again shows that $\eta|\eta \sim \Delta|\eta$, so by transitivity, $\Delta|\theta \sim \Delta|\eta$. By Axiom 0, we get that $\Delta|\theta \equiv \Delta|\eta$, as desired.

4. $\Delta|\theta \sqsupset \emptyset|\theta$ by Axiom 2.

From these conditions and the fact that Θ is finite, we get that E and F are qualitatively posterior and favoring equivalent. $\square$

B.4.1 Proof of Proposition 4

We prove two short lemmata first. The first is the qualitative analog of the fact that, if $P(B \cap C) > 0$, then $P(A \cap B|C) = 0$ if and only if $P(A|B \cap C) = 0$.

Lemma 13. Suppose $B \cap C \notin \mathcal{N}$. Then $A \cap B|C \sim \emptyset|D$ if and only if $A|B \cap C \sim \emptyset|D$ for all $D \notin \mathcal{N}$.

Proof. Since $B \cap C \notin \mathcal{N}$, we know $C \notin \mathcal{N}$ by Part 3 of Lemma 4. Thus:

$$\begin{aligned} A \cap B|C \sim \emptyset|D &\Leftrightarrow A \cap B|C \sim \emptyset|C && \text{by Part 2 of Lemma 6 since } C \notin \mathcal{N} \\ &\Leftrightarrow A \cap B \cap C|\Delta \sim \emptyset|\Delta && \text{by Lemma 7} \\ &\Leftrightarrow A|B \cap C \sim \emptyset|B \cap C && \text{by Lemma 7 since } B \cap C \notin \mathcal{N} \\ &\Leftrightarrow A|B \cap C \sim \emptyset|D && \text{by Part 2 of Lemma 6 since } D \notin \mathcal{N} \end{aligned}$$

$\square$

Lemma 14. Suppose $A \perp\!\!\!\perp_C B$ and $\emptyset|C \prec B|C$. Then $A \cap B|C \sim \emptyset|C$ if and only if $A|C \sim \emptyset|C$.

Proof. The right to left direction follows immediately from Part 3 of Lemma 4 since $A \cap B \subseteq A$.

In the left to right direction, first note that $B \cap C|C \sim B|C$ by Axiom 4 and that $B|C \succ \emptyset|C$ by assumption. Thus, $B \cap C|C \succ \emptyset|C$, from which it follows that $B \cap C \notin \mathcal{N}$. Thus:

$$\begin{aligned} A \cap B|C \sim \emptyset|C &\Leftrightarrow A|B \cap C \sim \emptyset|C && \text{by Lemma 13 since } B \cap C \notin \mathcal{N} \\ &\Leftrightarrow A|C \sim \emptyset|C && \text{since } A \perp\!\!\!\perp_C B \end{aligned}$$

$\square$

Proof of Proposition 4: Let's first show that if E is a member of $\mathcal{N}_{\preceq}$, then so is F. The proof of the converse is nearly identical because the lemmata we use are all biconditionals.

Suppose $E \in \mathcal{N}_{\preceq}$. Thus, $E \cap \{\theta\} \in \mathcal{N}$ for all θ by Part 3 of Lemma 4, as $E \cap \{\theta\} \subseteq E$. So by Axiom 2, it follows that $E \cap \{\theta\}|\Delta \sim \emptyset|\Delta$. Thus, for all θ such that $\{\theta\} \notin \mathcal{N}_{\preceq}$, we obtain that $E|\theta \sim \emptyset|\theta$ by Lemma 7.

Fix any arbitrary $\upsilon \in \Theta$ such that $\{\upsilon\} \notin \mathcal{N}_{\preceq}$. Below we'll show $F|\upsilon \sim \emptyset|\upsilon$. This will suffice to show $F \in \mathcal{N}_{\preceq}$ by the following reasoning. Since $F|\upsilon \sim \emptyset|\upsilon$, Lemma 7 entails that $F \cap \{\upsilon\}|\Delta \sim \emptyset|\Delta$. Since υ was arbitrary, it follows that $F \cap \{\theta\}|\Delta \sim \emptyset|\Delta$ for all θ such that $\{\theta\} \notin \mathcal{N}_{\preceq}$. Since Θ is finite by assumption, it follows from Axiom 5 that $F \cap H|\Delta \sim \emptyset|\Delta$ where $H = \{\theta \in \Theta : \{\theta\} \notin \mathcal{N}_{\preceq}\}$. Because Θ is finite, it's clear that $\neg H := \{\theta \in \Theta : \{\theta\} \in \mathcal{N}_{\preceq}\}$ is itself a member of $\mathcal{N}_{\preceq}$, and hence, $H \in \mathcal{S}$ is an almost-sure event. Thus, by Part 5 of Lemma 6, we obtain that $F \cap H|\Delta \sim F|\Delta$, and so $F|\Delta \sim \emptyset|\Delta$. Thus, by Axiom 2, $F \in \mathcal{N}_{\preceq}$ as desired.

Now we show that $F|\upsilon \sim \emptyset|\upsilon$ any arbitrary $\upsilon \in \Theta$ such that $\{\upsilon\} \notin \mathcal{N}_{\preceq}$, as we claimed. Above, we showed that $E|\upsilon \sim \emptyset|\upsilon$. From Axiom 0, it follows that $E|\upsilon \equiv \emptyset|\upsilon$. By the first assumption of the proposition, we know $E|\upsilon \equiv F \cap C_\upsilon|\upsilon$. Thus, $F \cap C_\upsilon|\upsilon \equiv \emptyset|\upsilon$. From our assumptions that $F \perp\!\!\!\perp_\upsilon C_\upsilon$ and the fact $C_\upsilon|\upsilon \sqsupset \emptyset|\upsilon$, Lemma 14 allows us to infer that $F|\upsilon \equiv \emptyset|\upsilon$. Since $\{\upsilon\} \notin \mathcal{N}_{\preceq}$, we can then apply Axiom 0 to conclude $F|\upsilon \sim \emptyset|\upsilon$, as desired.

The proof of the converse simply reverses the direction of all the inferences above. That is, if we assume $F \in \mathcal{N}_{\preceq}$, we can argue just as above that $F|\theta \sim \emptyset|\theta$ for all θ such that $\{\theta\} \notin \mathcal{N}$. From there, we use Lemma 14 to argue $F \cap C_\theta|\theta \sim \emptyset|\theta$ for all non-null θ. By the assumptions of the proposition, that entails $E|\theta \sim \emptyset|\theta$ for all non-null θ. By Lemma 7 and additivity, we can argue that $E|\Delta \sim \emptyset|\Delta$, and hence, $E \in \mathcal{N}_{\preceq}$ as desired. □

B.4.2 Proof of Proposition 5

The proof of Proposition 5 requires a few lemmata.

Lemma 15. Suppose $B_1, \ldots, B_n$ partition G. If $A_i|B_i \sim C_i|B_i$ for all $i \leq n$, then

$$\bigcup_{i \leq n} A_i \cap B_i|G \sim \bigcup_{i \leq n} C_i \cap B_i|G \text{ for all } i \leq n$$

Proof. We first use Lemma 7 on the premise, which yields $A_i \cap B_i|\Delta \sim C_i \cap B_i|\Delta$ for all $i \leq n$. Applying Axiom 5 n times yields $\bigcup_{i \leq n} A_i \cap B_i|\Delta \sim \bigcup_{i \leq n} C_i \cap B_i|\Delta$. Because $B_1, \ldots, B_n$ partition G, this is equivalent to $G \cap \bigcup_{i \leq n} A_i \cap B_i|\Delta \sim G \cap \bigcup_{i \leq n} C_i \cap B_i|\Delta$. Finally, we apply Lemma 7 again to get the desired result (note that $G \notin \mathcal{N}$ since $B_i \notin \mathcal{N}$ by the fact that $A_i|B_i \sim C_i|B_i$). □

The following lemma is the qualitative analog of the fact that, if $P(B) > 0$, then $P(A|B) = 0$ if and only if $P(A \cap B) = 0$. It is similar in spirit to Lemma 13.

Lemma 16. If $A|B \sim \emptyset|\Delta$, then $A \cap B|C \sim \emptyset|\Delta$ for all $C \notin \mathcal{N}$. The converse holds if $B \notin \mathcal{N}$.

Proof. In the left to right direction, suppose $A|B \sim \emptyset|\Delta$, and let $C \notin \mathcal{N}$ be arbitrary. We want to show $A \cap B|C \sim \emptyset|\Delta$. By Part 2 of Lemma 6, we know $\emptyset|\Delta \sim \emptyset|B$, and so $A|B \sim \emptyset|B$ by transitivity. Lemma 7 then entails that $A \cap B|\Delta \sim \emptyset|\Delta$, from which it follows that $A \cap B \in \mathcal{N}$. Hence, by Part 2 of Lemma 6, $A \cap B|C \sim \emptyset|\Delta$, as desired.

In the right to left direction, suppose $B \notin \mathcal{N}$ and that $A \cap B|C \sim \emptyset|\Delta$ for all $C \notin \mathcal{N}$. Then letting $C = \Delta$, we obtain that $A \cap B|\Delta \sim \emptyset|\Delta$. Thus, $A \cap B \in \mathcal{N}$, and Part 2 of Lemma 6 entails that $A \cap B|B \sim \emptyset|\Delta$. Axiom 4 then entails that $A|B \sim \emptyset|\Delta$, as desired. ☐

The next lemma outlines conditions under which we can "add" two qualitative equations by unioning the conditioning events. We will use this in the proof of Lemma 18.

Lemma 17. Suppose $Y \cap Y' = Z \cap Z' = \emptyset$ and that $X|Y \sim W|Z$ and $X|Y' \sim W|Z' \sim \emptyset|\Delta$. Then $X|Y \cup Y' \preceq W|Z \cup Z'$ if and only if $Y|Y \cup Y' \preceq Z|Z \cup Z'$.

Proof. Note that since $X|Y \sim W|Z$ and $X|Y' \sim W|Z'$, we know $Y, Y', Z, Z' \notin \mathcal{N}$ and can be conditioned on freely. Thus, by Part 3 of Lemma 4, it follows that $Y \cup Y'$ and $Z \cup Z'$ are non-null, and so we can condition on them freely as well. Now we proceed.

In the left to right direction, suppose $X|Y \cup Y' \preceq W|Z \cup Z'$. Suppose for the sake of contradiction that $Y|Y \cup Y' \npreceq Z|Z \cup Z'$. By Axiom 1, it follows that $Y|Y \cup Y' \succ Z|Z \cup Z'$.

Define:

$$\begin{aligned} A &= Y \cup Y', & A' &= Z \cup Z' \\ B &= Y & B' &= Z \\ C &= X \cap Y, & C' &= W \cap Z \end{aligned}$$

By Axiom 4, $C|B = X \cap Y|Y \sim X|Y$ and $C'|B' = W \cap Z|Z \sim W|Z$, and by assumption, $X|Y \sim W|Z$. Thus, $C|B \sim C'|B'$. Further, by assumption of the reductio, $B|A = Y|Y \cup Y' \succ W|Z \cup Z' = B'|A'$. So by Axiom 6b, it follows that $C|A \succ C'|A'$, i.e., that $X \cap Y|Y \cup Y' \succ W \cap Z|Z \cup Z'$. Further, because $Z \cap Z' = \emptyset$ and $W|Z' \sim \emptyset|\Delta$, it follows from Lemma 16 that $W \cap Z'|Z \cup Z' \sim \emptyset|\Delta$. Hence, $X \cap Y'|Y \cup Y' \succeq W \cap Z'|Z \cup Z'$ by Lemma 1. So we've shown that

$$\begin{gathered} X \cap Y|Y \cup Y' \succ W \cap Z|Z \cup Z', \text{ and} \\ X \cap Y'|Y \cup Y' \succeq W \cap Z'|Z \cup Z'. \end{gathered}$$

Because $Y \cap Y' = Z \cap Z' = \emptyset$, applying Axiom 5 to these equations yields $X \cap (Y \cup Y')|Y \cup Y' \succ W \cap (Z \cup Z')|Z \cup Z'$. Applying Axiom 4 to both sides of the equation yields $X|Y \cup Y' \succ W|Z \cup Z'$, which contradicts our assumption.

In the right to left direction, suppose $Y|Y \cup Y' \preceq Z|Z \cup Z'$. We must show $X|Y \cup Y' \preceq W|Z \cup Z'$. We first show that

$$X \cap Y|Y \cup Y' \preceq W \cap Z|Z \cup Z' \tag{23}$$

That argument will use the assumption that $X|Y \sim W|Z$. We then prove

$$X \cap Y'|Y \cup Y' \sim W \cap Z'|Z \cup Z' \tag{24}$$

That argument will use the assumption that $X|Y' \sim W|Z' \sim \emptyset|\Delta$. Because $Y \cap Y' = Z \cap Z' = \emptyset$ (and hence $(X \cap Y) \cap (X \cap Y') = (W \cap Z) \cap (W \cap Z') = \emptyset$), applying Axiom 5 to Equation 23 and Equation 24 yields

$$(X \cap Y) \cup (X \cap Y')|Y \cup Y' \preceq (W \cap Z) \cup (W \cap Z')|Z \cup Z'.$$

Applying Axiom 4 to both sides of that equation yields the desired conclusion that $X|Y \cup Y' \preceq W|Z \cup Z'$.

To prove Equation 23, define:

$$\begin{aligned} A &= Y \cup Y', & A' &= Z \cup Z' \\ B &= Y & B' &= Z \\ C &= X \cap Y, & C' &= W \cap Z \end{aligned}$$

By Axiom 4, $C|B = X \cap Y|Y \sim X|Y$ and $C'|B' = W \cap Z|Z \sim W|Z$. Since $X|Y \sim W|Z$ by assumption, we have $C|B \sim C'|B'$. And our assumption that $Y|Y \cup Y' \preceq Z|Z \cup Z'$ means that $B|A \preceq B'|A'$. From Axiom 6b, it follows that $C|A \preceq C'|A'$, i.e., that $X \cap Y|Y \cup Y' \preceq W \cap Z|Z \cup Z'$.

To prove Equation 24, notice that, since $X|Y' \sim \emptyset|\Delta$ by assumption, Lemma 16 immediately entails $X \cap Y'|Y \cup Y' \sim \emptyset|\Delta$. By the same reasoning, because $W|Z' \sim \emptyset|\Delta$ by assumption, it follows that $W \cap Z'|Z \cup Z' \sim \emptyset|\Delta$. By transitivity (specifically Axiom 1), $X \cap Y'|Y \cup Y' \sim W \cap Z'|Z \cup Z'$ as desired. □

Lemma 18. Suppose $B_1, \ldots, B_n$ partition G. Further, suppose $A|B_i \sim C_i|B_i$ for all $i \leq n$. If $A \cap G \notin \mathcal{N}$, then $B_i|A \cap G \sim B_i \cap C_i|\bigcup_{j \leq n} B_j \cap C_j$ for all $i \leq n$.

Lemma 18 can be thought of as an application of Bayes' theorem combined with the Law of Total Probability. If $C_i = C$ for all i, then quantitative analog of the lemma asserts the following: if $B_1 \ldots B_n$ partition G and $P(A|B_i) = P(C|B_i)$ for all i, then $P(B_i|A \cap G) = P(B_i|C \cap G)$ for all i. That claim follows easily from an application of Bayes' theorem and the Law of Total Probability.

Proof. We first prove that if $A \cap G \notin \mathcal{N}$, then $\bigcup_{j \leq n} B_j \cap C_j \notin \mathcal{N}$, and so the expressions in the statement of the lemma are well-defined.

Because $A \cap G = \bigcup_{i \leq n} A \cap B_i$, it follows from Part 4 of Lemma 6 that $A \cap B_i \notin \mathcal{N}$ for some $i \leq n$. Thus, $B_i \notin \mathcal{N}$ since $A \cap B_i \subseteq B_i$. Further, $A|B_i \succ \emptyset|B_i$, for otherwise, $A \cap B_i|\Delta \preceq \emptyset|\Delta$ by Lemma 7 and $A \cap B_i \in \mathcal{N}$, contradicting assumption. Finally, since $A|B_i \sim C_i|B_i$, it follows that $C_i|B_i \succ \emptyset|B_i$. It follows quickly, via the same sequence of lemmata that we just used, that $C_i \cap B_i \notin \mathcal{N}$, and hence, $\bigcup_{j \leq n} B_j \cap C_j \notin \mathcal{N}$ since $C_i \cap B_i \subseteq \bigcup_{j \leq n} B_j \cap C_j$.

Now we prove the lemma by induction on n. In the base case, when $n = 1$, $B_1 = G$. Then:

$$\begin{aligned} G|A \cap G &\sim A \cap G|A \cap G && \text{by Axiom 4} \\ &\sim C_1 \cap G|C_1 \cap G && \text{by Axiom 3} \end{aligned}$$

which is the desired equality.

For the inductive step, assume $B_1, \ldots, B_{n+1}$ partition G and $A|B_i \sim C_i|B_i$ for all $i \le n+1$. Let $G' = G \setminus B_{n+1}$. Clearly, $B_1, \ldots, B_n$ partition G' and $A|B_i \sim C_i|B_i$ for all $i \le n$. We consider three cases: either one of $A \cap G'$ or $A \cap B_{n+1}$ are not members of $\mathcal{N}$ or neither of them are (since $A \cap G \notin \mathcal{N}$, they cannot both be members of $\mathcal{N}$).

Case 1: $A \cap G', A \cap B_{n+1} \notin \mathcal{N}$ We handle the difficult case first. By the inductive hypothesis:

$$B_i|A \cap G' \sim B_i \cap C_i| \bigcup_{j \le n} B_j \cap C_j \text{ for all } i \le n \tag{25}$$

Equation 25 also holds for $i = n+1$ because both $B_{n+1} \cap (A \cap G')$ and $(B_{n+1} \cap C_{n+1}) \cap \bigcup_{j \le n} B_j \cap C_j$ are empty. Thus, we can use Axiom 4 and Lemma 2 to get Equation 25. Further, it's easy to show that for all $i \le n+1$:

$$B_i|A \cap B_{n+1} \sim B_i \cap C_i|B_{n+1} \cap C_{n+1} \tag{26}$$

If $i = n+1$, one can simply use Axiom 4 to turn the left side into the sure event and then use Axiom 3, just like in the base case. If $i \ne n+1$, then note that $B_i \cap (A \cap B_{n+1}) = (B_i \cap C_i) \cap (B_i \cap C_{n+1}) = \emptyset$. So like above, we use Axiom 4 and Lemma 2 to get the equality. Our goal is to show:

$$B_i|A \cap G \sim B_i \cap C_i| \bigcup_{j \le n+1} B_j \cap C_j \text{ for all } i \le n+1 \tag{27}$$

Equation 27 would follow if we could union the conditioning events in Equation 25 and Equation 26. Lemma 17 was designed for that purpose. Let i be arbitrary and define:

$$\begin{aligned} X &= B_i & W &= B_i \cap C_i \\ Y &= A \cap B_{n+1} & Z &= B_{n+1} \cap C_{n+1} \\ Y' &= A \cap G & Z' &= \bigcup_{k \le n} B_k \cap C_k \end{aligned}$$

We want to apply Lemma 17, so we first check all the conditions. Note that $Y \cap Y' = Z \cap Z' = \emptyset$, since the B_i's are disjoint. Further, $X|Y \sim W|Z$ is simply Equation 26 and $X|Y' \sim W|Z'$ is Equation 25. Lastly, if $i = n+1$, $X|Y' = B_i|A \cap G' \sim \emptyset|\Delta$ by the disjointness of B_{n+1} and G', along with Axiom 4, and Lemma 2. If $i \le n$, $X|Y = B_i|A \cap B_{n+1} \sim \emptyset|\Delta$, again because B_i and B_{n+1} are disjoint. In this case, we can simply swap Y with Y' and Z with Z' when applying the lemma. With the conditions satisfied, Lemma 17 tells us that Equation 27 holds if and only if $Y|Y \cup Y' \sim Z|Z \cup Z'$ i.e.:

$$A \cap B_{n+1}|A \cap G \sim B_{n+1} \cap C_{n+1}| \bigcup_{k \le n+1} B_k \cap C_k \tag{28}$$

To prove Equation 28, we first show:

$$A \cap B_{n+1}|G \sim B_{n+1} \cap C_{n+1}|G \tag{29}$$

Define:

$$\begin{aligned} X &= G & X' &= G \\ Y &= B_{n+1} & Y' &= B_{n+1} \\ Z &= A \cap B_{n+1} & Z' &= B_{n+1} \cap C_{n+1} \end{aligned}$$

Clearly, $Y|X \sim Y'|X'$, and $Z|Y = A \cap B_{n+1}|B_{n+1} \sim A|B_{n+1}$ by Axiom 4. Similarly, $Z'|Y' \sim C_{n+1}|B_{n+1}$. By the assumption that $A|B_i \sim C_i|B_i$ for all $i \leq n+1$, we get $Z|Y \sim Z'|Y'$. So Axiom 6b tells us that $Z|X \sim Z'|X'$, which is Equation 29.

We now prove Equation 28 by contradiction. Suppose $A \cap B_{n+1}|A \cap G \not\sim B_{n+1} \cap C_{n+1}|\bigcup_{k \leq n+1} B_k \cap C_k$, and without loss of generality, say $A \cap B_{n+1}|A \cap G \succ B_{n+1} \cap C_{n+1}|\bigcup_{k \leq n+1} B_k \cap C_k$. Now, define:

$$\begin{aligned} X &= G & X' &= G \\ Y &= A \cap G & Y' &= \bigcup_{k \leq n+1} B_k \cap C_k \\ Z &= A \cap B_{n+1} & Z' &= B_{n+1} \cap C_{n+1} \end{aligned}$$

By Lemma 15 (letting $A_i = A$ for all i), we know that $A \cap G|G \sim \bigcup_{k \leq n+1} B_k \cap C_k|G$, i.e. $Y|X \sim Y'|X'$. Note that $Z|Y \succ Z'|Y'$ by assumption. Since this shows that $Z \notin \mathcal{N}$, Axiom 6b entails that $Z|X \succ Z'|X'$, contradicting Equation 29.

Case 2: $A \cap G' \in \mathcal{N}, A \cap B_{n+1} \notin \mathcal{N}$ Notice that since $A \cap G' \in \mathcal{N}$, $A \cap B_i \in \mathcal{N}$, for all $i \leq n$, by Lemma 6. Thus, $B_i \cap C_i \in \mathcal{N}$, by our assumption. So, for all $i \leq n$, our goal is to show that $B_i|A \cap G \sim \emptyset$. Since $A \cap B_i \in \mathcal{N}$, this follows by Axiom 4. All that remains is to show that $B_{n+1}|A \cap G \sim B_{n+1} \cap C_{n+1}|\bigcup_{j \leq n} B_j \cap C_j$. Notice that Equation 26 still holds in this case. We simply apply Part 6 of Lemma 6 to both sides of this equation, adding in $A \cap G'$ to the left hand side and $\bigcup_{j \leq n} B_i \cap C_i$ to the right hand side, since those events are null.

Case 3: $A \cap B_{n+1} \in \mathcal{N}, A \cap G' \notin \mathcal{N}$ Since $A \cap B_{n+1} \in \mathcal{N}$, $B_{n+1} \cap C_{n+1} \in \mathcal{N}$, by our assumption. So, $B_{n+1}|A \cap G \sim B_{n+1} \cap C_{n+1}|\bigcup_{j \leq n} B_j \cap C_j$ holds since both sides are equiprobable with the emptyset. Further, Equation 25, the inductive hypothesis, still holds in this case. Just as before, we simply apply Part 6 of Lemma 6 to both sides of this equation, adding in $A \cap B_{n+1}$ to the left hand side and $B_{n+1} \cap C_{n+1}$ to the right hand side. $\square$

With these lemmas, the proof of Proposition 5 is very short.

Proof of Proposition 5. We apply Lemma 18. Let $G = \Theta_s = \{\theta_1, \ldots, \theta_n\}$ and $B_i = \theta_i$. Let $A = E$ and $C_i = F \cap C_{\theta_i}$. Note that $E \notin \mathcal{N}$ and Θ_s is a sure event; thus, by Lemma 5, $E \cap \Theta_s = A \cap G \notin \mathcal{N}$. Further, $A|B_i \sim C_i|B_i$. Thus, the premises of the lemma are satisfied, and $B_i|A \sim B_i \cap C_i|\bigcup_{j\leq n} B_j \cap C_j$ is the desired conclusion. □

B.4.3 Proving Proposition 6

The proof of Proposition 6 requires the following lemmata.

Lemma 19. If $X|Z \preceq \emptyset|Z$ and $Y \cap Z \notin \mathcal{N}$, then $X|Y \cap Z \sim \emptyset|Y \cap Z$.

Proof. Because $X|Z \sim \emptyset|Z$, Lemma 7 entails that $X \cap Z|\Delta \preceq \emptyset \cap Z|\Delta = \emptyset|\Delta$. By Lemma 3, we know that $X \cap Y \cap Z|\Delta \preceq X \cap Z|\Delta$ because $X \cap Y \cap Z \subseteq X \cap Z$. By transitivity, $X \cap Y \cap Z|\Delta \preceq \emptyset|\Delta$. We can then apply Lemma 7 (since $Y \cap Z \notin \mathcal{N}$) to obtain $X|Y \cap Z \preceq \emptyset|Y \cap Z$. Also, by Lemma 1, $X|Y \cap Z \succeq \emptyset|Y \cap Z$. Thus, we get $X|Y \cap Z \sim \emptyset|Y \cap Z$ as desired. □

Lemma 20. Suppose $X = X_1 \cup X_2$, where $X_1 \cap X_2 = \emptyset$. Suppose $X_1|Y \cap G \cap Z \sim X_2|Y \cap (G \setminus Z)$. Further, suppose that both

i. $X_1|(G \setminus Z) \sim \emptyset|Z$, and

ii. $X_2|(G \setminus Z) \sim \emptyset|Z$.

Then $X|Y \cap G \sim X_1|Y \cap G \cap Z \sim X_2|Y \cap (G \setminus Z)$.

Proof. First, suppose that $X \cap Y \cap G \in \mathcal{N}$. Then, $X|Y \cap G \sim \emptyset|Y \cap G$. We also claim that $X_1|Y \cap G \cap Z \sim \emptyset|Y \cap G \cap Z$, which gives us the desired result by Lemma 2. Suppose this wasn't the case – since $Y \cap G \cap Z \notin \mathcal{N}$, we know that $X_1|Y \cap G \cap Z \succ \emptyset|Y \cap G \cap Z$. We can use Axiom 5 to add this with $X_2|Y \cap G \cap Z \succeq \emptyset|Y \cap G \cap Z$ (Lemma 1), getting $X|Y \cap G \cap Z \succ \emptyset|Y \cap G \cap Z$. Note that by Lemma 6, that since $X \cap Y \cap G \in \mathcal{N}$ we know that $X|Y \cap G \cap Z \sim \emptyset|Y \cap G \cap Z$, we thus get the desired contradiction.

We know assume $X \cap Y \cap G \notin \mathcal{N}$. By applying Lemma 19 (since $Y \cap (G \setminus Z) \notin \mathcal{N}$) and Lemma 2 to (i), we get that $X_1|Y \cap (G \setminus Z) \sim \emptyset|Y \cap G \cap Z$. Now, we apply Axiom 5 as follows:

$$\emptyset|Y \cap G \cap Z \sim X_1|Y \cap (G \setminus Z) \tag{30}$$

$$X_1|Y \cap G \cap Z \sim X_2|Y \cap (G \setminus Z) \qquad \text{Given} \tag{31}$$

$$X_1|Y \cap G \cap Z \sim X_1 \cup X_2|Y \cap (G \setminus Z) \qquad \text{Axiom 5 on Equation 30 and Equation 31} \tag{32}$$

We apply the same steps to (ii). Using Lemma 19 (since $Y \cap G \cap Z \notin \mathcal{N}$) and

Lemma 2, we get $X_2|Y \cap G \cap Z \sim \emptyset|Y \cap (G \setminus Z)$. Then, we apply Axiom 5:

$$X_2|Y \cap G \cap Z \sim \emptyset|Y \cap (G \setminus Z) \tag{33}$$

$$X_1|Y \cap G \cap Z \sim X_2|Y \cap (G \setminus Z) \quad \text{Given} \tag{34}$$

$$X_1 \cup X_2|Y \cap G \cap Z \sim X_2|Y \cap (G \setminus Z) \quad \text{Axiom 5 on Equation 33 and Equation 34} \tag{35}$$

Because $X_1|Y \cap G \cap Z \sim X_2|Y \cap (G \setminus Z)$, from Equation 32 and Equation 35 we get $X|Y \cap G \cap Z \sim X|Y \cap (G \setminus Z)$.

Now, suppose for the sake of contradiction that $X|Y \cap G \not\sim X|Y \cap G \cap Z$. We know that $Y \cap G \cap Z \notin \mathcal{N}$ by the premise, which also means (via Lemma 4) that $Y \cap G \notin \mathcal{N}$. Thus, either $X|Y \cap G \succ X|Y \cap G \cap Z$ or vice versa; without loss of generality, assume $X|Y \cap G \succ X|Y \cap G \cap Z$. Now, we apply Lemma 8 (since we are assuming $X \cap Y \cap G \notin \mathcal{N}$) with:

$$\begin{aligned} A &= Y \cap G \\ B &= X \cap Y \cap G \qquad\qquad B' = Y \cap G \cap Z \\ C &= X \cap Y \cap G \cap Z \end{aligned}$$

Note that the necessary containment relations hold. Since $B|A \succ C|B'$, we get $B'|A \succ C|B$, or $Y \cap G \cap Z|Y \cap G \succ X \cap Y \cap G \cap Z|X \cap Y \cap G$. After applying Axiom 4, we get:

$$Z|Y \cap G \succ Z|X \cap Y \cap G \tag{36}$$

Since $X|Y \cap G \cap Z \sim X|Y \cap (G \setminus Z)$, we know from our contradicting assumption that $X|Y \cap G \succ X|Y \cap (G \setminus Z)$. We now proceed as before, applying Lemma 8 with:

$$\begin{aligned} A &= Y \cap G \\ B &= X \cap Y \cap G \qquad\qquad B' = Y \cap (G \setminus Z) \\ C &= X \cap Y \cap (G \setminus Z) \end{aligned}$$

Note that the necessary containment relations hold. Since $B|A \succ C|B'$, we get $B'|A \succ C|B$, or (after Axiom 4):

$$G \setminus Z|Y \cap G \succ G \setminus Z|X \cap Y \cap G \tag{37}$$

Applying Axiom 5 to Equation 36 and Equation 37 gives us $G|Y \cap G \succ G|X \cap Y \cap G$. Applying Axiom 4 gives us $\Delta \succ \Delta$, which is a contradiction. □

Lemma 21. Suppose $B_1, \ldots, B_n$ partition G, and that $A \cap B_i \notin \mathcal{N}$, for all i. Further, suppose $A \perp\!\!\!\perp_{B_i} C_i$ with respect to $\preceq$ and for all $i, j \leq n : C_i|B_i \sim C_j|B_j$. Then, for all $k \leq n$,

$$\bigcup_{i \leq n} C_i \cap B_i|A \cap G \sim C_k|B_k$$

Proof. Since $A \perp\!\!\!\perp_{B_i} C_i$ for all $i \leq n$, we have by Lemma 12 that $C_i \perp\!\!\!\perp_{B_i} A$. We now proceed by induction on n. When $n = 1$, note that $G = B_1$. So:

$$\begin{aligned}
\bigcup_{i \leq n} C_1 \cap B_i | A \cap G &= C_1 \cap G | A \cap G \\
&\sim C_1 | A \cap G \text{ by Axiom 4} \\
&\sim C_1 | G \text{ because } C_1 \perp\!\!\!\perp_G A
\end{aligned}$$

For the inductive step, suppose $B_1, \ldots, B_{n+1}$ partition G, and define $G' = \{B_1, \ldots, B_n\}$. Let i be arbitrary.

Note that the conditions of the inductive hypothesis are satisfied with respect to G' and $B_1, \ldots, B_n$. Thus, we get:

$$\bigcup_{j \leq n} C_j \cap B_j | A \cap G' \sim C_i | B_i$$

We also know that:

$$\begin{aligned}
C_{n+1} \cap B_{n+1} | A \cap B_{n+1} &\sim C_{n+1} | A \cap B_{n+1} && \text{by Axiom 4} \\
&\sim C_{n+1} | B_{n+1} && \text{since } A \perp\!\!\!\perp_{B_i} C_i \text{ for all } i \\
&\sim C_i | B_i && \text{since } C_i | B_i \sim C_j | B_j \text{ for all } i, j \leq n+1
\end{aligned}$$

So we've shown that (1) $\bigcup_{j \leq n} C_j \cap B_j | A \cap G' \sim C_i | B_i$ and (2) $C_{n+1} \cap B_{n+1} | A \cap B_{n+1} \sim C_i | B_i$. We now apply Lemma 20 with $X_1 = \bigcup_{j \leq n} C_j \cap B_j$, $X_2 = C_{n+1} \cap B_{n+1}$, $Y = A$, $Z = G'$ and $G = G$. We have shown (from (1) and (2)) that $X_1 | Y \cap G \cap Z \sim X_2 | Y \cap (G \setminus Z)$. Note that $X_1 | (G \setminus Z) \sim \emptyset | Z$ by Axiom 4, as $X_1 \cap (G \setminus Z) = \emptyset$. Similarly, $X_2 | G \cap Z \sim \emptyset$. With the conditions satisfied, we get $X | Y \cap G \sim X_1 | Y \cap G \cap Z$ or $\bigcup_{j \leq n+1} C_j \cap B_j | A \cap G \sim \bigcup_{j \leq n} C_j \cap B_j | A \cap G'$. And from (1), we get $\bigcup_{j \leq n+1} C_j \cap B_j | A \cap G \sim C_i | B_i$, as desired. □

Proof of Proposition 6. Let $\Theta_F = \{\theta_1, \ldots, \theta_n\}$ and let $B_i = \{\theta_i\}$. Thus, $G = \Theta_F$. Let $A = F$ and $C_i = C_{\theta_i}$. The conditions of Lemma 21 are satisfied by the assumptions of the proposition, and thus $\bigcup_{\eta \in \Theta_F} C_\eta \cap \eta | F \sim C_\theta | \theta$ for all $\theta \in \Theta_F$. Since $F \perp\!\!\!\perp_\theta C_\theta$, it follows from Lemma 12 that $C_\theta \perp\!\!\!\perp_\theta F$ and so $C_\theta | \theta \sim C_\theta | F \cap \theta$. Thus, we've shown:

$$\bigcup_{\eta \in \Theta_F} C_\eta \cap \eta | F \sim C_\theta | F \cap \theta \tag{38}$$

Now, we apply Lemma 8 with:

$$\begin{aligned}
&A = F \\
&B = \bigcup_{\eta \in \Theta_F} F \cap C_\eta \cap \eta \qquad B' = F \cap \theta \\
&C = F \cap C_\theta \cap \theta
\end{aligned}$$

Equation 38 says $B|A \sim C|B'$. So by Lemma 8, we get $B'|A \sim C|B$, which is the desired conclusion. □

References

Shiri Alon and Ehud Lehrer. Subjective multi-prior probability: A representation of a partial likelihood relation. *Journal of Economic Theory*, 151: 476–492, 2014.

James O. Berger. Robust Bayesian analysis: sensitivity to the prior. *Journal of statistical planning and inference*, 25(3):303–328, 1990. Publisher: Elsevier.

James Orvis Berger and Robert L. Wolpert. The likelihood principle. In *Ims Lecture Notes-Monograph*, volume 6. Institute of Mathematical Statistics, 1988.

David R. Bickel. The strength of statistical evidence for composite hypotheses: Inference to the best explanation. *Statistica Sinica*, pages 1147–1198, 2012.

Allan Birnbaum. On the foundations of statistical inference. *Journal of the American Statistical Association*, 57(298):269–306, 1962.

Anthony William Fairbank Edwards. *Likelihood*. CUP Archive, 1984.

Wade Edwards, Harold Lindman, and Leonard J. Savage. Bayesian Statistical Inference for Psychological Research. In Joseph B. Kadane, editor, *Robustness of Bayesian Analysis*, volume 4 of *Studies in Bayesian Econometrics*, pages 1–60. Elsevier Science Publishing Company Inc., Amsterdam: North Holland, 1984.

Michael J. Evans, Donald AS Fraser, and Georges Monette. On principles and arguments to likelihood. *Canadian Journal of Statistics*, 14(3):181–194, 1986.

D. A. S. Fraser, G. Monette, and K. W. Ng. Marginalization, likelihood and structured models. *Multivariate Analysis*, 6:209–217, 1984. Publisher: North-Holland Amsterdam.

Andrew Gelman, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. *Bayesian data analysis*. CRC press, 2013.

Ian Hacking. *Logic of statistical inference*. Cambridge University Press, 1965.

Joseph B. Kadane, editor. *Robustness of Bayesian Analysis*, volume 4 of *Studies in Bayesian Econometrics*. Elsevier Science Publishing Company Inc., Amsterdam: North Holland, 1984.

Charles H. Kraft, John W. Pratt, and Abraham Seidenberg. Intuitive probability on finite sets. *The Annals of Mathematical Statistics*, pages 408–419, 1959.

David H. Krantz, R. Duncan Luce, Patrick Suppes, and Amos Tversky. *Foundations of Measurement Volume I: Additive and Polynomial Representations*. Dover Publications, Mineola, NY, December 2006a. ISBN 978-0-486-45314-9.

David H. Krantz, R. Duncan Luce, Patrick Suppes, and Amos Tversky. *Foundations of Measurement Volume II: Geometrical, Threshold, and Probabilistic Representations*. Dover Publications, Mineola, N.Y, December 2006b. ISBN 978-0-486-45315-6.

Howard Raiffa and Robert Schlaifer. *Applied statistical decision theory*. Wiley Classics Library. Wiley-Interscience, 1961. Publisher: Division of Research, Graduate School of Business Adminitration, Harvard ....

Richard Royall. *Statistical evidence: a likelihood paradigm*, volume 71. Chapman & Hall/CRC, 1997.

Leonard J. Savage, George Barnard, Jerome Cornfield, Irwin Bross, I. J. Good, D. V. Lindley, C. W. Clunies-Ross, John W. Pratt, Howard Levene, and Thomas Goldman. On the foundations of statistical inference: Discussion. *Journal of the American Statistical Association*, 57(298):307–326, 1962. Publisher: JSTOR.

Elliott Sober. *Evidence and evolution: The logic behind the science*. Cambridge University Press, 2008.

Peter Walley. *Statistical reasoning with imprecise probabilities*. Number 42 in Monographs on Statistics and Applied Probability. Chapman & Hall, New York, 1991.

Sérgio Wechsler, Carlos A. de B. Pereira, Paulo C. Marques F., Marcelo de Souza Lauretto, Carlos Alberto de Bragança Pereira, and Julio Michael Stern. Birnbaum's Theorem Redux. *AIP Conference Proceedings*, 1073(1):96–100, November 2008. ISSN 0094-243X. doi: 10.1063/1.3039028. URL https://aip.scitation.org/doi/abs/10.1063/1.3039028. Publisher: American Institute of Physics.