An Axiomatic Theory of
Inductive Inference

Luciano Pomatto and Alvaro Sandroni*†

This article develops an axiomatic theory of induction that speaks to the recent debate on Bayesian orgulity. It shows the exact principles associated with the belief that data can corroborate universal laws. We identify two types of disbelief about induction: skepticism that the existence of universal laws of nature can be determined empirically, and skepticism that the true law of nature, if it exists, can be successfully identified. We formalize and characterize these two dispositions toward induction by introducing novel axioms for subjective probabilities. We also relate these dispositions to the (controversial) axiom of σ-additivity.

1. Introduction. We seek an axiomatic understanding of specific problems of induction. Informally, induction is taken to mean the process of using empirical evidence to validate general claims, and, for our purposes, it is critical to differentiate between two types of epistemic skepticism about induction.[1]

One may doubt it is possible to know whether nature abides by any law.[1] Any empirical regularity may be a temporary fluke. Hence, patterns can suggest, but not prove, the existence of universal laws. So, one may ascribe nonvanishing odds to the idea that nature does not follow any law, no matter how numerous and consistent the data may grow to be. We refer to this disposition as *Humean skepticism*, with the caveat that we do not claim to provide a complete representation of Hume's (and other authors') actual statements.

1. The definition of "law" is subjective, as we make clear in the main text.

In addition, even if it is taken for granted that nature abides by a law, one may be skeptical that such a law can be inferred with arbitrarily high precision, even when the data grow without bounds. Let us say that in each period either 0 or 1 must occur and that 1 has been observed every period, over a long time, say t periods. The data are consistent with the law "nature produces only 1" and with the law "nature produces 1 until period t and 0 afterward," among (infinitely) many other laws. So, one may maintain a nonvanishing doubt that empirical evidence can validate a specific law, even under the assumption that the data follow one. We refer to this form of skepticism as *Goodman's skepticism*, with the same caveat as above.

We consider a probabilistic framework in which an agent, named Bob, is endowed with a *coherent view of the world* (i.e., a finitely additive probability measure) over *paths* (i.e., infinite binary sequences). As the data unfold, Bob updates his view of the world through Bayes's rule.[2] No restrictions are placed on which paths may be produced. So, no relationship between past and future is, a priori, required (apart from the idea that either 0 or 1 occurs each period). Bob is not dogmatic about induction either. He believes that the data may or may not follow eternal laws.

Under the lenses of this formal framework, we formalize Hume's and Goodman's skepticisms by introducing two novel axioms for subjective probabilities. These axioms refer to Bob's belief as the data unfold and become arbitrarily numerous. Bob's view of the world is *inductive in the sense of Hume*, as we define it, if, under data compatible with laws, Bob expects to become almost convinced that nature indeed follows laws. This axiom rules out Humean skepticism about induction. Bob's view of the world is *inductive in the sense of Goodman* if he expects to successfully identify nature's law up to a vanishing degree of error, conditional on nature abiding by one. This axiom rules out Goodman's skepticism about induction.

A natural starting line of inquiry is the extent of the connection between the two problems of induction. We start by asking whether Goodman's skepticism implies Hume's skepticism and the converse implication. Neither is true. Some coherent views exhibit Goodman's skepticism but not Hume's skepticism, and, conversely, some coherent views exhibit Hume's skepticism but not Goodman's skepticism. Thus, these two types of skepticism are not logically nested.

Of particular interest are the coherent views that express skepticism in the sense of Goodman but not in the sense of Hume. If, say, confronted with the question of whether the data are generated by a Turing machine, such

2. This framework follows de Finetti's (1970) viewpoint that inference involves personal judgments of likelihood that must be formalized in a coherent way. See de Finetti (1970) for a connection between coherent views of the world and Dutch books. We make no original attempt to justify Bayesianism and subjectivism.

views of the world express conviction that with enough data it is possible to make this determination with near certainty. In spite of this remarkable confidence on the capacity of Bayes's rule to address this apparently insurmountable inference problem, the same view of the world, if confronted with the (arguably simpler) question of which Turing machine generates the data, assuming that one does, remains skeptical that this determination can be made with arbitrarily high precision.

The celebrated theorems of Lévy (1937), Doob (1949), and Blackwell and Dubins (1962) make clear that under σ-additivity a Bayesian must believe that his opinion about a given hypothesis will converge to the truth. In particular, σ-additivity excludes both Hume's and Goodman's skepticism, and therefore it implies a form of "Bayesian orgulity" (Belot 2013). Different results were obtained by Juhl and Kelly (1994), Kelly (1996), and Elga (2016), among others, who have shown that there exist non-σ-additive coherent views of the world that allow for Humean skepticism. Hence, in the absence of σ-additivity, epistemic skepticism is allowed.

Our results reveal a complex relationship between skepticism and subjective probability. There are non-σ-additive coherent views that rule out Humean skepticism and others yet that rule out Goodman's skepticism. Thus, the spectrum of coherent views is rich enough to allow, at the same time, both orgulity and skepticism about induction. In particular, in the absence of σ-additivity, orgulity and skepticism are allowed. Orgulity is not an exclusive property of σ-additivity and may hold with or without it. This is a difficulty for a clear-cut theory of induction that seeks the root causes of orgulity and skepticism about induction.

We show that while a lack of σ-additivity does not assure Hume's skepticism and Goodman's skepticism, it always assures skepticism in at least one of these two ways. This is demonstrated by the structure theorem for coherent views of the world. It shows that a coherent view is inductive in the sense of Hume and in the sense of Goodman if and only if it is σ-additive. Thus, σ-additivity is *the* definitive condition that assumes away both Hume's and Goodman's skepticism about induction. It is not necessary to rule out either type of skepticism, but it is required to rule out both types simultaneously.

The interpretation of the structure theorem requires considerable care. The equivalence between induction and σ-additivity may suggest that the problems of how to conceptually justify either induction or σ-additivity are, in fact, one and the same problem and that σ-additivity is the root and only cause of the conviction in the ultimate success of induction. This reading of the structure theorem may prove incomplete. Consider an alternative approach in which the focus is not on using data to ultimately (i.e., in the limit as data grow) uncover eternal laws of nature but on making predictions within a practical (i.e., bounded) future. Consider the case in which a long sequence of 1s has been observed. One may wonder whether "nature produces only

1s." One may also wonder whether "nature will produce only 1s for the next 1,000 periods." Our last result concerns the latter case, in which Bob remains agnostic about the validity of universal claims but asks whether regularities in the past can be used to make sharp predictions about a bounded future. This result shows that any coherent view of the world, no matter how it is formed, must be confident that multiple repetitions of Bayes's rule transform pattern data into a near infallible guide to a bounded future.

Moreover, after enough data there must be high confidence on limited, but correct, inductive inferences. This holds even if, a priori, no assumption is made on the relationship between past and future in the sense that the data may unfold according to any path, including those without patterns. It also holds even if Bob is a skeptic in regard to the use of data to ultimately validate specific or general laws. Eventually there must be high confidence that the past is a limited, but successful, guide to the future. This conclusion follows from conditional probability alone and holds for any coherent view of the world. Thus, some confidence in inductive inference follows from coherence.

The article speaks to the recent debate on "Bayesian orgulity," originated with Belot (2013, 2017). Central to Belot's thesis is the argument that the convergence results of Levy, Doob, and Blackwell and Dubins are proof that Bayesianism implies epistemic arrogance. The debate has spurred different views. Huttegger (2015) argued that the issue of convergence to the truth should be put in the context of a long but finite horizon. Weatherson (2015) revisited Belot's argument from the perspective of Bayesian imprecise probability. The work closest to this article is Elga (2016), who showed the existence of non-σ-additive subjective probabilities expressing epistemic humility.

This article is also connected to the work of Kelly (1996), who formalized the connection between inductive inference and finitely additive probabilities, to the work of Gilboa and Samuelson (2012), who analyzed how subjectivity can enhance inductive inference, and to Al-Najjar, Pomatto, and Sandroni (2014), who study how different dispositions toward induction can affect incentive problems.

2. Basic Concepts and Results

2.1. Patterns and Coherence. An agent named Bob observes, in every period, one of two possible outcomes, 0 or 1. The set $\Omega = \{0, 1\}^\infty$ is the set of all *paths* or infinite histories of outcomes. Given a path ω and a time t, we denote by ω^t the set of paths that share with ω the same first t outcomes. We call ω^t a *finite history*. We fix an algebra Σ of subsets of Ω (subsets of Ω mentioned in the text belong to Σ, even when not stated explicitly). The agent is

endowed with a finitely additive probability P on Σ.[3] The measure P captures Bob's subjective viewpoint on how the outcomes will evolve. We refer to P as a *coherent view of the world*.

Some paths are governed by a law or pattern, and some are not. For instance, the path $1_\infty = (1, 1, 1, ...)$ follows the law "nature produces only the outcome 1." A classic example of a pattern is given by *periodic* paths, defined by repeated cycles as in $(1, 0, 1, 0, ...)$ or $(1, 1, 0, 0, 1, 1, ...)$, or, more generally, *eventually periodic* paths (i.e., sequences that are periodic after some point in time). Both examples are subsumed by the class of *computable* paths, which consists of all sequences that can be generated by a Turing machine (i.e., all paths that are the output of some finite program running on a computer with unlimited storage).

In order to speak of induction it is critical to demarcate between paths governed by a law and paths that do not follow any discernible pattern. This distinction can be made in many different ways, and the precise way in which this determination should be made is orthogonal to the central questions in this article. So, we need not take a definitive stance of this matter. Instead, we assume that the final determination of what constitutes a law is subjective. That is, Bob determines which set of paths $A \subseteq \Omega$ are the ones that abide by a law. The complement of A is the set of paths that according to Bob do not follow any pattern. For simplicity, we often refer to paths in A as *laws* and to paths not in A as *nonlaws*.

We make the following assumptions on A and P.

Assumption 1. A is countable.

While flexible enough to capture many formal definitions of pattern, including the set of periodic, eventually periodic, or computable paths, the assumption is not, however, without loss of generality. It greatly simplifies the analysis because it rules out both conceptual and technical difficulties that are outside the scope of this article. The main implication of assumption 1 is that it allows a view of the world to assign strictly positive probability to each law-like path. If, for example, A was uncountable, then Bob would have to assign zero probability to most individual laws. Formally:

Remark 1. For any coherent view of the world, there can be, at most, countably many paths with strictly positive probability.

3. That is, a function $P : \Sigma \to [0, 1]$ such that $P(\Omega) = 1$ and for every pair of disjoint sets E_1 and E_2 in Σ it satisfies $P(E_1 \cup E_2) = P(E_1) + P(E_2)$.

The result applies to Bob's view of the world both before and, by Bayes's rule, after the data are observed.

An alternative approach, which allows capturing more complex inference problems, is to consider nondeterministic laws. In section 7 we discuss this alternative approach and, in particular, the difficulties it involves.

Assumption 2. $P(\{\omega\}) > 0$ for every $\omega \in A$, and $P(A^c) > 0$.

Bob believes that any law in his set A is, a priori, possible. Bob also does not rule out the possibility that nature does not follow any pattern. This assumption enables Bayesian inferences about universal laws.[4] Assumption 2 simplifies the notation and the statement of some of the results but can be substantially weakened. Formally, all results in the article continue to hold if their statements are modified by replacing the condition "for every $\omega \in A$" with "for every $\omega \in A$ such that $P(\{\omega\}) > 0$."

Assumption 3. Given any finite history ω^t, $A \cap \omega^t \neq \varnothing$ and $A^c \cap \omega^t \neq \varnothing$.

Given any finite history ω^t, no matter how complex or simple it may be, there are infinitely many laws that are compatible with it (i.e., there are infinitely many laws $\omega \in A$ such that the first t outcomes are equal to ω^t) as well as uncountably many nonlaws that are also compatible with it. So, for any data, Bob can never rule out the hypothesis that nature abides by laws or the hypothesis that it does not. This captures the idea that there are many different ways in which past and future can relate to each other. The history 1, 1, 1, 1, 1 is equally compatible with the law "always 1" and with the law "1 in the first 5 periods and 0 afterward." In sum, the assumption ensures that it is not possible to deduce conclusively, from any finite data, whether nature abides by laws or, if so, which law. Hence, it makes clear that induction, in this article, refers to probabilistic inferences that can approach certainty but never reach it in finite time. This assumption is also satisfied by all canonical definitions of patterns, and so is useful for the interpretation of the results. However, our results remain unchanged under the weaker condition that there exists (at least) one law $\bar{\omega}$ with the property that for every t there is a law $\omega \in A$ distinct from $\bar{\omega}$ such that $\omega^t = \bar{\omega}^t$. So, upon observing t outcomes matching the path $\bar{\omega}$, Bob cannot conclude with certainty that the law, if it exists, must be $\bar{\omega}$.

Finally, we emphasize that while our main examples of laws and patterns refer to celebrated ideas such as Turing machines and periodicity, our results would continue to hold even if Bob had an eccentric understanding

4. As is well known, Bayesian inference about a hypothesis requires the latter to have initial positive probability. See, e.g., Broad (1918), Wrinch and Jeffreys (1919), and Edgeworth (1922), among others. See also Zabell (2011) for these and other references.

of what is a law or pattern. That is, none of our results depend on the labels given to laws and nonlaws, nor do they depend on the nature of the paths that are categorized as laws and nonlaws (provided that assumption 1 on the existence of at most countably many laws holds). The key point is that whatever Bob's understanding of what constitutes laws and patterns might be, he privileges paths in the set A by assigning strictly positive probability to each of them. This is a nonjudgmental, but meaningful, differentiation of laws and nonlaws because, as we discussed, only countably many paths can have strictly positive probability.

We fix for the remainder of the article a set of paths A satisfying assumptions 1 and 3. We also restrict the attention to views of the world that satisfy assumption 2.

2.2. Induction and the Separation Theorem. We now formalize specific forms of induction.

> **Definition 1.** A coherent view of the world P is *inductive in the sense of Hume* if for every path $\omega \in A$,

$$P(A|\omega^t) \to 1 \text{ as } t \to \infty. \tag{1}$$

From sufficient data with a pattern, Bob ultimately concludes, with probability approaching certainty, that nature must follow some law. A view of the world that violates (1) is such that the probability of the set A of law-like paths remains bounded away from 1, regardless of the number of realizations. Any such worldview captures what we refer to as *Humean skepticism*: Bob maintains a nonvanishing doubt that perhaps nature does not work through eternal laws, no matter how consistent and numerous the data he observes.

At each point in time, the observed finite history ω^t is consistent with a path following a pattern as well as with a path that does not follow a pattern. So, by Bayes's rule, even a view of the world that is inductive in the sense of Hume will always attach nonzero probability to the event that nature does not abide by laws (under assumption 2). What distinguishes between skepticism and inductivity in the sense of Hume is whether Bob's doubt on the regularity of nature vanishes as the number of observations that exhibit a pattern goes to infinity.

> **Definition 2.** A coherent view of the world P is *inductive in the sense of Goodman* if for every path $\omega \in A$,

$$P(\{\omega\}|A \cap \omega^t) \to 1 \text{ as } t \to \infty. \tag{2}$$

If it is granted that nature abides by some law and sufficient data with a pattern are observed, Bob infers nature's true law with increasing precision and

ultimately concludes it is eternal. A view of the world that violates (2) captures what we refer to as *Goodman's skepticism*: even assuming that an underlying law exists and that extensive evidence is available, Bob remains skeptical he will ever be able to perfectly single out the data-generating law with arbitrarily high confidence.

As in the case of induction in the sense of Hume, at no point in time is Bob's inference solved perfectly. He will always attach nonzero odds to multiple paths. However, a view of the world that is inductive in the sense of Goodman is confident Bob must ultimately assign probability close to 1 to the law generating the data.

The distinction we make here need not be seen as the formal counterpart of the classic and the new riddle of induction (see Goodman [1955] and Stalker [1994], for a discussion), and the above terminology is used mostly as a mnemonic device. Fundamentally, we ask two direct inference questions: Within the present probabilistic framework can one tell, from sufficient data and with arbitrary precision, (1) whether nature must abide by a law and (2) if so, which law?

We now examine the logical connection between these two questions.

> **The Separation Theorem.** There exist coherent views of the world that are inductive in the sense of Hume but not in the sense of Goodman and views that are inductive in the sense of Goodman but not in the sense of Hume.

The Separation Theorem shows that Hume's skepticism and Goodman's skepticism are not logically nested. One does not imply the other. In the appendix we provide simple examples of views satisfying only one of the properties. Given the separation theorem, it is meaningful to consider those coherent views of the world that express both types of faith in induction.

> **Definition 3.** A coherent view of the world P is *inductive* if it is inductive in the sense of Hume and is inductive in the sense of Goodman.

Under an inductive view of the world, skepticism about induction vanishes. Bob interprets evidence consistent with a pattern as a sign of the existence of an underlying law of nature and expects further evidence to allow him to single out the correct law with virtually exact precision. So, inductive views express great confidence in the power of empirical evidence to predict the future. This can be expressed as follows:

> **Definition 4.** A coherent view of the world P is confident that enough pattern data transform the past into a near infallible guide to the future if for every path $\omega \in A$,

$$P(\{\omega\}|\omega^t) \to 1 \ \text{ as } \ t \to \infty. \tag{3}$$

So, conditional on sufficient long pattern data ω^t, the future is forecasted with an arbitrarily high degree of certainty.

> **Remark 2.** A coherent view of the world P is inductive if and only if it is confident that enough pattern data transform the past into a near infallible guide to the future.

So, partial induction is the necessary and sufficient condition for confidence that sufficient pattern data are a near perfect guide to the future. Remark 2 delivers an initial characterization of induction that will prove useful.

3. Orgulity and σ-Additive Coherent Views. This section examines inductive properties of σ-additive coherent views. These results are known and adapted to our framework. We refer to known results as "propositions" and to novel ones as "theorems."

> **Proposition 1.** If a coherent view of the world is σ-additive then it is inductive.

The proof of this result can be found in Kelly (1996). Under σ-additivity, after multiple observations consistent with a pattern, Bob infers nature's underlying law with arbitrary accuracy and concludes with almost certainty that nature cannot follow a different law. However, σ-additivity entails even stronger forms of faith in induction.

> **Definition 5.** A coherent view of the world P is *completely inductive in the sense of Hume* if
>
> $$P(A|\omega^t) \to 1 \ \text{ as } \ t \to \infty, \ \text{ for every path } \ \omega \in A,$$
> and
> $$P(A^c|\omega^t) \to 1 \ \text{ as } \ t \to \infty, \ \text{ for } P\text{-almost every path } \ \omega \text{ in } A^c. \tag{4}$$

A coherent view of the world that is completely inductive in the sense of Hume expresses full confidence that, with sufficient data, laws and nonlaws can be distinguished empirically and with near certainty. So, complete induction in the sense of Hume is an expression of confidence that a remarkably difficult inference problem can be resolved with arbitrarily high precision.

> **Proposition 2.** Any σ-additive coherent view of the world is completely inductive in the sense of Hume.

This result has led Belot (2013) to speak of "Bayesian orgulity." The basic inference problem is difficult. Yet, σ-additive coherent views are confident that finite, but long enough, data suffice to determine with arbitrarily high precision whether nature is governed by a law.

In addition,

> **Definition 6.** A coherent view of the world P is *completely inductive* if it is completely inductive in the sense of Hume and inductive in the sense of Goodman.

Combining propositions 1 and 2 yields:

> **Corollary 1.** If a coherent view of the world P is σ-additive then it is completely inductive.

Under σ-additivity, Bob must express the following viewpoint on induction:

> I do not know whether nature works through laws or not, but given sufficient data I will find out with an arbitrarily high degree of certainty. If nature generates the data based on a law, I will ultimately conclude that nature works through laws and uncover the law nature abides by, up to a vanishing error. If the data are not governed by a law, then, in the long run, I will become near certain that nature does not follow laws. This is true even though any finite data are simultaneously consistent with countably many laws and uncountably many nonlaws.

So, under σ-additivity, Bob believes that Bayes's rule resolves these essential problems of induction. With sufficient data, nature's law is eventually uncovered. A false inference of laws, when nature follows none, is unlikely. The intuition behind these results is as follows: First, let us assume, for simplicity, that nature abides either by the law "always 1" or by a law "1 until period t and 0 thereafter," for some $t > 0$. No sequence of 1s, either large or small, suffices to infer nature's law conclusively, but there is a crucial difference between a short and a long sequence. Ex ante, the odds of the law "always 1" are fixed and strictly positive. The odds of the law "1 until some period $t \geq m$ and 0 thereafter" are arbitrarily small if m is sufficiently high. It is here that the assumption of σ-additivity is used. Under σ-additivity, such tail events must be unlikely. It now follows, by Bayes's rule, that conditional on a sufficiently long sequence of 1s, the likelihood of the law "always 1" eventually dominates the likelihood of any competing standing theories. Thus, under σ-additivity, Bob cannot express Goodman's skepticism.

The intuition regarding Hume's skepticism is related but not identical. Assume, for simplicity, that nature either abides by the law "always 1" or

does not abide by any law. Once again, no sequence of 1s, either large or small, suffices for conclusive inference. For any sequence of 1s, no matter how long, there are still many nonlaws that are consistent with it. However, the set of nonlaws that are consistent with consecutive 1s until period t shrinks monotonically to the empty set as t goes to infinity. This follows because no nonlaw is consistent with an infinite sequence of 1s. So, under σ-additivity, the ex ante odds of the set of standing nonlaws (i.e., those consistent with data of consecutive 1s until period t) goes to zero as t goes to infinity. Hence, by Bayes's rule, conditional on a sufficiently long sequence of 1s, the relative likelihood of the law "always 1" is much higher than the competing and still standing nonlaws. Thus, under σ-additivity, Bob cannot express Hume's skepticism. Finally, the intuition regarding property (4) is also similar. The set of laws consistent with nonpattern data of length t shrinks monotonically to the empty set as t goes to infinity (because no law is consistent with an infinite sequence of nonpattern data). Thus, under σ-additivity, it is unlikely that laws are consistent with long nonpattern data. Hence, property (4) holds and so does complete induction in the sense of Hume.

4. Orgulity and General Coherent Views. As shown, σ-additive coherent views rule out skepticism about induction. We now consider Bob's conclusions about the ultimate fate of multiple repetitions of Bayes's rule for general, not necessarily σ-additive, coherent views of the world. We start with an important result, a corollary of Elga (2016) (related results can also be found in Juhl and Kelly [1994] and Kelly [1996]):[5]

> **Proposition 3.** Let $\varepsilon > 0$. There exists a coherent view of the world P such that
>
> $$P(A|\omega^t) \leq \varepsilon \text{ for every } t \text{ and every } \omega \in A.$$

The view P displays a complete failure of induction in the sense of Hume. Under P, no evidence can overturn Bob's initial pessimistic belief on the existence of laws. Hence, σ-additivity suffices to rule out Hume's skepticism about induction, and this condition cannot be completely disposed of. Elga (2016) shows that not all coherent views are inductive in the sense of Hume. But, the separation theorem shows that some non-σ-additive coherent views are inductive in the sense of Hume. Moreover, there are also coherent views that are not σ-additive but nevertheless are inductive in the sense of Goodman. Lack of σ-additivity does not assure skepticism in the sense of Hume

---

5. The construction in Elga (2016) does not immediately apply to our framework (where assumptions 1–3 hold). For completeness, we provide an alternative construction in the appendix.

and does not assure skepticism in the sense of Goodman either. Other strong forms of induction can also be obtained without σ-additivity.

> **The Complete Humean Induction Theorem.** There exists a coherent view of the world P that is not σ-additive but is completely inductive in the sense of Hume.

Insomuch as confidence about Humean induction must be granted under σ-additivity, the same confidence must also be granted without σ-additivity, for some coherent views of the world. An example of such a view can be found in the appendix.

Consider, for instance, the case in which A is the set of computable paths. The Complete Humean Induction Theorem shows that some, but not all, coherent views of the world express the belief that even a fundamental problem such as whether nature can be reduced to a Turing machine can be solved (up to a vanishing error) empirically, even in the absence of σ-additivity. In this sense, Bayesian orgulity is not restricted to σ-additivity. It extends to other coherent views of the world as well.

5. The Axiomatization of Induction. The Separation and the Complete Humean Induction theorems present a difficulty for the development of a crisp theory of inductive inference. The difficulty is that confidence on solving induction problems is a product of not only well-understood conditions, such as σ-additivity, but also properties coherent views might have, which are less understood and intuitively less clear. The Complete Humean Induction Theorem is particularly challenging because it shows that confidence on empirical solutions to strong forms of inference problems can be obtained under conditions other than σ-additivity. However, let $\bar{\Sigma}$ be the smallest algebra that contains all finite histories and all singletons $\{\omega\}$ for $\omega \in A$. This is the smallest algebra that allows the expression of property (3), which is equivalent to a view P being inductive. The key point of this algebra is as follows: it is possible to obtain property (1) and also property (2) without σ-additivity. It is even possible to combine properties (1) and (4) without σ-additivity (and, hence, produce complete Humean induction). However, on $\bar{\Sigma}$, it is not possible for property (3) without σ-additivity. This makes σ-additivity not only sufficient but necessary for partial induction (and, hence, for complete induction as well). Thus,

> **The Structure Theorem.** A coherent view of the world P is inductive if and only if is σ-additive on $\bar{\Sigma}$.

The Structure Theorem is a full characterization result that delivers an axiomatic understanding of induction. The key result is the demonstration that while lack of σ-additivity does not assure skepticism in the sense of Good-

man and it does not ensure skepticism in the sense of Hume either, it always assures skepticism in at least one of these two senses. So, on $\bar{\Sigma}$, any result that holds without σ-additivity holds under some skepticism over induction. Conversely, results that require σ-additivity, also require induction.

The collection $\bar{\Sigma}$ is smaller than the σ-algebras commonly used in probability theory. While σ-algebras are mathematically convenient under σ-additivity, they do not play a particular role under finite additivity. What makes $\bar{\Sigma}$ appealing in the context of induction is that $\bar{\Sigma}$ is the simplest (i.e., the smallest) algebra that allows distinguishing between inductive and non-inductive views of the world. Small algebras such as $\bar{\Sigma}$ have an additional advantage. Because $\bar{\Sigma}$ is countable, finitely additive measures can be defined on P using only elementary mathematics and without invoking the (uncountable) Axiom of Choice.

6. Pragmatism, Induction, and de Finetti.

6. Pragmatism, Induction, and de Finetti. So far, we have focused on induction in the sense of the empirical validation of eternal laws of nature. There are, however, other perspectives on induction, such as the one in which Bob is concerned with making accurate predictions about the practical future, rather than uncovering universal laws of nature or even questioning their existence.[6]

If a law or theory makes predictions that are accurate within some finite horizon, then the theory predicts as if it were correct. Thus, the argument goes, data need not uncover the actual data-generating process. Nor do data need to reveal whether a law exists. It only needs to allow for accurate predictions for the practical future. To fix ideas, we refer to this perspective as *pragmatism*, with no claim that our narrow use of this terminology comprehends most associations with this word.

We now revisit the different problems on induction but from a more pragmatic perspective. In doing so we take a shortcut in the conceptual development. We define pragmatic inductive views as requiring that enough pattern data lead to a near infallible guide to a bounded future, instead of first making a distinction between induction in the sense of Hume and Goodman and then obtaining accurate predictions as a result of both conditions as we did in remark 2.

> **Definition 7.** A coherent view of the world P is *pragmatically inductive* if, for every path $\omega \in A$ and every natural number k,

$$P\left(\omega^{t+k}|\omega^t\right) \to 1 \quad \text{as} \quad t \to \infty. \tag{5}$$

So, with enough pattern data, Bob is convinced that the next outcomes can be predicted with near certainty. This follows, in Bob's belief, even if nature

6. See Russell (1912, chap. 6) for a discussion of induction that clearly distinguishes between the two perspectives.

abides by no laws or if it abides by a law that cannot be inferred from the data. The only claim is that after enough pattern data nature behaves as if it abides by a (data-inferred) law for a bounded but arbitrarily long future.

We now turn to the concept of complete induction in the sense of Hume, from the pragmatic perspective. Let $\mathcal{U}$ be the set of unions of finite histories. So, a set $U \in \mathcal{U}$ is a union of finite histories such as ω^t, where $\omega \in \Omega$ and t is a natural number. Any arbitrarily complex set $E \subseteq \Omega$ can be approximated in terms of finite histories by choosing a set $U \in \mathcal{U}$ such that $E \subseteq U$.[7]

> **Definition 8.** A coherent view of the world P is *pragmatically completely inductive in the sense of Hume* if for any set $U \in \mathcal{U}$ such that $A \subseteq U$,
>
> $$P(U|\omega^t) \to 1 \text{ as } t \to \infty \text{ on every } \omega \text{ in } A,$$
>
> and for any set $V \in \mathcal{U}$ such that $A^c \subseteq V$,
>
> $$P(V|\omega^t) \to 1 \text{ as } t \to \infty \text{ on } P\text{-almost every } \omega \text{ in } A^c.$$

Given the requirement for any set in $\mathcal{U}$ that contains A or A^c, there is, in particular, the same requirement for sets arbitrarily close to A or A^c. Sufficient pattern data lead to near certainty of finite histories associated with laws, and sufficient nonpattern data lead to near certainty of finite histories associated with nonlaws. Combining the two definition yields

> **Definition 9.** A coherent view of the world P is *pragmatically completely inductive* if it is pragmatically inductive and pragmatically completely inductive in the sense of Hume.

So, in particular, enough pattern data lead to a near infallible guide to a bounded future, and enough nonpattern data lead to near certainty of future finite histories associated with nonlaws.

> **The Pragmatic Induction Theorem.** Every coherent view of the world is pragmatically completely inductive.

Unlike the previous results, the Pragmatic Induction Theorem holds for all coherent views of the world. No matter how coherent beliefs are formed, they must express confidence that mechanical repetitions of Bayes's rule transform sufficiently numerous pattern data into a near infallible guide to a bounded future. In the case of nonpattern data then, provided that the data are sufficiently long, there must be confidence, approaching certainty, of an

7. For instance, the set A^c of paths not following a pattern can be written as $A^c = \cap_n U_n$, where U_n is a decreasing sequence in $\mathcal{U}$.

observable future associated with nonlaws. This holds without any other assumption such as σ-additivity. Therefore, any coherent view of the world contains a seed of orgulity.

The concerns one may have about the orgulity of Bayesians may not go away, at least completely, by abandoning σ-additivity. The Pragmatic Induction Theorem relies on multiple repetitions of Bayes's rule alone; hence, it holds with or without σ-additivity. However, the extent to which this remaining form of orgulity is a difficulty for the Bayesian paradigm is a question beyond the scope of this article. According to one viewpoint, the cases of successful inference that follow from the repetition of Bayes's rule can be seen as a desideratum that provides support to the Bayesian approach. According to a different viewpoint, the Pragmatic Induction Theorem can be seen as an expression of excessive confidence of the same paradigm. This article does not resolve this fundamental tension, but it helps to make precise the conditions under which orgulity holds.

While the Pragmatic Induction Theorem relies only on coherence and Bayes's rule, it is embedded in a standpoint that can be traced back to de Finetti. The key conceptual point advanced by de Finetti is that the Bayesian perspective on inference effectively solves the problem of induction. As de Finetti (1970, 201) stated:[8] "In the philosophical arena, the problem of induction, its meaning, use and justification, has given rise to endless controversy, which, in the absence of an appropriate probabilistic framework, has inevitably been fruitless, leaving the major issues unresolved. It seems to me that the question was correctly formulated by Hume. . . . In our formulation, the problem of induction is, in fact, no longer a problem: we have, in effect, solved it without mentioning it explicitly. Everything reduces to the notion of conditional probability."

In this sense, the Pragmatic Induction Theorem can be seen as formalization of de Finetti's viewpoint on induction. However, to the best our knowledge, de Finetti never made a distinction between the two basic inference problems (i.e., does nature abides by laws, and if so which one?) and never examined these problems in a formal model. While pragmatism is the additional element necessary for the formalization of this viewpoint, there is a yet more basic contribution. De Finetti mostly wrote about induction in the context, as in de Finetti (1969), of exchangeable beliefs (i.e., beliefs such that the order in which different outcomes occur over time is irrelevant). Exchangeability not only rules out elementary laws such as "1 until period t and 0 afterward," it is also a critical assumption on the data and, hence, an assumption on how past and future must relate to each other. In contrast, in the Pragmatic Induction Theorem, confidence on limited, but successful, induc-

8. See de Finetti (1970), chaps. 11.1.5 and 11.2.1. For de Finetti's perspective on induction, see also de Finetti (1970, 1972).

tive inference about the future holds without assumptions on how the past and the future must relate to each other. The conclusions about the future depend on the data, but there is no restriction on the data-generating process itself.

7. Extensions. This article dealt with some inductive inference problems but left others unexamined. Perhaps the most basic limitation in this article is that the data-generating processes are deterministic. A natural extension could go as follows: the Blackwell-Dubins theorem extends proposition 1 to stochastic data-generating processes. Let us say that there are countably many (possibly stochastic) data-generating processes $P_1, P_2, P_3, \ldots$, and Bob's belief (a prior over $\{P_1, P_2, P_3, \ldots\}$) assigns, ex ante, strictly positive probability to each of them. If all probabilities are σ-additive, then Bob's predictions eventually will become indistinguishable from the data-generating process, no matter which one.

In spite of the power of the Blackwell-Dubins theorem, new difficulties arise in the case of stochastic data-generating processes. For example, if two processes are identical in all but the first period, then it may be impossible to empirically determine which process runs the data. This determination is not relevant for predicting the future after period 1 (see Lehrer and Smorodinsky [1996] and Acemoglu, Cherzonukov, and Yildiz [2016] on this problem). Other difficulties may prove currently intractable. The Blackwell-Dubins theorem relies heavily on σ-additivity. For general coherent views, there are some conceptual advances, and some analytical methods for Bayesian learning were developed in Pomatto, Al-Najjar, and Sandroni (2014). With some effort, these techniques can be applied to prove a version of the Pragmatic Induction Theorem for stochastic data-generating processes. The Complete Humean Induction and the Separation theorems are existence results and so still hold when the set of data-generating processes is expanded. The main hurdle is the Structure Theorem. For a counterpart of that result, one must find an algebra on which induction is equivalent to σ-additivity when the data-generating processes can be stochastic. This is a (very) difficult problem.

Appendix

Proof of the Separation Theorem. We now provide examples of views that are inductive in the sense of Hume but not in the sense of Goodman or are inductive in the sense of Goodman but not in the sense of Hume. Fix a σ-additive measure $P_\sigma = \Sigma_{\omega \in A} \beta_\omega \delta_\omega$, where each δ_ω is the measure putting probability 1 on a path ω and β_ω are strictly positive weights such that $\Sigma_{\omega \in A} \beta_\omega = 1$. Being σ-additive, it is inductive by proposition 1.

We start with the following result.

Lemma 1. There exists a finitely additive probability S satisfying the following two properties:

$$S(\omega^t) = P_\sigma(\omega^t) \text{ for every } \omega^t;$$

$$S(A) = 0.$$

So, any finite history has the same probability under S as under P. However, under S almost every path will eventually cease to follow a pattern.

Proof of Lemma 1. Let $\mathcal{F}$ be the algebra generated by all finite histories. Consider the algebra $\mathcal{A}$ generated by $\mathcal{F}$ and the set A. As proved in Łoś and Marczewski (1949), a set $E \subseteq \Omega$ belongs to $\mathcal{A}$ if and only if it is of the form $E = (F_1 \cap A) \cup (F_2 \cap A^c)$, where F_1, F_2 belong to $\mathcal{F}$. Let M be defined as

$$M((F_1 \cap A) \cup (F_2 \cap A^c)) = P_\sigma(F_2),$$

for every set $(F_1 \cap A) \cup (F_2 \cap A^c)$ in $\mathcal{A}$. It can be easily verified that M is a well-defined probability measure on $\mathcal{A}$. Let S be any measure extending M from $\mathcal{A}$ to Σ (see, e.g., Łoś and Marczewski [1949] for a proof that such an extension exists). By construction, S satisfies the desired properties. QED

The mixture $Q = (1/2)P_\sigma + (1/2)S$ satisfies assumptions 1 and 2. It is inductive in the sense of Goodman but not in the sense of Hume. The intuition for why S is inductive in the sense of Goodman is as follows: when conditioning on A the measure Q reduces to the σ-additive measure P_σ, which is inductive. Formally, because $S(A) = 0$ then for every $\omega \in A$ we have

$$Q(\{\omega\}|A \cap \omega^t) = \frac{(1/2)P_\sigma(\{\omega\} \cap A)}{(1/2)P_\sigma(A \cap \omega^t) + (1/2)S(A \cap \omega^t)} = P_\sigma(\{\omega\}|\omega^t),$$

for each t. The measure P_σ is σ-additive hence inductive, so $P_\sigma(\{\omega\}|\omega^t)$ converges to 1 for every $\omega \in A$. Hence, Q is inductive in the sense of Goodman. To see that it is not inductive in the sense of Hume, notice that for every $\omega \in A$, we have

$$Q(A|\omega^t) = \frac{P_\sigma(A \cap \omega^t) + S(A \cap \omega^t)}{P_\sigma(\omega^t) + S(\omega^t)} = \frac{P_\sigma(A \cap \omega^t)}{2P_\sigma(\omega^t)} = \frac{1}{2}.$$

Hence, $Q(A|\omega^t)$ remains equal to 1/2 no matter how large t is. So, Q is not inductive in the sense of Hume.

We now construct an example of a measure inductive in the sense of Hume but not in the sense of Goodman. As implied by assumption 3, we can fix a path $\bar{\omega} \in A$ with the property that for every t we can find another path $\bar{\omega}_t \in A$ distinct from $\bar{\omega}$ such that $\bar{\omega}_t^t = \bar{\omega}^t$ (so $\bar{\omega}_t$ and $\bar{\omega}$ coincide on the first t outcomes but differ on some later outcome). As is well known, there exist finitely additive probability measures that assign probability 0 to each single path but probability 1 to the whole set $\{\bar{\omega}_1, \bar{\omega}_2, ...\}$ (see, e.g., Rao and Rao 1983). Let R be such a measure. We consider the mixture

$$P = \frac{1}{2}P_\sigma + \frac{1}{2}R.$$

It satisfies assumptions 1 and 2. In addition,

$$P(A|\omega^t) = \frac{P_\sigma(A \cap \omega^t) + R(A \cap \omega^t)}{P_\sigma(\omega^t) + R(\omega^t)} = 1,$$

since $P_\sigma(A) = R(A) = 1$. To see that P is not inductive in the sense of Goodman, consider the finite history $\bar{\omega}^t$. Bayes's rule implies

$$P(\{\bar{\omega}\}|A \cap \bar{\omega}^t) = \frac{P_\sigma(\{\bar{\omega}\})}{P_\sigma(\bar{\omega}^t) + R(\bar{\omega}^t)}.$$

By definition the measure R assigns probability 0 to every finite set of paths. Hence, $R(\{\bar{\omega}_k : k \geq 1, \bar{\omega}_k \in \bar{\omega}^t\}) = R(\{\bar{\omega}_k : k \geq 1\})$ for every t, so that $R(\bar{\omega}^t) = 1$. Therefore,

$$P(\{\bar{\omega}\}|A \cap \bar{\omega}^t) = \frac{P_\sigma(\{\bar{\omega}\})}{P_\sigma(\bar{\omega}^t) + 1}.$$

As $t \to \infty$, σ-additivity implies that $P_\sigma(\bar{\omega}^t)$ converges to $P_\sigma(\{\bar{\omega}\})$, so $P(\{\bar{\omega}\}|A \cap \bar{\omega}^t)$ converges to 1/2. Hence, P is inductive in the sense of Hume but not in the sense of Goodman.

Proof of the Complete Humean Induction Theorem. The proof follows the same argument in the second part of the proof of the Separation Theorem. Let P_σ and R be defined as in the above proof, and let $\tilde{\omega}$ be a path such that $\tilde{\omega} \notin A$ and $\tilde{\omega}^1 \neq \bar{\omega}^1$ (since A is countable, such a path exists). Consider the mixture $P = (1/3)P_\sigma + (1/3)R + (1/3)\delta_{\tilde{\omega}}$. As shown above, we have $P(A|\omega^t) \to 1$ as $t \to \infty$ for every path $\omega \in A$. Given the path $\tilde{\omega}$, we have that for every $t > 1$,

$$P(A^c|\tilde{\omega}^t) = \frac{P_\sigma(A^c \cap \tilde{\omega}^t) + R(A^c \cap \tilde{\omega}^t) + 1}{P_\sigma(\tilde{\omega}^t) + R(\tilde{\omega}^t) + 1},$$

since $\tilde{\omega}^t \neq \bar{\omega}^t$, then $R(\tilde{\omega}^t) = R(\{\bar{\omega}_k : \bar{\omega}_k \in \tilde{\omega}^t\}) = 0$. Therefore,

$$P(A^c|\tilde{\omega}^t) = \frac{P_\sigma(A^c \cap \tilde{\omega}^t) + 1}{P_\sigma(\tilde{\omega}^t) + 1},$$

since $\tilde{\omega} \notin A$, then $P_\sigma(\tilde{\omega}^t) \to 0$, so $P(A^c|\tilde{\omega}^t) \to 1$. Therefore, P is completely inductive in the sense of Hume. To see that P is not σ-additive, notice that for every n, we have

$$P(\{\bar{\omega}_k : k \geq n\}) = \frac{1}{3}\sum_{k \geq n} P_\sigma(\{\bar{\omega}_k\}) + \frac{1}{3}R(\{\bar{\omega}_k : k \geq n\}).$$

Because R assigns probability 0 to every finite set of paths, we have $R(\{\bar{\omega}_k : k \geq n\}) = 1$ for every n. Hence, $P(\{\bar{\omega}_k : k \geq n\}) \geq 1/3$ for every n, even if $\cap_n \{\bar{\omega}_k : k \geq n\} = \varnothing$. Hence, P is not σ-additive.

Proof of the Structure Theorem. We denote by $\mathcal{F}$ the algebra generated by all finite histories. Hence, $\mathcal{F} \subseteq \bar{\Sigma} \subseteq \Sigma$. A result related to the next lemma appears in Al-Najjar et al. (2014).

Lemma 2. A set E belongs to $\bar{\Sigma}$ if and only if there exists a set F belonging to $\mathcal{F}$ such that the symmetric difference $E\Delta F$ is finite and included in A.

Proof. Let $\mathcal{E}$ be the collection of sets E for which there exists a set $F \in \mathcal{F}$ such that the symmetric difference $E\Delta F$ is finite and included in A. We prove that $\mathcal{F} \subseteq \bar{\Sigma}$. Let E and $F \in \mathcal{F}$ be such that $E\Delta F$ is finite and included in A. Because $E\backslash F$ is finite and included in A and $\bar{\Sigma}$ is an algebra containing each singleton $\{\omega\}$ for paths in A, then $F \cup (E\backslash F) \in \bar{\Sigma}$. Similarly, $F\backslash E \in \bar{\Sigma}$, and so $E = (F \cup (E\backslash F))\backslash(F\backslash E) \in \bar{\Sigma}$. We now show that $\bar{\Sigma} \subseteq \mathcal{F}$. It follows from the definition that $\mathcal{E}$ satisfies $\mathcal{F} \subseteq \mathcal{E}$ and $\{\omega\} \in \mathcal{E}$ for each $\omega \in A$. We now prove that $\mathcal{E}$ is an algebra. Let $E \in \mathcal{E}$ be such that $E\Delta F$ is finite and included in A for some $F \in \mathcal{F}$. Because $E^c\Delta F^c = E\Delta F$ and $F^c \in \mathcal{F}$, it follows that $E^c \in \mathcal{E}$. Now let $E_1, E_2 \in \mathcal{E}$, and fix $F_1, F_2 \in \mathcal{F}$ such that $E_1\Delta F_1$ and $E_2\Delta F_2$ are finite and included in A. Let $E = E_1 \cup E_2$ and $F = F_1 \cup F_2$. Then, $E\Delta F \subseteq (E_1\Delta F_1) \cup (E_2\Delta F_2)$. Hence, $E\Delta F$ is finite and satisfies $E\Delta F \subseteq A$. Thus, $\mathcal{E}$ is closed under union and complementation. Therefore, $\mathcal{E}$ is an algebra. So, $\bar{\Sigma} \subseteq \mathcal{E}$. Thus, $\bar{\Sigma} = \mathcal{E}$. QED

We can now proceed with the proof. Let P be σ-additive. As shown in, for instance, Shiryaev (1996, 134), σ-additivity implies that P must satisfy $P(\omega^t) \to P(\{\omega\})$ as $t \to \infty$, for every $\omega \in A$. Therefore, $P(\{\omega\}|\omega^t) = P(\{\omega\})/P(\omega^t) \to 1$ whenever $P(\{\omega\}) > 0$. So, by remark 2, P is inductive in the sense of Hume and in the sense of Goodman. Conversely, suppose

P is inductive in both senses. We now show it is σ-additive on $\bar{\Sigma}$. Let μ be the restriction of P on $\mathcal{F}$. The measure μ is σ-additive on $\mathcal{F}$ (see the discussion in Rao and Rao [1983], example 10.4.2). So, by Carathéodory's theorem it admits a σ-additive extension P_μ on the σ-algebra generated by $\mathcal{F}$. In order to show that P is σ-additive (on $\bar{\Sigma}$), we prove that $P_\mu(E) = P(E)$ for every $E \in \bar{\Sigma}$.

Let $E \in \bar{\Sigma}$ and choose a set $F \in \mathcal{F}$ such that $E \Delta F$ is finite and included in A. By additivity, any measure Q satisfies

$$Q(E) = Q(F) + \sum_{\omega \in E-F} Q(\{\omega\}) - \sum_{\omega \in F-E} Q(\{\omega\}). \qquad \text{(A1)}$$

By construction, P_μ and P coincide on $\mathcal{F}$. Hence, $P(F) = P_\mu(F)$. Since P is inductive, for every $\omega \in A$, by remark 2 it satisfies $P(\{\omega\}|\omega^t) = P(\{\omega\})/P(\omega^t) \to 1$; that is, $P(\{\omega\}) = \lim_t P(\omega^t)$. The σ-additivity of P_μ and the fact P and P_μ coincide on $\mathcal{F}$ imply

$$P_\mu(\{\omega\}) = \lim_t P_\mu(\omega^t) = \lim_t P(\omega^t) = P(\{\omega\}),$$

for every $\omega \in A$. In particular, this holds for every $\omega \in E \Delta F$. We can therefore conclude from (A1) that

$$\begin{aligned}
P_\mu(E) &= P_\mu(F) + \sum_{\omega \in E-F} P_\mu(\{\omega\}) - \sum_{\omega \in F-E} P_\mu(\{\omega\}) \\
&= P(F) + \sum_{\omega \in E-F} P(\{\omega\}) - \sum_{\omega \in F-E} P(\{\omega\}) \\
&= P(E).
\end{aligned}$$

Because E is arbitrary, it then follows that P and P_μ coincide on $\bar{\Sigma}$. Hence, P is σ-additive on $\bar{\Sigma}$.

Proof of the Pragmatic Induction Theorem. Endow Ω with the product topology, and let $\mathcal{B}$ be the Borel σ-algebra generated. Let $\mathcal{F}$ be, as before, the algebra generated by all finite histories. Given any coherent view of the world P (satisfying, as usual, assumptions 1 and 2) consider the restriction μ of P on $\mathcal{F}$. Following the proof of the Structure Theorem, the measure μ admits a σ-additive extension P_σ on $\mathcal{B}$.

We now show that P is pragmatically inductive. For each $\omega \in A$ we have $P_\sigma(\{\omega\}) > 0$. To see this, notice that σ-additivity implies $P_\sigma(\{\omega\}) = \lim_t P_\sigma(\omega^t)$. For each t, we have $P_\sigma(\omega^t) = P(\omega^t) \geq P(\{\omega\}) > 0$. Hence, $P_\sigma(\{\omega\}) > 0$. Therefore, by σ-additivity, $P_\sigma(\{\omega\}|\omega^t) \to 1$ as $t \to \infty$. Since $P_\sigma(\omega^{t+K}|\omega^t) \geq P_\sigma(\{\omega\}|\omega^t)$, we conclude that $P_\sigma(\omega^{t+K}|\omega^t) \to 1$ as $t \to \infty$. Because $P_\sigma(\omega^{t+K}|\omega^t) = P(\omega^{t+K}|\omega^t)$ for every t, we conclude that P is pragmatically inductive.

The result that P is pragmatically completely inductive in the sense of Hume can be proved as a consequence of the following general principle: for every set $U \in \mathcal{U}$ and every history ω^t, we have

$$P(U|\omega^t) \geq P_\sigma(U|\omega^t).$$

We now prove this claim. The collection $\mathcal{U}$ of unions finite histories forms a base for the topology. Since the product topology is separable, each $U \in \mathcal{U}$ can be written as $U = \cup_{n=1}^{\infty} h_n$, where each h_n is a finite history. For each m, we have that $\cup_{n=1}^{m} h_n$ belongs to $\mathcal{F}$; hence,

$$P(U) \geq P\left(\bigcup_{n=1}^{m} h_n\right) = P_\sigma\left(\bigcup_{n=1}^{m} h_n\right).$$

Since $\cup_{n=1}^{m} h_n \uparrow U$ as $m \to \infty$, σ-additivity implies $P_\sigma(\cup_{n=1}^{m} h_n) \uparrow P_\sigma(U)$ as $m \to \infty$. Therefore, $P(U) \geq P_\sigma(U)$. For each t and path ω, the set $U \cap \omega^t$ is open, and the same argument as above implies that $P(U \cap \omega^t) \geq P_\sigma(U \cap \omega^t)$. Because P_σ and P coincide on $\mathcal{F}$, we also have $P(\omega^t) = P_\sigma(\omega^t)$. Hence, $P(U|\omega^t) \geq P_\sigma(U|\omega^t)$, as claimed.

Because P_σ is σ-additive, it is completely inductive in the sense of Hume. So, if $A \subseteq U$ and $A^c \subseteq V$, then $P_\sigma(U|\omega^t) \to 1$ for every $\omega \in A$, and $P(V|\omega^t) \to 1$ for P-almost every path $\omega \in A^c$. Since $P(U|\omega^t) \geq P_\sigma(U|\omega^t)$ and $P(V|\omega^t) \geq P_\sigma(V|\omega^t)$, it then follows that P is pragmatically completely inductive in the sense of Hume.

Proof of Other Results in the Text

Proof of Remark 1. The proof of this result is standard and included only for the sake of completeness. Let $D = \{\omega : P(\{\omega\}) > 0\}$ be the set of paths to which P attaches strictly positive probability. The additivity of P implies that for each positive integer k, the set $D_k = \{\omega : P(\{\omega\}) > k^{-1}\}$ must be finite. Hence, $D = \cup_{k=1}^{\infty} D_k$ is countable. QED

Proof of Remark 2. Assumptions 1, 2, and 3 imply that for each ω and t, the conditional probabilities $P(\cdot|\omega^t)$ and $P(\cdot|\omega^t \cap A)$ are well defined. In addition, by the law of total probability, for each $\omega \in A$ we have

$$P(\{\omega\}|\omega^t) = P(\{\omega\}|\omega^t \cap A)P(A|\omega^t),$$

for each $\omega \in A$. Hence, as $t \to \infty$, it follows that $P(\{\omega\}|\omega^t) \to 1$ if and only if $P(\{\omega\}|\omega^t \cap A)P(A|\omega^t) \to 1$. That is, if and only if $P(\{\omega\}|\omega^t \cap A) \to 1$ and $P(A|\omega^t) \to 1$. QED

Proof of Proposition 3. Let $\varepsilon \in (0, 1)$, and let P_σ be a σ-additive measure that satisfies assumptions 1–3. Using lemma 1, let S be a probability measure that satisfies $S(\omega^t) = P_\sigma(\omega^t)$ for every history, but $S(A) = 0$. Let $P = \varepsilon P_\sigma + (1 - \varepsilon)S$. Then, for every $\omega \in A$ and every t, we have

$$P(A|\omega^t) = \frac{\varepsilon P_\sigma(A \cap \omega^t) + (1 - \varepsilon)S(A \cap \omega^t)}{P_\sigma(\omega^t)} = \frac{\varepsilon P_\sigma(A \cap \omega^t)}{P_\sigma(\omega^t)} \leq \varepsilon.$$

QED

REFERENCES

Acemoglu, D., V. Cherzonukov, and M. Yildiz. 2016. "Fragility of Asymptotic Agreement under Bayesian Learning." *Theoretical Economics* 11:187–227.

Al-Najjar, N., L. Pomatto, and A. Sandroni. 2014. "Claim Validation." *American Economic Review* 104 (11): 3725–36.

Belot, G. 2013. "Bayesian Orgulity." *Philosophy of Science* 80 (4): 483–503.

———. 2017. "Objectivity and Bias." *Mind* 126 (503): 655–95.

Blackwell, D., and L. Dubins. 1962. "Merging of Opinions with Increasing Information." *Annals of Mathematical Statistics* 33 (3): 882–86.

Broad, C. D. 1918. "On the Relation between Induction and Probability." Pt. 1. *Mind* 27 (4): 389–404.

de Finetti, B. 1969. "Initial Probabilities: A Prerequisite for Any Valid Induction." *Synthese* 20 (1): 2–16.

———. 1970. *Theory of Probability*. Vol. 2. New York: Wiley.

———. 1972. *Probability, Induction, and Statistics*. New York: Wiley.

Doob, J. L. 1949. "Application of the Theory of Martingales." In *Le calcul des probabilites et ses applications*, 23–27. Paris: Centre national de la recherche scientifique.

Edgeworth, F. Y. 1922. "The Philosophy of Chance." *Mind* 31 (123): 257–83.

Elga, A. 2016. "Bayesian Humility." *Philosophy of Science* 83 (3): 305–23.

Gilboa, I., and L. Samuelson. 2012. "Subjectivity in Inductive Inference." *Theoretical Economics* 7 (2): 183–215.

Goodman, N. 1955. *Fact, Fiction and Forecast*. Cambridge, MA: Harvard University Press.

Huttegger, S. 2015. "Bayesian Convergence to the Truth and the Metaphysics of Possible Worlds." *Philosophy of Science* 82 (4): 587–601.

Juhl, C., and K. T. Kelly. 1994. "Reliability, Convergence, and Additivity." In *PSA 1994: Proceedings of the 1994 Biennial Meeting of the Philosophy of Science Association*, 181–89. East Lansing, MI: Philosophy of Science Association.

Kelly, K. T. 1996. *Logic of Reliable Inquiry*. Oxford: Oxford University Press.

Lehrer, E., and R. Smorodinsky. 1996. "Merging and Learning." *Statistics, Probability and Game Theory* 30:147–68.

Lévy, P. 1937. *Theorie de l'Addition des Variables Aléatoires*. Paris: Gauthier-Villars.

Łoś, J., and E. Marczewski. 1949. "Extensions of Measures." *Fundamenta Mathematicae* 1 (36): 267–76.

Pomatto, L., N. Al-Najjar, and A. Sandroni. 2014. "Merging and Testing Opinions." *Annals of Statistics* 42 (3): 1003–28.

Rao, K. P. S., and M. Rao. 1983. *Theory of Charges*. New York: Academic Press.

Russell, B. 1912. *The Problems of Philosophy*. Oxford: Oxford University Press.

Shiryaev, A. N. 1996. *Probability*. New York: Springer.

Stalker, D. F. 1994. *Grue! The New Riddle of Induction*. Chicago: Open Court.

Weatherson, B. 2015. "For Bayesians, Rational Modesty Requires Imprecision." *Ergo, an Open Access Journal of Philosophy* 2 (20).

Wrinch, D., and H. Jeffreys. 1919. "On Some Aspects of the Theory of Probability." *Philosophical Magazine* 38 (228): 715–31.

Zabell, S. L. 2011. "Carnap and the Logic of Inductive Inference." In *Handbook of the History of Logic*, ed. Dov M. Gabbay, John Woods, and Akihiro Kanamori, 265–309. Amsterdam: Elsevier.