Moderate Bayesianism Is Incompatible With Equal Treatment

August 26, 2016

Abstract

*Equal Treatment*, recently articulated and defended by Susanna Rinard (forthcoming), holds that beliefs are rational or irrational in exactly the same way that actions are. According to Bayesian expected utility theory an action is rational if and only if it maximizes expected utility. Thus, Bayesians who embrace Equal Treatment are led to the view that a belief is rational if and only if it maximizes expected utility. I show that, under a widely accepted condition, this view is inconsistent with almost all forms of Bayesianism. In fact, any form of Bayesianism consistent with Equal Treatment must endorse Extreme Permissivism.

1 Introduction

1.1 Equal Treatment

What is it that makes it rational to bring an umbrella when it is raining, or to study for an upcoming exam? Philosophers have come up with many different accounts of what makes an action rational, but from a sufficiently abstract point of view most accounts endorse the view that rational action is inextricably tied up with the well-being of agents. Thus, an action is rational for an agent insofar as the action –
somehow or other – helps the agent (or perhaps someone else) attain what they want or should want.

However, many epistemologists reject the view that agent well-being is relevant to whether it is rational for an agent to hold a belief. The popular yet uncompromising doctrine of *Evidentialism* says that it is rational for an agent to believe a proposition if and only if the proposition is supported by the evidence; whether believing the proposition makes the agent feel good or has other beneficial consequences is entirely irrelevant.

Other philosophers believe that it is sometimes rational to hold a belief that is not supported by the evidence, but that in that case the kind of rationality at play is pragmatic rather than epistemic. Thus the holding of a belief may be pragmatically rational and yet epistemically irrational, or the other way around. But, conceptually, there is an important distinction to be drawn between epistemic rationality on the one hand and pragmatic rationality on the other.

Susanna Rinard (forthcoming) disagrees. Rinard’s goal is to argue that there is only one standard of rationality, and that all states an agent may occupy, including holding a certain belief or choosing a certain action, should be evaluated according to that same standard. Rinard does not endorse any particular account of rationality, but instead argues for the more abstract thesis that the true account of rationality – whatever it happens to be – must apply in the same way to all states. Briefly, then, *Equal Treatment* is the thesis that,

*Equal Treatment* All states (including choosing an action or holding a belief) are rational or irrational in precisely the same way.

Since Equal Treatment is an abstract thesis, Rinard’s arguments proceed on a rather abstract level; for example, she argues that Equal Treatment is better than its competitors because it is theoretically simpler and more elegant. Although we may agree with Rinard that Equal Treatment has various theoretical advantages relative

---

1Rinard routinely refers to actions such as, e.g., mowing the lawn as “states” throughout her paper; I will follow her practice in using “state” as a generic umbrella term that covers beliefs, actions, and perhaps other things (though I will focus only on actions and beliefs).
to the alternatives, Rinard fails to address the question of whether Equal Treatment is actually compatible with any of the well-developed theories of rationality on offer. I will argue that one of the most influential frameworks for rational belief management and decision making – namely, Bayesian expected utility theory, or “Bayesianism” for short – is hard to reconcile with Equal Treatment. In the next subsection, I briefly describe the tenets of Bayesian expected utility theory; In Sections 2 and 3, I show that – given a widely accepted condition – most forms of Bayesianism are inconsistent with Equal Treatment; I end Section 3 by suggesting that my conclusion is borne out by the arguments provided by philosophers who work on epistemic utility theory.

1.2 A brief description of Bayesian expected utility theory

Bayesians embrace the view that beliefs vary in strength, and they use the real numbers between 0 and 1 to represent agents’ degrees of belief in propositions in the following way: for any given agent, that agent’s beliefs in a set of propositions is represented as a function – called the agent’s “credence function” – from the set of propositions to \([0, 1]\), where 0 represents certain disbelief and 1 represents certain belief. If the credence function is a probability function, it is said to be “coherent.” All Bayesians agree that being coherent is a minimal rational requirement on agents’ degrees of belief, but most do not think coherence is sufficient since it’s not hard to come up with coherent credence functions that nonetheless seem irrational. For example, it’s quite possible to have a credence function that assigns 1 (i.e. certainty) to the proposition that the Earth is made of silicon although, intuitively, such a belief does not seem rationally defensible. Versions of Bayesianism that take coherence to be necessary \textit{and} sufficient for rationality endorse \textit{Extreme Permissivism} in the sense that any probability function counts as rational. In the next section, I will show that (under a widely accepted condition) any version of Bayesianism consistent with Equal Treatment must endorse Extreme Permissivism.

In addition to prescribing how agents should rationally manage their beliefs, Bayesianism comes equipped with a theory of rational decision making, namely
Bayesian expected utility theory. A Bayesian expected utility calculation requires two components: first, a utility function, \( U \) that specifies the value of the various possible states, \( S_i \), of the world; and second a probability function, \( p \) that specifies the probability of those possible states. An agent who has both a utility function and a probability function may then calculate the expected utility of performing action \( A \) (or being in state \( A \)) relative to \( p \) and \( U \) with the following formula:\(^2\)

\[ U(A) = \sum_i p(S_i)U(S_i&A) \quad (1.1) \]

Bayesian expected utility theory now offers the following criterion for rational action:

**Principle of Maximum Utility** An action is rational for an agent if and only if the action maximizes expected utility for the agent.

Note that the Principle of Maximum Utility assumes that the agent comes to the table already equipped with a probability function; after all, expected utility must be calculated relative to a probability function. According to Equal Treatment, however, beliefs – and thus probability functions, which are the mathematical representations of graded belief states – are rational or irrational in exactly the same way that actions are. In a Bayesian framework, Equal Treatment therefore implies that there must be a criterion analogous to the Principle of Maximum Utility for probability functions. I explore what such a criterion should look like in the next section.

### 2 Bayesian expected utility theory and Equal Treatment

As a first pass, we may try the following formulation, which follows the wording of the Principle of Maximum Utility as closely as possible:

\(^2\)The following formula requires the assumption that the actions and states of the world are independent according to \( p \). This is an assumption I will, for the sake of simplicity, make in the remainder of the paper.
**PMU for Probability Functions** Probability function $p$ is rational for an agent if and only if $p$ maximizes utility for the agent.

But the above formulation is not sufficiently spelled out. Imagine that you are entertaining various probability functions and that you are trying to figure out which one of them is rational for you to adopt. The PMU for Probability Functions, as stated, requires that you first choose a probability function, $p'$, relative to which you may then perform the expected utility calculations. But how do you pick which $p'$ to use? Your choice of $p'$ will determine which of the probability functions under consideration receives the highest expected utility, so it is an important choice.

Presumably you should use the $p'$ that it is rational for you to use, since whether $p$ maximizes utility relative to just any distribution, rational or irrational, is uninteresting. Indeed, as we will see in Section 3, almost all philosophers accept conditions that entail that every probability function maximizes utility relative to some probability function, in which case the question of whether $p$ maximizes utility relative to some distribution is not only uninteresting, but has a trivially true answer.

So let’s grant that expected utility needs to be calculated relative to the $p'$ that it is rational for you to use. According to the PMU for Probability Functions, that means you should use the $p'$ that maximizes your expected utility. But now you face the same problem as before: the expected utility of $p'$ must be calculated relative to some probability function, $p''$, so, again, the question arises of which $p''$ to use. Once again, you presumably ought to use the $p''$ that it is rational for you to use, and once again this requires – by the PMU for Probability Functions – that you already have a probability function, $p'''$, that you may use to calculate the expected utility of $p''$. A potentially infinite regress now looms: you are faced with a chain of probability functions, $p_1, p_2, \ldots$, where, for every $p_i$ in the chain, in order to figure out whether $p_i$ is rational, you must first find the $p_{i+1}$ that it is rational for you to use.

I can only see two ways of avoiding the infinite regress. Either it must be the case that $p_i = p_{i+1}$ at some point in the chain of probability distributions, so that the expected utility of $p_i$ is calculated relative to itself, or else it must be the case that at some point in the chain, $p_i$, it is no longer necessary to calculate the expected
utility of $p_i$ in order to determine whether $p_i$ is rational. The second alternative is immediately ruled out by Equal Treatment. According to Equal Treatment, if any state is rational in virtue of maximizing expected utility, then every state must be rational in virtue of maximizing expected utility.

That leaves the first alternative. If there is some probability distribution in the chain such that $p_i = p_{i+1}$, then the regress clearly stops. Thus, if the regress stops at $nth$ level of the chain, then you are able to determine that, say, $p_1$ is rational because $p_1$ maximizes expected utility relative to $p_2$, and that $p_2$ is rational because it maximizes expected utility relative to $p_3$, ... etc. all the way up to $p_n$, which is rational because it maximizes utility relative to $p_n$. The previous line of thought leads to the following spelled-out and refined formulation of PMU for Probability Functions:

**Refined PMU for Probability Functions**: Probability function $p$ is rational for an agent if and only if there is a finite chain of probability functions, $p = p_1, p_2, \ldots p_n$, such that, for each $i < n$, $p_i$ maximizes utility relative to $p_{i+1}$, and moreover $p_n$ maximizes utility relative to $p_n$.

Unfortunately, as I will show below, given a widely accepted necessary condition on utility functions, Refined PMU for Probability Functions renders every probability function rational; in other words, under a very plausible and ubiquitous assumption, Refined PMU for Probability Functions entails what I earlier called Extreme Permissivism.

3 Proper Utility Functions

The idea that Bayesian expected utility theory may be used to justify constraints on credence functions has become an influential position in recent years: inspired by Joyce’s (1998) argument for Probabilism, philosophers working on so-called “epistemic utility theory” have provided arguments for, among other things, Conditionalization (Greaves and Wallace, 2006), Conglomerability and Reflection (Easwaran,
2013\(^3\), the Principal Principle (Pettigrew, 2012, 2013), and the Principle of Indifference (Pettigrew, 2014).

In principle Bayesianism leaves it open for agents to use all kinds of utility functions, but the idea behind epistemic utility theory is that – in an epistemic context – agents’ utility functions ought to satisfy certain very plausible constraints. Most important is the constraint that credence functions be judged by how accurate they are. Here, “accuracy” is defined as distance from the truth according to some appropriate distance measure. These distance measures thus play the role of utility functions; in the literature, they are typically known as “scoring rules.” It is a minimal requirement on any scoring rule that the credence function that assigns 1 to every truth and 0 to every falsehood have maximum utility.

Another widely accepted requirement on scoring rules is that they be “proper.” A scoring rule is proper if and only if its expected value when calculated relative to \( p \) is maximized by \( p \) itself, whenever \( p \) is a probability function. Just about all formal epistemologists who currently work on epistemic utility theory work exclusively with proper scoring rules.\(^4\) Furthermore, as far as I know, all of the scoring rules that people have provided independent arguments for in both philosophy and in statistics have turned out to be proper. So proper utility functions are widespread, to say the least.\(^5\)

Now, suppose your utility function is proper. Given any arbitrary probability function, \( p \), it follows from the fact that your utility function is proper that \( p \) must maximize expected utility relative to itself. But now there trivially is a chain of probability functions that satisfies the right-hand-side of the biconditional in Refined PMU for Probability Functions: namely the chain that consists of just \( p \). Thus, according to the criterion, \( p \) is rational for you. But \( p \) was arbitrary. Thus, you are led to accept Extreme Permissivism: by your lights, every probability function is

\(^3\)Easwaran also strengthens Greaves and Wallace’s result.

\(^4\)All but one of the examples mentioned in the previous paragraph requires propriety or something that amounts to propriety in order for the results to go through. The exception is Pettigrew (2012), which Pettigrew (2013) repudiates.

\(^5\)Several philosophers have also provided arguments for propriety or properties closely related to propriety, e.g. Oddie (1997), Gibbard (2007), and Joyce (2009).
rational.

More generally, if we assume – along with almost all philosophers and statisticians – that propriety is a non-negotiable necessary condition on utility functions, then Extreme Permissivism follows from Refined PMU for Probability Functions. Thus, any form of Bayesianism that rejects Extreme Permissivism must consequently reject Refined PMU for Probability Functions. But since Refined PMU for Probability Functions is just the Bayesian precisification of Equal Treatment, it follows that any moderate form of Bayesianism that rejects Extreme Permissivism must also reject Equal Treatment. Thus, whatever theoretical virtues it may have, moderate Bayesians cannot accept Equal Treatment.

An interesting consequence of the above argument is that whatever extra constraints in addition to the probability axioms stronger forms of Bayesianism include – such as, say, the Principle of Indifference or the Principal Principle – those extra constraints cannot be justified by arguments that appeal to Bayesian expected utility maximization alone. Additional assumptions must necessarily play an important role, or the arguments must reject the Bayesian principle of maximum expected utility.

Inspecting the various epistemic utility arguments that have been given for, e.g., the Principle of Indifference and the Principal Principle confirms this to be the case. In particular, in his argument for the Principle of Indifference, Pettigrew eschews the usual Bayesian criterion of utility maximization in favor of the Minimax criterion (Pettigrew, 2014). Furthermore, in his argument for the Principal Principle, Pettigrew (2013) calculates the expected utility of credence functions relative to the possible objective chance functions, which surely cannot be justified by a Bayesian expected utility calculation. My point here is not to pick on Pettigrew, but merely to point out that philosophers who work on epistemic utility theory do violate Equal Treatment in their arguments, which is to be expected from the discussion in this paper.
4 Conclusion

I have argued that Bayesianism is hard to reconcile with Equal Treatment, but there are a few ways in which the two can be made to be compatible. One could “bite the bullet” and accept Extreme Permissivism. Alternatively, one could reject the assumption that utility functions must be proper. Yet another option is to somehow change the Bayesian principle of maximum utility, perhaps by specifying some unique probability function that should be used in all expected utility calculations.

Those who do not find the above options appealing face a choice: they need to give up either Bayesianism or Equal Treatment. I take no stance here on what the better option is.

References


