The Bayesian and the Dogmatist

Brian Weatherson, October 17

references etc incomplete – talk versioon

There is a lot of philosophically interesting work being done in the borderlands between traditional and formal epistemology. It is easy to think that this would all be one-way traffic. When we try to formalise a traditional theory, we see that its hidden assumptions are inconsistent or otherwise untenable. Or we see that the proponents of the theory had been conflating two concepts that careful formal work lets us distinguish. Either way, the formalist teaches the traditionalist a lesson about what the live epistemological options are. I want to argue, more or less by example, that the traffic here should be two-way. By thinking carefully about considerations that move traditional epistemologists, we can find grounds for questioning some presuppositions that many formal epistemologists make.

To make this more concrete, I'm going to be looking at a Bayesian objection to a certain kind of dogmatism about justification. It has been urged that the incompatibility of dogmatism with a kind of Bayesianism is a reason to reject dogmatism. I rather think that it is reason to question the Bayesianism. To put the point slightly more carefully, there is a simple proof that dogmatism (of the kind I envisage) can't be modelled using standard Bayesian modelling tools. Rather than conclude that dogmatism is therefore flawed, I conclude that we need better modelling tools. I'll spend a fair bit of this paper on outlining a kind of model that (a) allows us to model dogmatic reasoning, (b) is motivated by the epistemological considerations that motivate dogmatism, and (c) helps with a familiar problem besetting the Bayesian.

I'm going to work up to that problem somewhat indirectly. I'll start with looking at the kind of sceptical argument that motivates dogmatism. I'll then briefly rehearse the argument that shows dogmatism and Bayesianism are incompatible. Then in the bulk of the paper I'll suggest a way of making Bayesian models more flexible so they are no longer incompatible with dogmatism. I'll call these new models *dynamic Keynesian* models of uncertainty. I'll end with a brief guide to the virtues of my new kind of model.

1. Sceptical Arguments

Let *H* be some relatively speculative piece of knowledge that we have, say that G. E. Moore had hands, or that it will snow in Alaska sometime next year. And let *E* be all of our evidence about the external world. I'm not going to make many assumptions about what *E* contains, though we will have to say one thing. But for now *E* will stay fairly schematic. Now a fairly standard sceptical argument goes something like this. Consider a situation *S* in which our evidence is unchanged, but in which *H* is false, such as a brain-in-vat scenario, or a zombie scenario, or a scenario where the future does not resemble the past. Now a fairly standard sceptical argument goes something like this.

1. To know *H* you have to be in a position to know you aren't in *S*
2. You aren't in a position to know that you aren't in *S*
3. So, you don't know *H*

There are a few immediate responses one could make, but which I'm going to dismiss without argument *here*. Timothy Williamson claims the set-up is incoherent, because *H* is part of your evidence, and evidence is always true, so there is no situation where we have the evidence we actually have and *H* is false. Fred Dretske claims that the kind of closure principle that motivates premise 1 fails when dealing with 'heavyweight' propositions like the claim that you aren't in *S*. Sceptics of course say that the conclusion is true. Contextualists say that different reactions are appropriate to different tokens of the argument, and other epistemologists might say that different reactions are appropriate to different types of the argument, for instance accepting the conclusion when *H* is about the future, and taking the setup to be incoherent when *H* is something we can see to be true. Instead of surveying these options, I want to look at responses that question premise 2.

It is sometimes claimed that we should accept premise 2, or at least take it seriously, because it is *intuitive*. This might be backed up by a straw poll, either of students, or colleagues, or folks in Central Park, who share the intuition. The thought is that this intuition is just like the intuition that subjects in Gettier cases don't know certain things about their co-workers, or that various people do and don't know about Tom Grabit's thievery. It is a bedrock data point that we don't know we aren't in *S*. Now data can be misleading. There are such things as bad machine readings. So those who argue that premise 2 is a data point don't say that we have to accept it, but they do say that its being intuitive gives it a special epistemic status.

I used to agree. But now I think that taking all intuitive propositions to have this special status involves a subtle philosophical confusion. Not all intuitive propositions are alike, and the differences matter epistemologically. Some propositions are *self-evident*. I mean this quite literally; they are their own evidence. To go looking elsewhere for evidence for the propositions, for reasons to back them up, is to misjudge one's epistemological position. Bedrock claims in ethics and political philosophy seem like good cases of self-evident propositions. So I tend to think Jefferson got it exactly right in the opening lines of the Declaration of Independence in describing his ethical starting points (all men are equal, they have certain inalienable rights) as *self-evident*.[1]

What Jefferson called *self-evident*, some contemporary philosophers might call *intuitive*. But the intuitive well outruns the self-evident. It also includes what I'll dub the *reasonable*. That is, it includes

[1] I don't think it is self-evident that there is a creator, so I disagree with Jefferson to that extent. I also don't think that it is self-evident that there is a right to revolution when these inalienable rights are being violated, but I think it is a little tricky to say whether Jefferson is claiming that it is.

things we find compelling for (independent) reasons of a broadly philosophical nature. These reasons may not be ones that we can readily articulate, or even are consciously aware of. But in an important sense reasonable propositions are mental conclusions, not mental premises. We can tell the reasonable and the self-evident apart by looking at the appropriate reaction to those who disagree. If someone doesn't agree with what you find reasonable, you try to give them your reasons. If someone doesn't agree with what you find self-evident, you might make sure they've understood what you are saying, or make sure they aren't crazy, but pretty soon you might have to accept that some people just disagree. If the disagreement matters politically, it might be the point where discussion stops, and fighting begins.

Now in certain epistemological examples, that is exactly what happens. If two people just simply disagree about the right intuitive reaction to a particular case, they pretty soon won't have much to say. One might be able to convince the other that their intuition is misleading, but even that will rely on a shared stock of epistemic judgments. But the sceptical case is different. Every time I've seen a philosopher try to present something like the argument from sceptical cases, there's always a brief statement of a reason, often incomplete or hard to parse, backing up premise 2. If someone simply refuses to accept that premise 2 has any plausibility (as I'm disposed to do), the response is to offer reasons to accept it. This is notably not the reaction to someone who rejects Jeffersonian principles of equality, where it is hard to know how to even begin to reason. And even if you can offer reasons, I rather doubt you'll be offering your own reasons for believing in equality.

When a premise is self-evident, we are well justified in accepting it in argument, at least in the absence of defeating evidence. But in controversial cases, *reasonable* premises do not deserve such special treatment. When a reasonable premise leads us astray, the natural thing to do is to look to our reasons, and see if they were faulty ones. So what reasons can we give for accepting premise 2?

I have an intermittent interest in producing a complete taxonomy of the arguments that have been (less than clearly) offered in support of this premise. It's a bit of a job, because there are so many overlapping arguments. I think, though this is far from clear, that there are four major classes. Here (in the footnote) are the first three, with a representative example of each.

Discrimination Argument

1. Someone in *S* can't discriminate her situation from yours
2. Indiscriminability is symmetric
3. If you can't discriminate your situation from *S*, you can't know you're not in *S*
4. So you can't know you're not in *S*

Evidential (or Underdetermination) Argument

1. Someone in *S* has the same evidence as you do
2. What you can know supervenes on what your evidence is
3. So, you can't know you are not in *S*

Dialectical Argument

1. There is no non-circular argument to the conclusion that you aren't in *S*
2. If you were able to know you're not in *S*, you would be able to produce a non-circular argument that concluded that you aren't in *S*
3. So you can't know that you aren't in *S*

I won't say much about these arguments, save that I think in each case the second is very implausible. I suspect that most non-philosophers who are moved by sceptical arguments are tacitly relying on one or other of these arguments, but confirming that would require a more careful psychological study than I could do. And saying just which great philosophers had accepted which of these three arguments would require more knowledge of epistemology's history than I have. But set all that aside, because there's a fourth argument that is more troubling. This argument takes its inspiration from what we might call Hume's exhaustive argument for inductive scepticism. Hume said that we can't justify induction inductively, and we can't justify it deductively, and that *exhausts* the justifications, so we can't justify induction. A similar kind of argument helps out the general sceptic.

Exhaustive Argument

1. If you know you aren't in *S*, you know this a priori, or a posteriori on the basis of your evidence *E*
2. You can't know you aren't in *S* a posteriori on the basis of your evidence *E*
3. You can't know you aren't in *S* a priori
4. So, you can't know you aren't in *S*

This seems to be a really interesting argument to me. To make things simpler, I'll stipulate that by a posteriori knowledge, I just mean knowledge that isn't a priori. That makes the first premise pretty secure, as long as we're assuming classical logic. Lots of philosophers take its third premise for granted. In fact, when I try to argue the third premise is false, several people look at me as if I'd endorsed modal realism. But I won't push that line here. Instead I'll look at denials of the second premise.

2. Dogmatism and a Bayesian Objection

Someone who denies the second premise says that your empirical evidence can provide the basis for knowing that you aren't in S, even though you didn't know this a priori. I'm going to call any such person a *dogmatist*. This is a considerably broader use of the term than is standard, but my reasons for this semantic heterodoxy will become clear shortly. The dogmatist is not a sceptic, so the dogmatist believes that you can know H on the basis of E. The dogmatist also believes a closure principle, so the dogmatist also believes you can know $E \supset H$. If the dogmatist thought you could know $E \supset H$ a priori, they'd think that you could know a priori that you weren't in S. (This follows by another application of closure.) But they think that isn't possible, so knowing $E \supset H$ a priori isn't possible. Hence you know $E \supset H$ a posteriori, and since E is your evidence, you know it on the basis of your evidence E.

If we reflect on the fact that E is your total evidence, then we can draw two conclusions. The first is that the dogmatist thinks that you can come to know H on the basis of E even though you didn't know in advance that if E is true, then H is true. You don't, that is, need *antecedent* knowledge of the conditional $E \supset H$ in order to be able to learn H from E. That's why I'm calling them a dogmatist. You can learn H on the basis of E without antecedently knowing that E is, even materially, indicative of H's truth. The second point is that the dogmatist is now running head on into a piece of Bayesian orthodoxy.

To see the problem, note that we can easily prove (A), for arbitrary E, H and K.[2]

(A) $Pr(E \supset H \mid E \wedge K) \leq Pr(E \supset H \mid K)$, with equality iff $Pr(E \supset H \mid E \wedge K) = 1$

[2] Proof:

1.	$E \supset H \dashv\vdash ((E \supset H) \wedge E) \vee ((E \supset H) \wedge \neg E)$	Logic
2.	$Pr(E \supset H \mid K) = Pr(((E \supset H) \wedge E) \vee ((E \supset H) \wedge \neg E) \mid K)$	1
3.	$Pr(((E \supset H) \wedge E) \vee ((E \supset H) \wedge \neg E) \mid K) = Pr((E \supset H) \wedge E \mid K) + Pr((E \supset H) \wedge \neg E \mid K)$	Prob axiom
4.	$Pr((E \supset H) \wedge E \mid K) = Pr(E \supset H \mid E \wedge K)\, Pr(E \mid K)$	Prob axiom
5.	$Pr((E \supset H) \wedge \neg E \mid K) = Pr(E \supset H \mid \neg E \wedge K)\, Pr(\neg E \mid K)$	Prob axiom
6.	$Pr(E \supset H \mid K) = Pr(E \supset H \mid E \wedge K)\, Pr(E \mid K) + Pr(E \supset H \mid \neg E \wedge K)\, Pr(\neg E \mid K)$	2, 3, 4, 5
7.	$Pr(E \supset H \mid \neg E \wedge K) = 1$	Prob axiom
8.	$Pr(E \supset H \mid E \wedge K) \leq 1$	Prob axiom
9.	$Pr(E \supset H \mid K) \geq Pr(E \supset H \mid E \wedge K)\, Pr(E \mid K) + Pr(E \supset H \mid E \wedge K)\, Pr(\neg E \mid K)$	6, 7, 8
10.	$Pr(E \mid K) + Pr(\neg E \mid K) = 1$	Prob axiom
11.	$Pr(E \supset H \mid K) \geq Pr(E \supset H \mid E \wedge K)$	9, 10

Note that line 1 is only valid in classical logic. In fact (A) is not a theorem of the constructivist probability calculus that I discussed in earlier work. There is an interesting question here of whether the dogmatist could successfully avoid the Bayesian objection by changing logics, but I won't go into that here. I'm grateful to David Jehle for these observations.

It is clear enough from the proof that line 6 is an equality iff line 3 is an equality, so we have proven (A). Now some authors have inferred from this something like (B) from (A).[3]

(B) It is impossible to go from not being in a position to know $E \supset H$ to being in a position to know it just by receiving evidence E.

The transition here should raise an eyebrow or two. (A) is a principle of probability statics. (B) is a principle of epistemological dynamics. To get from (A) to (B) we need a principle linking probability and epistemology, and a principle linking statics and dynamics. Fortunately, orthodox Bayesian confirmation theory offers us suggestions for both principles. We'll write $Cr(A)$ for the agent's credence in A, and $Cr_E(A)$ for the agent's credence in A when updated by receiving evidence E.

LEARNING. If $Cr_E(A) \leq Cr(A)$, then it is impossible to go from not being in a position to know A to being in a position to know it just by receiving evidence E.

BAYES. $Cr_E(A) = Cr(A \mid E)$. That is, learning goes by conditionalisation.

A quick browse at any of the literature on Bayesian confirmation theory will show that these principles are both widely accepted by Bayesians. I'm going to accept LEARNING at least for the sake of argument. I'm going to argue instead that the inference from (A) to (B) fails because BAYES fails. But the mere falsity of BAYES would not be enough to save the dogmatist. We don't need BAYES to be true in its full generality. A single instance of it would do. If a principle I'll call LOWER is true, then dogmatism in the sense I'm defending it fails.

LOWER. $Cr_E(E \supset H) \leq Cr(E \supset H)$.

LOWER follows from a particular instance of BAYES, namely that $Cr_E(E \supset H) = Cr(E \supset H \mid E)$, and principle (A). It is sufficient to run the argument against dogmatism. On the other hand, if LOWER is false then the argument fails. I'm going to argue that LOWER is indeed false.

Now there is a bad argument around here that the dogmatist might make. It might be argued that since the Bayesian approach (including BAYES) involves so much idealisation it could not be applicable to real agents. That's a bad argument because the Bayesian approach might provide us with a good model for real agents, and models can be useful without being scale models. As long as the Bayesian model is

[3] Roger White and Stewart Cohen endorse probabilistic arguments against people who are, in my sense, dogmatists. John Hawthorne also makes a similar argument when arguing that certain conditionals, much like $E \supset H$, are a priori.

the most appropriate model in the circumstances, then we can draw conclusions for the real world from facts about the model. The problem arises if there are alternative models which seem to fit just as well, but in which principles like LOWER are not true. If there are alternative models that seem better suited (or at least just as well suited) to modelling the situation of initial evidence acquisition, and those models do not make LOWER true, then we might think the derivation of LOWER in the Bayesian model is a mere consequence of the arbitrary choice of model. In the next section I will develop just such a model. I won't argue that it is the best model, let alone the only alternative to the Bayesian model. But it does not imply LOWER, and I will argue that it is as good for these purposes as the Bayesian model.

3. Bayes and Keynes

The traditional Bayesian model of a rational agent starts with the following two principles.

- At any moment, the agent's credal states are represented by a probability function.
- From moment to moment, the agent's credal states are updated by conditionalisation on the evidence received.

Over recent decades many philosophers have been interested in models that relax those assumptions. One particular model that has got a lot of attention (from e.g. Isaac Levi, Richard Jeffrey, Bas van Fraassen, Alan Hájek and many others) is what I'll call the *static Keynesian model*. This model has the following features.

- At any moment, the agent's credal states are represented by a set of probability functions, called their representor.
- The agent holds that *p* is more probable than *q* iff the probability of *p* is greater than the probability of *q* according to all probability functions in their representor. The agent holds that *p* and *q* are equally probable iff the probability of *p* is equal to the probability of *q* according to all probability functions in their representor.
- From moment to moment, the agent's credal states are updated by conditionalising each of the functions in the representor on the evidence received.

The second point is the big attraction. It allows that the agent need not hold that *p* is more probable than *q*, or *q* more probable than *p*, or that *p* and *q* are equally probable, for arbitrary *p* and *q*. And that's good because it isn't a rationality requirement that agents make pairwise probability judgments about all pairs of propositions. Largely because of this feature, I argued in an earlier paper that this model could be use to formalise the key philosophical ideas in Keynes's *Treatise on Probability*. That's the reason I call this a 'Keynesian' model.

The modifier ‘static’ might seem a little strange, because the agent’s representor does change when she receives new evidence. But the change is always of a certain kind. Her ‘hypothetical priors’ do not change. If at t_1 her evidence is E_1 and her representor R_1, and at t_2 her evidence is E_2 and her representor R_2, then there is a ‘prior’ representor R_0 such that the following two claims are true for all probability functions *Pr*.

- $Pr \in R_1 \leftrightarrow [\exists \mathrm{Pr}_0 \in R_0\text{: } \forall p\ (Pr(p) = \mathrm{Pr}_0(p \mid E_1))]$
- $Pr \in R_2 \leftrightarrow [\exists \mathrm{Pr}_0 \in R_0\text{: } \forall p\ (Pr(p) = \mathrm{Pr}_0(p \mid E_2))]$

That is, there is a set of probability functions such that the agent’s representor at any time is the result of conditionalising each of those functions on her evidence. I’ll call any model with this property a static model, so the model described above is the static Keynesian model.

Now there is a lot to like about the static Keynesian model, and I have made extensive use of it previous work. It is a particularly useful model to use when we need to distinguish between risk and uncertainty in the sense that these terms are used in Keynes’s 1937 article “The General Theory of Employment”.[4] The traditional Bayesian model assumes that all propositions are risky, but in real life some propositions are uncertain as well, and in positions of radical doubt, where we have little or no empirical evidence, presumably most propositions are extremely uncertain. And using the static Keynesian model does not mean we have to abandon the great work done in Bayesian epistemology and philosophy of science. Since a Bayesian model is a (degenerate) static Keynesian model, we can say that in many circumstances (namely circumstances where *uncertainty* can be properly ignored) the Bayesian model will be appropriate. Indeed, these days it is something like a consensus among probabilists or Bayesians that the static Keynesian model is a useful generalisation of the Bayesian model. For example in Christensen 2005 it is noted, almost as an afterthought, that the static Keynesian model will be more

[4] The clearest statement of the distinction that I know is from that paper.

> By ‘uncertain’ knowledge, let me explain, I do not mean merely to distinguish what is known for certain from what is only probable. The game of roulette is not subject, in this sense, to uncertainty; nor is the prospect of a Victory bond being drawn. Or, again, the expectation of life is only slightly uncertain. Even the weather is only moderately uncertain. The sense in which I am using the term is that in which the prospect of a European war is uncertain, or the price of copper and the rate of interest twenty years hence, or the obsolescence of a new invention, or the position of private wealth owners in the social system in 1970. About these matters there is no scientific basis on which to form any calculable probability whatever. We simply do not know. Nevertheless, the necessity for action and decision compels us as practical men to do our best to overlook this awkward fact and to behave exactly as we should if we had behind us a good Benthamite calculation of a series of prospective advantages and disadvantages, each multiplied by its appropriate probability, waiting to be summed. (114-5)

realistic, and hence potentially more useful, than the traditional Bayesian model. Christensen doesn't appear to take this as any kind of *objection* to Bayesianism, and I think this is just the right attitude.

But just as the static Keynesian is more general than the Bayesian model, there are bound to be interesting models that are more general than the static Keynesian model. One such model is what I call the *dynamic* Keynesian model. This model has been used by Seth Yalcin to explicate some interesting semantic theories, but to the best of my knowledge it has not been used for epistemological purposes before. That should change. The model is like the static Keynesian model in its use of representors, but it changes the way updating is modelled. When an agent with representor *R* receives evidence *E*, she should update her representor by a two step process.

- Replace *R* with U(*R*, *E*)
- Conditionalise U(*R*, *E*), i.e. replace it with {*Pr*(· | *E*): *Pr* is in U(*R*, *E*)}

In this story, U is a function that takes two inputs: a representor and a piece of evidence, and returns a representor that is a subset of the original representor. Intuitively, this models the effect of learning, via getting evidence *E*, what evidential relationships obtain. In the static Keynesian model, it is assumed that before the agent receives evidence *E*, she could already say which propositions would receive probabilistic support from *E*. All of the relations of evidential support were encoded in her conditional probabilities. There is no place in the model for learning about fundamental evidential relationships. In the dynamic Keynesian model, this is possible. When the agent receives evidence *E*, she might learn that certain functions that were previously in her representor misrepresented the relationship between evidence and hypotheses, particularly between evidence *E* and other hypotheses. In those cases, U(*R*, *E*) will be her old representor *R*, minus the functions that *E* teaches her misrepresent these evidential relationships.

The dynamic Keynesian model seems well suited to the dogmatist, indeed to any epistemological theory that allows for fundamental evidential relationships to be only knowable a posteriori. There's something very odd about the idea that learning should always go by conditionalisation. It implies that we start inquiry with a complete epistemological game plan, that our conditional probabilities reflect a strategy for dealing with whatever evidence we get. This is obviously an idealisation, and it is interesting to look at models that don't have such a property, that (quite reasonably) allow agents to adjust the plan as the game develops. As we might put it in (American) football terms, there is nothing wrong with sometimes having to call an audible. And which audible to call depends on what the data are. More formally, anyone who thinks that any fundamental epistemological relations can only be known a posteriori should be interested in having such a model that allows for this. The dynamic Keynesian model is one such model.

4. In Defence of Dynamism

In this final section I want to argue for three things. First, the dynamic Keynesian model has a distinctive virtue, the virtue of all theories that say our knowledge that sceptical possibilities do not obtain is a posteriori. Second, the dynamic Keynesian theory does not have a vice you might think it has; it doesn't make us throw out all philosophical worked based on conditionalisation. Third, once we take the model seriously, we see that the dogmatist has an answer to the objection from LEARNING and LOWER. I'll also briefly mention some other virtues the model has and vices it lacks.

4.1 The Problem of the Priors

One really nice consequence of the dynamic Keynesian approach is that it lets us say what the representor of an agent with no empirical information should be. Say a proposition is *a priori certain* iff it is a priori that all rational agents assign credence 1 to that proposition. Then the representor of the agent with no empirical evidence is $\{Pr: \forall p$: If p is a priori certain, then $Pr(p) = 1\}$. Apart from assigning probability 1 to the a priori certainties, the representor is silent. Hence it treats all propositions that are not a priori certain in exactly the same way. This kind of symmetric treatment of propositions is not possible on the traditional Bayesian conception for logical reasons. (The reasons are set out in the various discussions of the paradoxes of indifference, going back to Bertrand 1888.) Such a prior representor is strictly speaking consistent with the static Keynesian approach, but it yields highly implausible results, since it does not allow learning.[5] So on the static Keynesian model, this attractively symmetric prior representor is not available.

I think one of the motivations of anti-dogmatist thinking is the thought that we *should* be able to tell a priori what is evidence for what. If it looking like there is a cow in front of us is a reason to think there is a cow in front of us, that should be knowable a priori. I think the motivation for this kind of position shrinks a little when we realise that an a priori prior that represented all the connections between evidence and hypotheses would have to give us a lot of guidance as to what to do (epistemically speaking) in worlds quite unlike our own. Moreover, there is no reason we should have lots that information. So consider, for a minute, a soul in a world with no spatial dimensions and three temporal dimensions, where the primary source of evidence for souls is empathic connection with other souls from which they get a (fallible) guide to those souls' mental states. When such a soul conditionalises on the evidence "A soul

[5] More precisely, conditionalising on E has no effect on the distribution of values of Pr(p) among functions in the representor for any p not made a priori certain by E. (We'll say p is made a priori certain by E iff $E \supset p$ is a priori certain.) So if this is our starting representor, we can't even get probabilistic evidence for things that are not made certain by our evidence. (The argument here goes by a little quickly, because I've defined representors in terms on unconditional probabilities and this leads to complications to do with conditionalising on propositions of zero probability. A better thing to do, as suggested by Hájek 2003, is to take conditional probability as primitive. If we do this we'll define representors as sets of conditional probability functions, and the a priori representor will be {Pr: If $p \supset q$ is a priori certain, then $Pr(q \mid p) = 1$}. Then everything I said above will be true.)

seems to love me" (that's the kind of evidence they get) what should their posterior probability be that there is indeed a soul that loves them? What if the souls have a very alien mental life, so they instantiate mental concepts very unlike our own, and souls get fallible evidence of these alien concepts being instantiated through empathy? I think it's pretty clear we *don't* know the answers to these questions. Note that to answer this question we'd have to know which of these concepts were grue-like, and which were projectable, and there is no reason to believe we are in a position to know that.

Now those souls are presumably just as ignorant about the epistemologically appropriate reaction to the kinds of evidence we get, like seeing a cow or hearing a doorbell, as we are about their evidence. The dynamic Keynesian model can allow for this, especially if we use the very weak prior representor described above. When we get the kind of evidence we actually get, the effect of U is to shrink our representors to sets of probability functions which are broadly speaking epistemically appropriate for the kind of world we are in. Before we got that evidence, we didn't know how we should respond to it, just like the spaceless empathic souls don't know how to respond to it, just like we don't know how to respond to their evidence.

It is a commonplace observation that (a) prior probabilities are really crucial in Bayesian epistemology, but (b) we have next to no idea what they look like. I call this the problem of the priors, and note with some satisfaction that the dynamic Keynesian model avoids it. Now a cynic might note that all I've done is replace a hand-wavy story about priors with a hand-wavy story about updating. That's true, but nevertheless I think this is progress. The things I'm being deliberately unclear about, such as what U should look like for *E* such as "Some other non-spatial tri-temporal soul seems to love me" are things that (a) my theory says are not a priori knowable, and (b) I don't have any evidence concerning. So it isn't surprising that I don't have much to say about them. It isn't clear that the traditional Bayesian can offer any story, even by their own lights, as to why they are less clear about the structure of the prior probability conditional on such an *E*.

4.2 Fertility Arguments

It might be worried that the great success of the Bayesian model in Bayesian philosophy of science over the last few decades gives us a reason to keep it. I tend to be very impressed by fertility arguments in philosophy. The best way to see that some kind of model works is to see it work, after all. If adopting the dynamic Keynesian required us to undo all the progress that Bayesian philosophers have made, that would be a strong argument against adopting it. Fortunately, this is not the case.

Although the dynamic Keynesian model is different to the traditional Bayesian model, it is not strictly speaking a competitor to the Bayesian model. Static Keynesian models are special cases of dynamic Keynesian models, and traditional Bayesian models are special cases of static Keynesian models. So even if in general we should model learning using a dynamic Keynesian model, that does not suggest that it is wrong to ever use traditional Bayesian models. It just has to be the case that (a) there are some conditions such that when they are satisfied, we don't need to extra generality the dynamic Keynesian model provides, and (b) those conditions are satisfied in the particular application we are looking at.

In general, we can use a static Keynesian model rather than a dynamic Keynesian model if the situation being modelled does not involve learning about fundamental epistemological relationships. That is, if the situation being modelled is not such that the new evidence agents get provides them with justification for epistemological theses they were previously not justified in believing. For what U represents is just this kind of learning; learning what kinds of hypotheses are justified by what evidence. And we can use a traditional Bayesian model rather than a static Keynesian model if the distinction between risk and uncertainty is not relevant to the case. Now it is very plausible that in the cases where Bayesian modelling has proven most useful, such as in explaining away Hempel's paradox of confirmation, those two conditions are satisfied. But we should not expect that the conditions will *always* be satisfied.

In particular, in the case at issue here, *neither* condition is plausibly satisfied. Unless we assume that facts about which evidence justifies belief in which propositions is a priori, we should think this *is* a situation where getting new evidence involves epistemological, not just empirical, learning. And if there is any situation where the risk/uncertainty distinction matters, it is the situation the Bayesian is trying to model with their hypothetical priors, i.e. the situation of an agent with no empirical evidence. So there is strong reason to think that, even if we are sympathetic to a broadly Bayesian approach to issues in philosophy of science, that we need a more general approach here. But if we use more general models, such as dynamic Keynesian models, the anti-dogmatist argument does not go through. So I conclude that this anti-dogmatist argument relies on improperly applying models that are appropriate to everyday cases, i.e. Bayesian models, to a situation not at all like the everyday, i.e. the situation of receiving for the first time empirical evidence about the external world.

4.3 Other Considerations

In this section I briefly mention two virtues of the dynamic Keynesian model, and note how I'd reply to two worries people might have about the model. For space reasons, this will be very compressed.

Because the dynamic Keynesian model does not verify LOWER, it is consistent with dogmatism. Now if you think that dogmatism is obviously false, you won't think this is much of an advantage. But I tend to think that dogmatism is one of the small number of not absurd solutions to a very hard epistemological problem with no obvious solution, so we should not rule it out pre-emptively. Hence I think our formal models should be consistent with it. So it's a very good thing that the dynamic

Keynesian model is consistent with it. Proving this is somewhat difficult, so I've relegated that to a footnote.[6]

The way the dynamic Keynesian model is set up, the ideal agent does two things when they get evidence *E*; cull the representor and conditionalise it. In practice, these might happen at different times. I think what happens when we see that old evidence supports a new theory is that we engage in some representor culling that we were previously entitled to do, but had not previously done. That's why old evidence can support new hypotheses. I don't want to claim this is the only solution to the notorious problem of old evidence, but on the other hand I don't think any traditional Bayesian solution to that problem is anywhere near as simple.

---

[6] To see whether this is true on the dynamic Keynesian model, we need to say what it is to *lower* the credence of some proposition. Since representors map propositions onto intervals rather than numbers, we can't simply talk about one 'probability' being a smaller number than another. (Strictly speaking, the story I've told so far does not guarantee that for any proposition p, the values that Pr(p) takes (for Pr in the representor) form an interval. But it is usual in more detailed presentations of the model to put constraints on the representor to guarantee that happens, and I'll assume we've done that.) On the static Keynesian model, the most natural move is to say that conditionalisation on *E lowers* the credence of *p* iff for all *Pr* in the representor, $Pr(p) > Pr(p \mid E)$. This implies that if every function in the representor says that *E* is negatively relevant to *p*, then conditionalising on *E* makes *p* less probable. Importantly, it allows this even if the values that $Pr(p)$ takes across the representor before and after conditionalisation overlap. So what should we say on the dynamic Keynesian model? The weakest approach that seems viable, and not coincidentally the most plausible approach, is to say that updating on *E* lowers the credence of *p* iff the following conditions are met:

- For all *Pr* in $U(R, E)$, $Pr(p \mid E) < Pr(p)$
- For all *Pr* in *R* but not in $U(R, E)$, there is a Pr' in $U(R, E)$ such that $Pr'(p \mid E) < Pr(p)$

It isn't too hard to show that for some models, updating on *E* does not lower the credence of $E \supset H$, if lowering is understood this way. The following is an extreme example, but it suffices to make the logical point. Let *R* be the minimal representor, the set of all probability functions that assign probability 1 to a priori certainties. And let $U(R, E)$ be the singleton of the following probability function, defined only over Boolean combinations of *E* and *H*: $Pr(E \wedge H) = Pr(E \wedge \neg H) = Pr(\neg E \wedge H) = Pr(\neg E \wedge \neg H) = ¼$. Then the probability of $E \supset H$ after updating is ¾. (More precisely, according to all *Pr* in $U(R, E)$, $Pr(E \supset H) = ¾$.) Since before updating there were *Pr* in *R* such that $Pr(E \supset H) < ¾$, in fact there were *Pr* in *R* such that $Pr(E \supset H) = 0$, updating on *E* did not *lower* the credence of $E \supset H$. So the dynamic Keynesian model does not, in general, have as a consequence that updating on *E* lowers the credence of $E \supset H$. This suggests that LOWER in general is not true.

It might be objected that if evidence *E* supports our knowledge that $E \supset H$, then updating on *E* should *raise* the credence of $E \supset H$. And if we define credence raising the same way we just defined credence lowering, updating on *E never* raises the credence of $E \supset H$. From a Keynesian perspective, we should simply deny that evidence has to raise the credence of the propositions known on the basis of that evidence. It might be sufficient that getting this evidence removes the uncertainty associated with those propositions. Even on the static Keynesian model, it is possible for evidence to remove uncertainty related to propositions without raising the probability of that proposition. A little informally, we might note that whether an agent with representor *R* is sufficiently confident in *p* to know that *p* depends on the lowest value that $Pr(p)$ takes for $Pr \in R$, and updating can raise the value of this 'lower bound' without raising the value of $Pr(p)$ according to all functions in *R*, and hence without strictly speaking *raising* the credence of *p*.

There are several arguments for conditionalisation, e.g. the Teller-Lewis Dutch Book argument and the Greaves-Wallace epistemic utility argument. Every such argument that I know of assumes that credences are numerical, i.e. assumes that we are working in a Bayesian rather than a Keynesian framework. So they couldn't be arguments for having a static rather than a dynamic Keynesian model. So I don't think there are any direct arguments for the superiority of the static model.

Finally, there might of course be models more general than the dynamic Keynesian model. There are many idealisations in the model. Some of them, such as logical omniscience, are obviously idealisations. Some of them, such as the assumption that we can separate epistemological learning from factual learning, are more contentious. But I'm sure there will be cases where a more general model than the dynamic Keynesian model is needed. That in no way undermines the usefulness of the dynamic Keynesian model, or restores the Bayesian argument against dogmatism.

4.7 Why Should We Care?

The sceptic's opening move was to appeal to our intuition that propositions like $E \supset H$ are unknowable. The response was that this is something that we should only accept for reasons, because it is not self-evident. The sceptic can respond with a wide range of arguments, four of which are mentioned above. Here we focussed on the sceptic's argument from exhaustion. $E \supset H$ isn't knowable a priori, because it could be false, and it isn't knowable a posteriori, because, on standard models of learning, our evidence *lowers* its credibility. My response is to say that this is an artefact of the model the sceptic (along with everyone else) is using. There's nothing wrong with using simplified models, in fact it is usually the only way to make progress, but we must be always wary that our conclusions transfer from the model to the real world. One way to argue that a conclusion is a mere artefact of the model is to come up with a model that is sensitive to more features of reality in which the conclusion does not hold. That's what I've done here. The dynamic Keynesian model is sensitive to the facts that (a) there is a distinction between risk and uncertainty and (b) we can learn about fundamental evidential connections. In the dynamic Keynesian model, it isn't true that our evidence lowers the probability of $E \supset H$. So the anti-sceptic who says that $E \supset H$ is knowable a posteriori, the person I've called the dogmatist, has a defence against this Bayesian argument. If the response is successful, then there may well be other applications of the dynamic Keynesian model, but for now I'm content to show how the model can be used to defend the dogmatic response to scepticism.