★

Explanationist constraints on rational degrees of belief

★

Jon Williamson

Draft of September 18, 2024

Abstract

In this paper I counter Alexander Bird's claim that Bayesian probabilities ought to be informed by explanatory considerations. Bird (2022) invokes an argument of Michael Huemer to conclude that inductive inference requires explanatory constraints on prior probabilities. I suggest that this argument is unsuccessful, on account of its appeal to David Lewis' Principal Principle. Bird goes on to interpret the probabilities in Bayes' theorem other than the prior as a measure of how well a hypothesis explains the evidence. I show that this interpretation faces a new version of the old evidence problem.

§1 Introduction

Proponents of inference to the best explanation—'explanationists'—have argued that explanatory considerations impose substantive constraints on rational degrees of belief. Bird (2022), for example, claims that explanatory considerations constrain both the prior probability $P(H)$ and the ratio $P(E|H)/P(E)$ in Bayes' theorem:[1]

$$P(H|E) = P(H) \times \frac{P(E|H)}{P(E)}.$$

Bird appeals to the argument of Huemer (2009) to justify the claim that explanatory considerations constrain the prior probability $P(H)$. In §2, I present a new problem for Huemer's argument: I show that the use of the Principal Principle to justify one part of the argument undermines another part of the argument.

Bird goes on to suggest that the ratio $P(E|H)/P(E)$ can be interpreted as a measure of how well H explains E. In §3, I put forward a new kind of 'old evidence problem' that casts doubt on this explanationist interpretation.

I conclude in §4 that these explanationist arguments fail to motivate the claim that Bayesianism should take explanatory considerations into account.

[1]This paper adopts the standard conventions that P is an epistemic probability function, i.e., a function that represents an agent's rational degrees of belief, $P(\cdot|\cdot)$ denotes conditional probability, E is total evidence and H is some hypothesis of interest. The negation of a proposition A will be denoted by $\bar{A}$. 'Prior' probabilities are probabilities assigned before the evidence E is obtained. $P(H|E)$ is the 'posterior' probability of H, i.e., the probability of H in the light of evidence E.

§2
The prior

§2.1. Huemer's argument

Huemer (2009) argues in detail that prior probabilities must be informed by explanatory considerations if objective Bayesianism is to provide an account of inductive inference.[2] This argument is endorsed by Poston (2014, §7.3) and Bird (2022, §8.3), among others. The argument can be sketched as follows.

Suppose there have been n repetitions of an experiment with two possible outcomes, A_i^+ and A_i^-, but there is no further information about the experiment. Objective Bayesians would seek to apply some version of the Principle of Indifference to such a scenario. Huemer considers two ways of applying the Principle of Indifference, which yield different prior probability functions:

Keynesian prior. The function P_K that gives each possible state of outcomes the same probability, $P_K(A_1^\pm \cdots A_n^\pm) = 1/2^n$.[3]

Laplacean prior. A function P_L that gives each possible number of positive outcomes the same probability, $P_L(0) = P_L(1) = \cdots = P_L(n) = 1/(n+1)$.

I have used the term 'Keynesian prior' above, because this assignment of degrees of belief best fits the Principle of Indifference as formulated by Keynes (1921), who argued that one should give the same probability to each member of a partition of possible outcomes only if those outcomes are indivisible. Given the problem formulation, the states $A_1^\pm \cdots A_n^\pm$ form a partition of indivisible outcomes. The numbers $0, \ldots, n$ of positive outcomes are not indivisible, because, for example, the outcome 1 can be realised in multiple ways—by the states $A_1^+ A_2^- A_3^- \cdots A_n^-$ and $A_1^- A_2^+ A_3^- \cdots A_n^-$ for instance.

The latter kind of prior is 'Laplacean' because it induces what is known as Laplace's rule of succession:

$$P_L(A_{n+1}^+ | A_1^+ \cdots A_n^+) = \frac{n+1}{n+2} > \frac{1}{2} = P_L(A_{n+1}^+),$$

for $n > 0$. The use of this rule can be interpreted as yielding a kind of inductive inference: learning n outcomes, all positive, raises the probability that the next outcome will be positive.

Huemer observes that

$$P_K(A_{n+1}^+ | A_1^+ \cdots A_n^+) = \frac{1}{2} = P_K(A_{n+1}^+)$$

and claims that this identity amounts to inductive scepticism: under the Keynesian prior, learning n outcomes, all positive, fails to raise the probability that the next

[2] Subjective Bayesians hold that that there are relatively weak constraints on degrees of belief—usually, just the axioms of probability and conditionalisation as an updating rule. According to subjective Bayesianism, different agents may have markedly different degrees of belief yet be equally rational. Objective Bayesians, in contrast, hold that constraints on epistemic probabilities are much stronger. Usually, objectivists advocate the Principal Principle (see §2.2) and some kind of principle of indifference, as explained below. Objective Bayesianism is of primary interest here, because explanationists advocate strong constraints on degrees of belief. See Howson and Urbach (1989) for an introduction to Bayesianism and Williamson (2010) for a defence of one version of objective Bayesianism.

[3] Here, a state $A_1^\pm \cdots A_n^\pm$ is a conjunction of outcomes that includes either A_i^+ or A_i^- for each $i = 1, \ldots, n$.

outcome will be positive above the prior value $\frac{1}{2}$. Thus, the claim is that the Laplacean prior permits inductive inference but the Keynesian prior does not.

Huemer (2009, §3.4) then argues that the Laplacean prior can be favoured on the grounds that it is *explanatorily more basic* than the Keynesian prior, for reasons considered below. He concludes that objective Bayesianism requires explanatory constraints on the prior if it is to provide a viable account of inductive inference.

§2.2. A problem for Huemer's argument

Huemer's argument that the Laplacean prior is explanatorily more basic appeals to David Lewis' Principal Principle (Lewis, 1980):

Principal Principle. If X says that the chance of A is x, and E is any proposition that is compatible with X and admissible, then $P(A|XE) = x$.

Huemer argues that chances are explanatorily prior to both states of outcomes and numbers of positive outcomes, so one should assign a uniform probability distribution over chances, rather than directly to states or to numbers of outcomes. By adopting such an assignment, and by applying the Principal Principle, Huemer (2009, §3.3) derives a Laplacean prior.

I shall not challenge that part of Huemer's argument here. What I shall challenge is the claim that the Keynesian prior necessarily precludes inductive inference. As we shall see, versions of Bayesianism that invoke the Principal Principle can use the Principal Principle to ensure that inductive inference is possible, regardless of the choice of prior.[4]

Consider the chance proposition X that features in the Principal Principle. Chance propositions tend to be established by appeal to sample statistics, observed symmetries, and/or physical theories. (By 'established' I just mean *added to the stock of evidence.*[5]) Take the use of sample statistics, for example. The use of confidence interval estimates of chances is common across the sciences: from a sample, one infers that the chance of A lies in some specific confidence interval around the observed proportion of positive outcomes in the sample. Exactly which interval one chooses depends on the level of inductive risk one is prepared to accept, though a default 95% confidence level is typical in many sciences. Having inferred that the chance lies in some suitable confidence interval, one will then treat this as evidence from which to infer further propositions. In the standard Bayesian framework, if E contains sample data and one establishes chance proposition X, which says that the chance of A lies within some specific confidence interval, then one must condition on XE. Applying the Principal Principle, we have that rational degree of belief in A, $P(A|XE)$, must also lie in that confidence interval.

Let us return to the experiment with two possible outcomes and the question of whether the Keynesian prior permits induction. Now, $A_1^+ \cdots A_n^+$ constitutes a sample that yields the value 1 for the proportion of positive outcomes. If one is prepared to establish some proposition X about the chance of A_{n+1} from this sample, one then needs to condition on X in addition to the sample data. The posterior probability of A_{n+1}^+ is thus the quantity $P_K(A_{n+1}^+|XA_1^+ \cdots A_n^+)$, not the

[4] I focus on David Lewis' original formulation of the Principal Principle here, as that is the most familiar formulation, but the following points would apply equally to other formulations, such as the Calibration principle of Williamson (2010).

[5] Bird (2022, Chapter 5) argues in favour of the claim of Williamson (2000, Chapter 9) that evidence is knowledge, E=K, but I do not presuppose any particular account of the nature of evidence here.

quantity $P_K(A_{n+1}^+|A_1^+\cdots A_n^+)$. For example, if one is prepared to establish that the chance is within some confidence interval that has lower bound $l > 1/2$, then the Principal Principle will force

$$P_K(A_{n+1}^+|XA_1^+\cdots A_n^+) \geq l > \frac{1}{2} = P_K(A_{n+1}^+).$$

We see then that in the presence of the Principal Principle, the Keynesian prior does not preclude induction after all. It is the Principal Principle here that accommodates induction, not the choice of prior.

Recall that Huemer argues that objective Bayesian inductive inference requires explanatory constraints on the prior in order to favour the Laplacean prior over the Keynesian prior, which precludes induction. But we have seen that the Keynesian prior does not preclude induction. Hence, Huemer's argument is undermined.

§2.3. Potential responses to the problem

In response to this problem for Huemer's argument, one might raise concerns about the claim that chance proposition X is established on the basis of the sample $A_1^+\cdots A_n^+$. For example, one might maintain that one would only be prepared to infer a chance proposition X if one had evidence E that the sample were random and that the outcomes of the experiment were independent and identically distributed. No matter—the key point still holds: in suitable circumstances (i.e., where one has this further evidence E that allows one to establish X and that is admissible and compatible with X, and where X specifies a lower bound $l > 1/2$ for the chance of a positive outcome), $P_K(A_{n+1}^+|XEA_1^+\cdots A_n^+) > 1/2$, by an application of the Principal Principle. Induction remains possible with the Keynesian prior, and this possibility is enough to undermine Huemer's argument.

Alternatively, one might suggest that one should *never* establish such a chance proposition X. (One might, for example, think that one should only take the deductive consequences of one's observations as evidence, and note that X does not follow deductively from the sample.) Perhaps the most one would be prepared to do is believe X to some appropriate degree short of 1—to degree 0.95, say. Then the above problem for Huemer's argument does not seem to get off the ground.

The Bayesian can respond to this second objection, however, by appealing to Jeffrey conditionalisation instead of Bayesian conditionalisation. Jeffrey conditionalisation allows the Bayesian to update probabilities in the light of a change in degree of belief in X that is induced by the sample $A_1^+\cdots A_n^+$. If P' is the new belief function, we have:

$$\begin{array}{rcl} P'(A_{n+1}^+) & = & P_K(A_{n+1}^+|XA_1^+\cdots A_n^+)P'(X) + P_K(A_{n+1}^+|\bar{X}A_1^+\cdots A_n^+)P'(\bar{X}) \\ & \geq & l \times 0.95 + P_K(A_{n+1}^+|\bar{X}A_1^+\cdots A_n^+) \times 0.05 \\ & > & 1/2 \end{array}$$

as long as $l > 10/19$. Thus learning from experience remains possible with the Keynesian prior, even if chance propositions are never added to the stock of evidence.[6]

[6]Two technical points are relevant here. Firstly, this use of Jeffrey conditionalisation requires certain 'rigidity' conditions to hold, namely that $P'(A_{n+1}^+|XA_1^+\cdots A_n^+) = P_K(A_{n+1}^+|XA_1^+\cdots A_n^+)$ and $P'(A_{n+1}^+|\bar{X}A_1^+\cdots A_n^+) = P_K(A_{n+1}^+|\bar{X}A_1^+\cdots A_n^+)$. These do plausibly hold here. Given these rigidity conditions, and given that $P'(A_1^+\cdots A_n^+) = 1$, so $P'(X) = P'(X|A_1^+\cdots A_n^+)$ and $P'(\bar{X}) = P'(\bar{X}|A_1^+\cdots A_n^+)$, this

So far, we have considered objections that seek to limit the establishing of chance propositions. Alternatively, one might be inclined to respond to the problem by rejecting objective chances or the Principal Principle itself. But to do that would undermine Huemer's argument that the Laplacean distribution is explanatorily more basic, which appeals to both chances and the Principal Principle.

We are left with a dilemma. With the Principal Principle, both P_K and P_L can accommodate induction perfectly well. Without the Principal Principle, there are no explanatory grounds for choosing P_L over P_K. Either way, Huemer's argument turns out to be inconclusive: it does not force the conclusion that objective Bayesian inductive inference requires explanatory constraints on the prior.[7]

§3
The ratio

§3.1. Explanatory virtue and old evidence

Even if the standard argument for explanatory constraints on the prior probability $P(H)$ does not succeed, one might yet claim that there are explanationist constraints on the ratio $P(E|H)/P(E)$ that features in Bayes' theorem, and hence on the posterior probability of H. In this vein, Bird (2022, p. 207) maintains that the ratio quantifies the 'external explanatory virtue of H relative to E'. The claim is that the ratio $P(E|H)/P(E)$ can be interpreted as a measure of how well H explains E: H explains E to the extent that this ratio is greater than 1.[8]

---

use of Jeffrey conditionalisation is just an instance of the theorem of total probability for P', and should therefore be uncontroversial.

Second, Jeffrey conditionalisation was developed to handle exogenous changes in degrees of belief: i.e., changes that are not captured simply by conditioning on other established propositions (see, e.g., Jeffrey, 2004, §3.2). In particular, one should not suppose that $P'(X)$ is simply identifiable with $P_K(X|A_1^+ \cdots A_n^+)$, even though the change in degree of belief in X is prompted by the sample $A_1^+ \cdots A_n^+$. Indeed, it is not possible that $P'(X) = 0.95 = P_K(X|A_1^+ \cdots A_n^+)$, on pain of contradiction: if $P_K(X|A_1^+ \cdots A_n^+) = 0.95$ and $l > 10/19$ then,

$$\begin{aligned}
1/2 &= P_K(A_{n+1}^+|A_1^+ \cdots A_n^+) \\
&= P_K(A_{n+1}^+|XA_1^+ \cdots A_n^+)P_K(X|A_1^+ \cdots A_n^+) \\
&\quad +P_K(A_{n+1}^+|\bar{X}A_1^+ \cdots A_n^+)P_K(\bar{X}|A_1^+ \cdots A_n^+) \\
&> 10/19 \times 0.95 \\
&= 1/2,
\end{aligned}$$

a contradiction. The objective Bayesian can rationalise the fact that $P'(X) \neq P_K(X|A_1^+ \cdots A_n^+)$ as follows (see Williamson, 2010, §4.2). Conditionalisation is only appropriate where certain requirements are met: in particular, the proposition θ that is conditioned upon needs to be 'simple' in the sense that all the information it provides can be captured by the constraint $P(\theta) = 1$. Here, the sample proposition $A_1^+ \cdots A_n^+$ is not simple because it not only tells us that $P(A_1^+ \cdots A_n^+) = 1$ but it also provides information about chances. Thus we should not expect that $P'(X) = P_K(X|A_1^+ \cdots A_n^+)$.

The key point is that Huemer's argument fails for any version of Bayesianism that adopts the Principal Principle and permits exogenous changes in beliefs about chance propositions, even if chance propositions are never fully established.

[7]One might wonder what grounds we might have, then, for choosing one of P_K and P_L over the other, and, more generally, how to avoid inconsistencies that might arise by applying the Principle of Indifference in mutually incompatible ways. While this is a bigger question than can be answered here, it is worth noting that in a Bayesian framework, different ways of applying the Principle of Indifference arguably lead only to subjectivity, not inconsistency (Williamson, 2010, Chapter 9).

[8]Bird's claim is that ratio as a whole has an explanationist interpretation, but that the individual probability $P(E|H)$ does not have a distinct explanationist interpretation, and nor does $P(E)$. Fur-

A problem arises for this interpretation of the ratio as a measure of explanatory virtue, however. The problem is that we only seek to explain previously established propositions. We do not, for instance, seek to explain why the moon is made of blue cheese, because the moon has not been established to be made of blue cheese. Now, for the Bayesian, any previously established proposition E must be fully believed, $P(E) = 1$. But if $P(E) = 1$ then $P(E|H) = 1$ for any H such that $P(H) > 0$. In which case,

$$\frac{P(E|H)}{P(E)} = 1,$$

i.e., H is deemed to be explanatorily neutral with respect to E. Thus, whenever E is a proposition that we might seek to explain, the ratio measure takes H to be explanatorily neutral with respect to E. But the explanationist will want to maintain that there are genuine cases in which $P(H) > 0$ and H is not explanatory neutral with respect to E. By modus tollens, then, one can only conclude that the ratio cannot be interpreted as a measure of how well H explains E.[9]

The upshot is that we have a new kind of 'old evidence problem': a precondition for interpreting the ratio as a measure of how well H explains E is that E should already be evidence, in which case the ratio measure deems H to be explanatorily neutral with respect to E.[10]

One might try to respond to this problem in one of two ways: by trying to tackle the problem head on (§3.2) or by attempting to avoid the problem altogether (§3.3). Both paths are beset with obstacles, as we shall see.

§3.2. Tackling the new old evidence problem

This new old evidence problem differs from the standard old evidence problem of Glymour (1981). The standard problem is that it is difficult to see how old evidence E can possibly confirm a hypothesis H. According to Bayesian confirmation theory, confirmation requires probability raising, i.e., $P(H|E) > P(H)$. But since E is old evidence, $P(E) = 1$, so $P(H|E) = P(H)$.

Despite the apparent differences between the new and the standard old evidence problems, there are also similarities. In particular, the two problems stem from the fact that when $P(E) = 1$ and $P(H) > 0$, $P(H|E)/P(H) = P(E|H)/P(E) = 1$. The question thus arises as to whether some response to the standard problem might help to resolve the new problem.

---

thermore, Bird takes informal explanationist reasoning to have largely heuristic value: it offers an approximation to more formal Bayesian reasoning, 'leading to the same or similar judgments reasonably frequently' (Bird, 2022, p. 208). Thus, the explanationist interpretation of the ratio saves us from having to calculate $P(E|H)$ and $P(E)$ individually: we need only multiply the ratio, which measures the external explanatory virtue of H relative to E (aka its 'external loveliness'), by the prior probability of H, which measures the internal explanatory virtue of H (aka its 'internal loveliness').

[9] Note that this problem is orthogonal to Lipton's distinction between inference to the best *actual* explanation and inference to the best *potential* explanation (Lipton, 1991, Chapter 4). Lipton argues that inference to the best explanation is fallible and should thus be construed as inference to the best potential explanation, not inference to the best actual explanation: the latter form of inference would be infallible because the actual explanation must be at least approximately true. Thus Lipton's concern is that one should not presuppose the truth of the hypothesis H that one infers. Here, in contrast, the concern is that Bird's interpretation of the ratio presupposes the truth of the evidence E. This concern arises for inference to the best potential explanation, just as it does for inference to the best actual explanation.

[10] This problem also besets a closely related measure of explanatory power due to Good (1960). See also McGrew (2003) and Schupbach and Sprenger (2011).

There are two main lines of response to the standard old evidence problem: one due to Daniel Garber and one due to Colin Howson (see, e.g., Sprenger, 2015). The approach of Garber (1983) is intended to deal with failures of logical omniscience: cases in which one does not initially realise that the total stock F of current evidence entails E. It is then not E itself but the proposition that F *entails* E that is of primary interest. The new old evidence problem does not hinge upon a failure of logical omniscience, however. The problem here is that explaining E presupposes that E is already established—it is no surprise that F entails E. Therefore, Garber's response to the standard problem does not extrapolate to the new problem.

Howson (1991) provided a second strategy for tackling the standard old evidence problem and Howson's strategy is more readily applicable to the new problem. Consider the extent to which one would have believed E counterfactually, were E not already in one's stock of evidence F. One might then take the ratio

$$\frac{P(E|H(F-E))}{P(E|F-E)}$$

to be a viable measure of how well H explains E, where $F-E$ is formed by contracting F in such a way that E cannot be established from $F-E$.

This contraction strategy faces two obstacles, however: one to do with auxiliary hypotheses and one to do with underdetermination.

The first obstacle is that the contraction strategy can fail to yield an accurate measure of how well H explains E. This is because contraction can remove propositions from F that inform how well H explains E. In particular, contraction can remove theoretical propositions from F that bear on the connection between H and E. For example, suppose that H is the Standard Model of particle physics and E is the proposition that the 2012 experiments of the Large Hadron Collider observed the Higgs Boson. Our total evidence F includes E, which was established early in 2013. The question is whether $P(E|H(F-E))/P(E|F-E)$ can be interpreted as a measure of how well H explains E. Note that H only explains E in the context of a multitude of auxiliary hypotheses about physics, experimentation and the proper functioning of the Large Hadron Collider. Now, many of these auxiliary hypotheses were also required to establish E in the first place, back in 2013, so, when forming the contraction $F-E$, one might reasonably remove some of these hypotheses in order to ensure that E cannot be established from $F-E$. But removing these crucial auxiliary propositions will underestimate how well H explains E, because H requires these auxiliaries to explain E. Hence, $P(E|H(F-E))/P(E|F-E)$ can turn out to be a poor measure of how well H explains E.

The second obstacle to the contraction strategy is one of underdetermination: there will usually be a great many ways to construct $F-E$. Indeed, the literature on belief contraction and revision posit conditions that a contraction operator should satisfy and it soon becomes clear that there are many reasonable contraction operators (see, e.g., Hansson, 2022). This underdetermination of contraction leads to the ratio measure being ill-defined.[11]

[11]In the context of subjective Bayesianism, this underdetermination is perhaps of no great concern. Subjective Bayesians hold that any prior probability function P is rationally permissible, and permissiveness with respect to the contraction operator is just more of the same thing. But as pointed out in note 2, subjectivism is not what is required here. The explanationist is looking for substantive constraints on rational degree of belief that are forced by explanatory considerations. For the explanationist, it is rationally required, not merely rationally permissible, that these constraints are satisfied. It is for this

One might object at this point that there are cases in which the contraction strategy appears relatively unproblematic. For example, if E happens to be the single most recently established proposition, one might take the rational belief function held just prior to establishing E as a proxy for $P(\cdot|F-E)$. In such a situation, there is arguably no need for a general rule for determining $F-E$.

In response, it is important to observe that what this suggestion is doing is treating the evidence just prior to establishing E as a candidate for $F-E$. While this may be a viable candidate, it remains but one candidate: other candidates for $F-E$ will usually also be rationally permissible. Thus, the problem of underdetermination remains. The explanatory virtue interpretation of the ratio requires a determinate way of retracting old evidence, yet there is none to be found because there are many equally viable candidates for $F-E$.

We have seen, then, that attempts to resolve the standard old evidence problem do not adequately resolve the new old evidence problem that arises for Bird's construal of the ratio as a measure of how well H explains E. Garber's approach is not applicable to the new problem. While Howson's contraction strategy might appear to offer a way out, it faces two substantial obstacles in the present context: a problem of auxiliary hypotheses and a problem of underdetermination.

§3.3. Avoiding the new old evidence problem

Rather than attempting to solve the new old evidence problem, one might try to avoid the problem altogether by denying that E is old evidence. In particular, one might attempt to disavow Bird's interpretation of E as *evidence*: this would avoid the claim that E must have been previously established. One could, for example, take explanatory constraints on the ratio $P(E|H)/P(E)$ to apply only to prior probabilities formulated before any evidence has been collected. In this situation, the proposition E is not evidence at all—neither old nor new. And without evidence, our new old evidence problem cannot bite.[12]

Unfortunately, this view is prone to the following rather different kind of problem. According to this view, the ratio does not measure the explanatory virtue of a hypothesis in relation to evidence—rather, it measures the explanatory virtue of one arbitrary proposition in relation to some other arbitrary proposition. Now, for most pairs of propositions, A,B, proposition A will be explanatorily neutral with respect to proposition B. For example, the proposition A that the 2012 experiments of the Large Hadron Collider observed the Higgs Boson does not explain the proposition B that the moon is made of blue cheese. If explanatory considerations are to constrain the ratio, we must have that $P(B|A)/P(B) = 1$ whenever A is explanatorily neutral with respect to B.

All these neutrality constraints force the prior probability function to satisfy many probabilistic independencies. In the above example, both A and its negation $\bar{A}$ are explanatorily neutral with respect to B, forcing $P(B|A) = P(B) = P(B|\bar{A})$. A similar constraint must hold for any two propositions A and B such that neither A nor $\bar{A}$ explains B. But these constraints quickly become untenable. A and B

reason that Huemer appeals to objective Bayesianism and Bird invokes what he calls 'super-objective Bayesianism' (Bird, 2022, p. 186). An ill-defined or very permissive measure of how well H explains E will not impose substantive enough constraints for the explanationist.

[12] Alternatively, by drawing an analogy with Lipton's distinction between actual and potential explanations (see note 9), one might phrase this objection as follows: one should not take E to be 'actual' evidence, but rather 'potential' evidence. Again, the key idea is to interpret the ratio before evidence is collected.

might be effects of a common cause: neither explains the other, but each raises the probability of the other; Bayesianism will fail to validate this probability raising if explanatory neutrality forces probabilistic independence. Worse, A might logically imply B without explaining B: explanatory neutrality would force $P(B|A) = P(B)$ while the axioms of probability force $P(B|A) = 1$, leading to inconsistency whenever $P(B) < 1$. Hence, this attempt to rescue the explanatory virtue interpretation arguably leads to problems of even greater magnitude than the new old evidence problem.

Incidentally, a sample $A_1^{\pm} \cdots A_n^{\pm}$ from the experiment of §2 cannot be said to *explain* the next outcome $A_{n+1}^{\pm}$, so explanatory neutrality would force $P(A_{n+1}^{\pm}|A_1^{\pm} \cdots A_n^{\pm}) = P(A_{n+1}^{\pm})$ for all $n \geq 1$. These independence constraints lead to the Keynesian prior over outcomes of the experiment, $P_K(A_1^{\pm} \cdots A_n^{\pm}) = 1/2^n$. Thus, any attempt to move away from Bird's interpretation of E as evidence threatens to further undermine his appeal to Huemer's argument for explanatory constraints on the prior, because this move favours the Keynesian prior over the Laplacian prior.

§4
Conclusion

Huemer's argument for explanationist constraints on the prior $P(H)$ rests on the claim that the Keynesian prior is necessarily non-inductive. In the presence of the Principal Principle, this claim is false: the Principal Principle can be exploited to ensure that any prior—the Keynesian prior included—can accommodate inductive inference. On the other hand, Bird's argument for explanationist constraints on the ratio $P(E|H)/P(E)$ faces a new kind of old evidence problem, and, as we have seen, attempts to tackle or avoid this problem face formidable obstacles. These problems undermine the claim that there are substantive explanationist constraints on Bayesian probabilities.

Acknowledgements

I am grateful to the anonymous referee for helpful comments. This research was supported by Leverhulme Trust grant RPG-2022-336 and UKRI CRCRM grant 25130.

Bibliography

Bird, A. (2022). *Knowing Science.* Oxford University Press, Oxford.

Garber, D. (1983). Old evidence and logical omniscience in Bayesian confirmation theory. In Earman, J., editor, *Testing Scientific Theories*, Minnesota Studies in the Philosophy of Science, pages 99–132. University of Minnesota Press, Minneapolis.

Glymour, C. (1981). Why I am not a Bayesian. In *Theory and evidence*, pages 63–93. University of Chicago Press, Chicago IL.

Good, I. J. (1960). The paradox of confirmation. *British Journal for the Philosophy of Science*, 11:145–149.

Hansson, S. O. (2022). Logic of Belief Revision. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy.* Metaphysics Research Lab, Stanford University, Spring 2022 edition.

Howson, C. (1991). The 'old evidence' problem. *British Journal for the Philosophy of Science*, 42(4):547–555.

Howson, C. and Urbach, P. (1989). *Scientific reasoning: the Bayesian approach.* Open Court, Chicago IL, second (1993) edition.

Huemer, M. (2009). Explanationist aid for the theory of inductive logic. *The British Journal for the Philosophy of Science*, 60(2):345–375.

Jeffrey, R. (2004). *Subjective probability: the real thing.* Cambridge University Press, Cambridge.

Keynes, J. M. (1921). *A treatise on probability.* Macmillan, London, 1973 edition.

Lewis, D. K. (1980). A subjectivist's guide to objective chance. In *Philosophical papers*, volume 2, pages 83–132. Oxford University Press (1986), Oxford. With postscripts.

Lipton, P. (1991). *Inference to the best explanation.* Routledge, second (2004) edition.

McGrew, T. (2003). Confirmation, heuristics, and explanatory reasoning. *The British Journal for the Philosophy of Science*, 54(4):553–567.

Poston, T. (2014). *Reason and Explanation: A Defense of Explanatory Coherentism.* Palgrave Innovations in Philosophy. Palgrave Macmillan, Basingstoke.

Schupbach, J. N. and Sprenger, J. (2011). The logic of explanatory power. *Philosophy of Science*, 78(1):105–127.

Sprenger, J. (2015). A novel solution to the problem of old evidence. *Philosophy of Science*, 82(3):383–401.

Williamson, J. (2010). *In defence of objective Bayesianism.* Oxford University Press, Oxford.

Williamson, T. (2000). *Knowledge and its limits.* Oxford University Press, Oxford.