Comments for Stefan

General: Stefan, can I ask about the order of the chapters? I can see that 1-3 must come first. Beyond that, the chapters are more or less independent essays. What is the rationale for the order in which you put them?

Here is an alternative ordering: put the counterexamples first, then the discussion of vague credences, then the discussion of geometry of reason (which is the most open-ended chapter and leads naturally into the conclusion and prospects for further work). So:

Ch. 4 = Wagner
Ch. 5 = Judy Benjamin
Ch. 6 = Augustin's concessions
Ch. 7 = Geometry of Reason

Your Introduction should have a section that offers a brief road map of the chapters. This way, we have:

Chaps 2-3: explaining the debate about PME, providing the basics about updating partial beliefs and the basic ideas of information theory, discussion of justification.

Chaps 4-5: defusing two leading counterexamples. [Explain in the Intro why you pick these two: each connects to one or two specific broader issues, e.g., epistemic entrenchment for JB. Thus, they are not merely accomplishing the "negative task" of defusing the objection – they are also showing how we can think about these broader issues in a positive and coherent way using information theory.] The Wagner should have a forward reference to the "Augustin's concessions" chapter.

Chaps 6-7: Analysis of two major "current debates" connected to information theory (and where bringing information theory to bear provides illumination) – vague credences and geometry of reason.

Chapter 8: Conclusion.

Chapter 1

General comments:

The chapter is short, but complicated enough that it might benefit from sub-sections. I think it also needs some re-organization. Themes are taken up, set aside, then taken up again. The strands should be separated where this is possible.

1.1 Information theory and probability kinematics

p. 1 First paragraph should just be sentence 1 (also – do vague credences belong here as a major claim, since they get just one chapter?). Put "A Bayesian framework…" with the second paragraph. But then I would try to add something like what you have at the bottom of p. 3 (the "powerful and simple idea" that is the core intuition of information theory), since that motivates the whole thesis.

The entire final paragraph (running on pages 1-2) should be a footnote to the first sentence of paragraph 3. As it stands, this interrupts the flow.

The discussion on pages 1-2 is about updating. Then you have two paragraphs that distinguish between the problem of priors and the problem of updating. And then you go back to updating and PME. I suggest that, right around the start of paragraph 3, you introduce the distinction between the two problems, even before you start talking about PME, and indicate that you will focus on the latter. Move the two paragraphs on priors to this spot.

Top of p. 4: Good place to start a new section

Section 1.2 The Principle of Maximum Entropy (PME): a short history.

Then end with:

Section 1.3 Dissertation outline

Here, you can group together some of the material on pages 4-5 with the material on pages 7-9 in your chapter outline.

Specific comments:

p. 3 Reference to the 'full employment theorem' is too cryptic – they won't know what you mean.

p. 4 van Fraassen
"probabilities (or ranks)": should "or ranks" be in a footnote, which also mentions that there are lots of ways to represent uncertainty?
Not sure about the Susan Haack reference, if you can't find a text.

End the sentence about Haack earlier: "years ago. This issue is also related…
"… the double task for indeterminate credences…": too cryptic. How about: "…also related to a dilemma for indeterminate credences that I develop in chapter NN: they cannot perform the double task of representing both epistemic uncertainty and all relevant features of the evidence."
Last sentence of this paragraph ("A large part"…) belongs with the next paragraph.

pp. 5-6 How does this material relate to the thesis?

p. 6 At the bottom of p. 6, you go back to epistemic entrenchment. Again, you should find a way to group this material together.

p. 7 Section 1.3 (Dissertation outline) should start here. But you need to say something about chapters 2 and 3.

p. 8 "…indeterminate creedal states."
Note: you brought up counterexamples, then you moved beyond counterexamples, and now you are back to counterexamples. Hence my earlier suggestion to group this with material on pages 4-5.

p. 9 First reference to JUP – won't be obvious to the reader that this is Jeffrey conditioning.

Chapter 2

General comments:

Most important: in a chapter that introduces PME and discusses its history, there is no statement of PME! Either it needs to be here (in section 2.2), or you need to explain why you are deferring this to a later chapter.

In the same light: do you want a brief review of standard conditioning and maybe even Jeffrey conditioning in this chapter (instead of appearing first in chapter 3 – it seems odd to consider them first and foremost as principles of information theory, as they appear in sections 3.2.4 and 3.2.5)?

It might be good to consolidate some of the things here with similar things in chapter 1 (e.g., discussion of epistemic entrenchment; reference to 'full employment').

Minor but annoying: most of the references are coming out: DATE,), as in (de Finetti, 1931, and 1937,). Should end as: …1937). This happens throughout the chapter.

Specific comments:

p. 10 …19th and 20th century bring us two responses…
delete and on decision problems
Some references on this page are partial – you just have a list of names, but no dates.
…justifies entertaining…

p. 11 …begins with John Maynard…
What about Carnap (under logical interpretation)?
Question: since you have indicated that you aren't going to talk about priors, does the contrast between Jeffreys and de Finetti make any difference to your project? That is: couldn't they both agree that there is a *logic* for updating? (I see that you clarify this point on the very next page: you don't care about priors.) However: did you not go over this distinction, and your decision to focus only on updating, in the Introduction? Maybe you need to check for repetition.
Why a Pyrrhic victory?

p. 12 Missing footnote TBD.
Alan H also considers conditional probability …
"on page 315": what does this mean?

p. 13 (For a more systematic… objectivism.) THIS SHOULD BE A FOOTNOTE.
Not sure the analogy with geometry helps – it's such a complex topic, and you have already given the analogy with deductive logic.

p. 13-14 Again, it seems odd to have this whirlwind survey of techniques for justifying standard conditioning and Jeffrey conditioning, without ever having reivewed (even very briefly) what these are. In one page, you could insert them

here as the core (at least, standard conditioning) of Bayesian updating and hence the best Bayesian candidate for inductive logic.

p. 14 Final remark about the logical interpretation is missing a footnote. But I'm not sure what this final paragraph of section 2.1 is doing here, since you have a) just been talking about ideas for justifying Bayesian updating principles, and b) you have been stressing that you are not committed to the full program of logical probability, and that what you are committed to is largely shared by the logical approach and the subjective approach.

p. 15 You need some definitions here (entropy, a statement of PME), or else you need to explain why they aren't here and will be given later.
There is a breathless quality about this section 2.2. You skip through the names very quickly but not really pausing long enough to explain things. If you have the chance, you can add some content.
"two epistemological dimensions": can you state what they are?

p. 16-17 I just want to say the same thing again. I seem to remember you explaining very well, somewhere, the sequence from standard conditioning to Jeffrey conditioning (a generalization) and then to PME (another generalization). A reader who starts this section expecting to be enlightened about information theory and PME will be puzzled not to see these things.

p. 17 "By then": can you explain what this means, or give a date?
"<u>has</u> been launched"? Actually, this is a good place to mention that the choice of the present tense for a chapter that is recounting a sequence of develops is a little awkward. It may be that you have no time now to change this – it is not that bad, and this is low priority, but I found it sometimes got in the way of my focusing on the ideas.
This would be a good place to distinguish between criticism that takes the form of counterexamples (Shimony, Judy Benjamin, Wagner) and more theoretical opposition (Uffink and Seidenfeld -- what is Seidenfeld's complaint, actually? – you don't tell us).
It's interesting that the late criticism does not produce any new counterexamples! Perhaps none are thought to be needed. But it's worth noting this.

p. 18 Again, this part is a bit breathless. Would it be possible, instead of running through the names quickly, instead to identify the main criticisms and devote a paragraph to each one? You could make formatting work for you: use bullet points or even numbered paragraphs, as in:

1) **Coarsening at random.** Grunwald…
2) **Excessive apriorism.** This is the concern…

Finally, why not mention epistemic entrenchment here as a third major late criticism (picking up on Judy Benjamin)? This would provide a bridge to the next

sub-section – otherwise, it's not clear why you are suddenly bringing in acceptance vs. probabilistic belief.

P 19 Section 2.4: you need to add an opening paragraph that motivates this section. Where does information theory fit in? Is it about acceptance, partial belief, or both? Somehow, we need to see why the debate is relevant for you. We get there by the last couple of sentences, but we need to see something at the start.

p. 20 I would divide the final paragraph of section 2.4 into two paragraphs, with the last one explaining how you think information theory can make a contribution to the reconciliation project.

p. 20, section 2.5

You need to explain something about why determinate vs. imprecise credences is a big distinction for your project. Maybe start out by framing the issue in terms of representation of uncertainty: then we have the two issues about priors – whether they are sharp, and whether they are objective.

Up to this point, you've wanted to be very non-commital about priors, so it comes as a surprise that you are so strongly wedded to determinate priors. Is this a form of apriorism? Maybe you can explain that it is not!

p. 21 …much as Jaynes intended it...

…of probabilities; see for example Zabell [insert semi-colon]

p. 22 …making one of information theory or probability…

…not robust, neither in quantity nor in quality… : what does this mean?

"information theory delivers the unique and across the board successful candidate for an objective updating mechanism in probability kinematics: PME and the companion principle of minimum cross-entropy. [you need to state what the candidate is]

The paragraph on Carnap may be too much detail in the wrong place. Just think about it.

p. 23 Regarding Shimony: maybe you can state why you aren't going to talk about his example.

… a good example of the kind of cross-fertilization…

A nice discussion of epistemic entrenchment – should this have come earlier on, since you have referred to epistemic entrenchment a number of times?

p. 24-5 The reference to full employment is a bit repetitive of what you said in chapter 1.

Chapter 3

p. 26 It would be good to add an opening paragraph that explains what you'll be doing in this chapter.

p. 26 "non-negative number representing partial belief": at the start of the paragraph, it sounded as if you would give clear distinction between full and partial belief, but this is not a definition of partial belief.
Instead of $\Omega = -(X \cap Y)$, why not use $X \cap Y = \phi$?

p. 27 (3.1) – add, provided $P(Y) \neq 0$

p. 28 Holmes' evidence [add apostrophe]

p. 31 ...the updated $(y', z') = (y'_1, z'_1)$ is continuous...

... to determine $(y', z') = (y'_2, z'_2)$.

p. 32 "...contra Landes and Williamson, who use this issue to stick the label of irrationality to PME in an otherwise solid defense". NOT CLEAR what you mean. Is that criticism that PME can't deal with a non-affine constraint?

...In the next subsection, I provide a sketch of the information-theoretic approach.

p. 33 Why not tell us what H_n is at the outset? "Let H_n be the entropy associated with a probability distribution P_n"

p. 34 Statement of "Original PME" is slightly awkward. Try: A rational agent accepts for her partial beliefs a probability distribution that satisfies evidential constraints and has maximal Shannon entropy.

p. 36 You should probably flag that this is your definition of PME (in contrast to p. 34) by adding REFINED PME (or something like this) at the start, and you should add a statement strongly emphasizing that all future references to PME refer to this definition.

Substantively, I'm confused about what PME means. We have: a) a principle about priors; b) a principle of updating; and c) both. The email I just sent you expressed confusion about whether your usage is consistent between a) and c), but here I see you have defined PME to mean b). And it may be that you are consistent. Perhaps just flagging the definition (as I suggested) is adequate.

Yet on the very next page (37-38), the discussion proceeds as if PME is just the synchronic norm.

But since you here accept a version of PME (maximal entropy) for synchronic contexts that are not updating, you seem to be at odds with the position you took earlier indicating that you would discuss only updating.

p. 38 Good discussion of compatibility.

Last line: …I will present Jaynes' method… [add apostrophe]

p. 39 The illustration of (3.19) with Judy Benjamin is so compressed as to be obscure.

p. 40 What is the function g?

p. 42 Should be: $\lambda = 1 - \ln \sum y_i$

p. 43 line 2: remember…

Actually, "remember" is not appropriate here. Try: (noting that in section 5.4, (5.11) is no longer required).

p. 44 What is an "effective" human reasoner? Do you mean, an *actual* human?

"This assumption [std. conditioning] is controversial, but only minimally so, as standard conditionalization is widely accepted by Bayesians."

In later chapters, I shall defend the thesis…

Neither sharp credences nor information-based update…

On the contrary, most Bayesians now appear to prefer indeterminate creedal states…

This section sets the stage by examining what a defence would need to do in order to be successful.

[Paragraph break at this point; "Partial belief epistemologists…" – move this sentence to the next paragraph.]

p. 45 I now turn to a more detailed…

…and the philosophical approaches, which…

p. 46 Can you add a final sentence to section 3.3.2 that reminds us that as shown in 3.2.4, there is no conflict?

…see the above comment on weakening…

p. 49 The [OMIT strength of] axioms, which require… are defensible and do not lead …

p. 50 …The disagreement is often about how far intuition can take us in this type of circumscription.

…beliefs of a rational agent that many may find…

p. 51 Not sure quite how the Salmon quotation is supposed to help. The Salmonesque worries seem to be about absurdity in what you have called relative prior probabilities, but your dissertation is about absurdity in the updating methods. Are you simply indicating that, like Salmon, you want to avoid absurdity?

Chapter 4

In very good shape!

p. 55 What are f_y, o_t?

p. 56 "the forecaster anticipates every instance of rain with a 100% forecast" – this seems stronger than what you said on p. 55, where it's enough for the forecaster to predict 30 days of rain out of 100.

…I shall argue that the failure of Jeffrey conditioning…

p. 57 …that the metric for a geometry of logic is Euclidean by default: [Add colon]

p. 58 Statement of the salient axioms: you need to tell us what they mean. What are I, w, etc? Perhaps the material on the next page that does this should come earlier.

p. 59 …only legitimate measures…

p. 63 "isomorphic manifold" ?

p. 66 GE_{xp}?

p. 77 "Levinstein… account falls far short…": good place to back this with a quotation

p. 84 Description of SKEW-ANTISYMMETRY could be made clearer

Chapter 5.

p. 99 Some repetition at the start of the chapter in first two paragraphs. For example, at the start of paragraph 2, it sounds as though this is your first reference to PME.

p. 100 "His derivation… is conceptually more complex than he assumes." You need an additional sentence or two that has more punch: "Specifically, his derivation relies upon an assumption that is rejected by advocates of PME: that rational agents can (indeed, should) have imprecise credences. In this chapter, I show that if agents have sharp credences, Wagner conditioning actually agrees with PME. In chapter 6, I examine the debate about imprecise credences."

p. 100 Why not state (I) and (L) up front as indented propositions? Start the paragraph with a sentence such as the following: "Wagner's argument hinges crucially on the following two propositions:

(I) ….

(L) …

Below, we show that PME and (W) are consistent given (L)…

p. 100 You need a sentence that explains your table (which should also probably be numbered – e.g., Fig 5.1). Explain: x means "inconsistent" and ✓ means "consistent".

p. 101 …for example [in] the Judy Benjamin case; see chapter 7. [Delete "in", add semi-colon]

… what we mean by objectivity. [Where does objectivity come in?]

…let us articulate (L) and PME. [What about (I)? Also, maybe you should clarify that you are re-stating PME from chapter 3, not articulating it for the first time.]

…some rules <u>about</u> which propositions…

p. 102 "In the remainder of this chapter I will provide a sketch of a formal proof for this claim. <u>(The case for (L) will be made in chapter 6.)</u>...

p. 103 Can you explain what $m(E)$ and $b(E)$ mean, as Dempster-Shafer provide a kind of intuitive analysis here? (I think Halpern does this as well.) Just a sentence or two.

p. 104 The long quotation from Wagner is supposed to illustrate your claim that Wagner's 'scathing verdict' depends upon PME and (I), but there is no hint of (I) in the quotation. I suspect it is the point about "all measures q bounded below by b", but PLEASE EXPLAIN after the quotation in what way Wagner is invoking (I).

p. 106 How do you get 60:40 for Wagner's PME solution?

…which <u>a</u> PME advocate… ["a" rather than "an"]

p. 106-7 "After the utterance… populated by prior probabilities according to (L). That this happens retrospectively may or may not be a problem…" And then further down the page, "Following (L), we shall populate the joint probability matrix… which is a perfect task for PME, as updating the joint probability… is a perfect task for *Infomin*."

I am confused. First, because PME now seems to be a principle that applies to priors (in direct contrast to *Infomin* which applies to updating) – and this goes against the way you re-defined PME in chapter 3 (see above). Second, because early in the thesis, you made it sound as if it is no part of your thesis to concern yourself with applications of PME to prior probabilities.

I suspect: you want PME to include BOTH updating via KL ("Infomin") and something you earlier called synchronic applications of PME that is to be distinguished from absolutely prior probabilities, and that it is the latter sort of application that you want here. You need to clarify, both in chapter 3 (perhaps giving the two cases clear names) and here (with reference back to chapter 3). Otherwise, it looks like you are suddenly bringing back the type of application of PME that you disavowed at the outset.

Finally: how DO you get the priors that you use to solve the Wagner problem?

p. 109 A fair bit of repetition on this page.

p. 111 Repetition towards the bottom of the page, I think.

p. 112 I think this is the key part of your argument, but I could not follow the argument starting on the last half of this page and running right to the end of the chapter. That may not be important. But anything you can do to help will be appreciated by the reader.

Chapter 6

Generally, this chapter is in good shape.

p. 118 …to address problems that sharp credences face in reflecting an agent's doxastic state

…there are now many Bayesians who permit…

Explain the terminology: Laplaceans, Booleans. Why these labels?

p. 121 "…without it Bayesian theory is a means with no ends…": this is too cryptic. Explain!

…has an explanation at hand for how partial beliefs…

p. 123 …or when C_1 arises from C_2 by conditioning [is this backwards?]

TBD footnote about one-third from the end needs to be filled in.

p. 125 "what I will call the double task": you need to explain what it is, or at least provide a forward reference to the relevant section

p. 127-8 Really clear; excellent explanation.

p. 128 To understand Kaplan's remarks, it would help a great deal to have a quick summary of the set-up (numbers of balls, etc.)

p. 135 "Indeterminacy imposes a double task": forward reference again to the relevant section.

p. 136 What is LP? What is PSET?

p. 138 …Joyce explicitly says that the REPETITION argument …

p. 138 "This worry supports Augustin's second concession…" How, if we are talking about Logan?

p. 139 I like the riposte a great deal. It's just what you should anticipate.

p. 142 Bottom of page: opaque. What are X, Y? I am guessing X = bet 1 and Y = bet 2, but spell this out.

p. 145 Hand urn argument is excellent.

Chapter 7

Another chapter that is basically in good shape.

Again, some redundancy at the start, as in "call this standard conditioning", as if we had not discussed this previously!

p. 150 …as likely to land in A_2 as to land in A_1…

p. 151 Diaconis and Zabell state: "any claims…"

p. 154 Check that your discussion of full employment is not repeating earlier discussions. Actually, this may be the first time you really explain it. If so, does it belong here?

p. 157 "conflicts with the symmetry requirements" – explain them!

Pp 162-3 Really nice discussion

p. 164 what is the "shaded line" ?

p. 165 What is the CAR condition? State it.

p. 167 Excellent summary to end section 7.4

p. 167 "no corresponding reason why we should update our probabilities using the powerset approach", i.e., we have no information that this approach corresponds to the scenario in which Judy gets her information. Is that what you mean?

p. 168 … formula for the powerset approach corresponding to the formula for PME approach… [What does this mean?]