

★

How uncertain do we need to be?

★

Jon Williamson

Draft of August 12, 2013

Abstract

Expert probability forecasts can be useful for decision making (§1). But levels of uncertainty escalate: however the forecaster expresses the uncertainty that attaches to a forecast, there are good reasons for her to express a further level of uncertainty, in the shape of either imprecision or higher order uncertainty (§2). Bayesian epistemology provides the means to halt this escalator, by tying expressions of uncertainty to the propositions expressible in an agent's language (§3). But Bayesian epistemology comes in three main varieties. Strictly subjective Bayesianism and empirically-based subjective Bayesianism have difficulty in justifying the use of a forecaster's probabilities for decision making (§4). On the other hand, objective Bayesianism can justify the use of these probabilities, at least when the probabilities are consistent with the agent's evidence (§5). Hence objective Bayesianism offers the most promise overall for explaining how testimony of uncertainty can be useful for decision making.

Interestingly, the objective Bayesian analysis provided in §5 can also be used to justify a version of the Principle of Reflection (§6).

Contents

- §1. To be uncertain or not to be?
- §2. The uncertainty escalator
- §3. Bayesian epistemology
- §4. Objectivity and Calibration
- §5. Matching objective Bayesian probabilities
- §6. Conclusions

§1

To be uncertain or not to be?

In November 2011 the UK Meteorological Office (Met Office) changed the way it gave weather forecasts, from mainly qualitative expressions of uncertainty—'rain at

times', 'scattered showers, mainly in the NW', 'up to 50mm in places', 'risk of heavy bursts'—to quantitative announcements of probabilities—'a 10% chance of...', '80% probability of disruption to...', 'a one-in-five chance that Heathrow Airport might get...', 'a 5% risk of... so it is probably worth the risk.'¹

The question arises as to whether that was the right thing to do. A previous attempt to introduce probability into Met Office forecasts did not augur well. In April 2009 the Met Office announced that a barbecue summer was 'odds on'.² In the face of widespread criticism in the UK press during a summer of floods in which rainfall in June, July and August turned out to be 40% above the long-term average, the Met Office noted: 'in April we said there was a 65% chance of temperatures above average and rainfall below average but that does leave a 35% chance that the opposite would be true'.³ That the British public were not ready for probabilistic forecasting was confirmed when, on 9th December 2011, the Met Office won a Golden Bull Award (awards which are meted out to 'the worst examples of written tripe') at the annual awards ceremony of the Plain English Campaign, for their use of 'probabilities of precipitation'.⁴

Even if the public mood was not in favour of probabilistic forecasts, the question remains as to whether it is nevertheless in the public interest to express uncertainty this way.

This question was discussed in a Royal Society symposium on *Handling uncertainty in science*, and the editors of the resulting proceedings gave the following defence:

if the public were more exposed to weather prediction as inherently probabilistic, perhaps there would be more acceptance of the simple fact that one's view about the risk of dangerous climate change should not be framed in the black and white terms of 'belief' and 'scepticism'. For one thing, belief should have no real role in science, and, in truth, all good scientists are inherently sceptical people. (Palmer and Hardaker, 2011)

This defence is unsatisfactory in two respects. First, there isn't much uncertainty that the climate is currently on a dangerous trajectory. Rather, uncertainty attaches to specific long-range climate predictions, and to the question of which policy interventions can best mitigate climate change. Neither of these latter forms of uncertainty have yet yielded to probabilistic forecasts about which one can be very confident. Hence there is a danger that probability is just being invoked to give a false sense of objectivity. The second questionable feature of this defence is the claim that belief should have no real role in science. While *faith*—understood as a propositional attitude that is impervious to changes in one's body of evidence—arguably should have no real role in science, *belief* in scientific claims makes scientific progress possible: scientists would not investigate theories unless they believed that such theories are likely to be true, or that such investigations are likely to be fruitful. As Ramsey (1926,

¹<http://research.metoffice.gov.uk/research/nwp/ensemble/uncertainty.html>.

²<http://www.metoffice.gov.uk/news/releases/archive/2009/summer2009>.

³<http://www.dailymail.co.uk/news/article-1202982/Met-Office-left-red-faced-Britains-forecast-barbecue-summer-turns-washout.html>. Toulmin would argue that the Met Office is using probability assertions precisely to guard against such comebacks: 'If they unreservedly forecast cloud later today and the skies remain clear, they can justifiably be rounded on by the housewife who has put off her heavy wash on account of their prediction.' (Toulmin, 1958, p. 47.)

⁴<http://www.plainenglish.co.uk/awards/golden-bull-awards/golden-bull-winners-2011.html>.

p. 183) pointed out, ‘whenever we go to the station we are betting that a train will really run, and if we had not a sufficient degree of belief in this we should decline the bet and stay at home’. It is important—in the context of climate change and elsewhere—to deliberate about how much belief to apportion a scientific hypothesis, and not to sweep the very idea of belief under the carpet.

This defence, then, merely serves to highlight the dangers of probabilistic forecasts: at worst, they can be used a rhetorical device to convey a false sense of objectivity, replacing apparently more subjective terms such as ‘belief’. So the question remains as to whether it is appropriate to announce probabilistic forecasts.

In contrast to the line taken by the Royal Society symposium, the Met Office responded to public unease as follows:

Often people want to make a decision, such as whether to put out their washing to dry, and would like us to give a simple yes or no. However, this is often a simplification of the complexities of the forecast and may not be accurate. By giving PoP [Probability of Precipitation] we give a more honest opinion of the risk and allow you to make a decision depending on how much it matters to you. For example, if you are just hanging out your sheets that you need next week you might take the risk at 40% probability of precipitation, whereas if you are drying your best shirt that you need for an important dinner this evening then you might not hang it out at more than 10% probability. PoP allows you to make the decisions that matter to you. (Press Release, 9th December 2011.)⁵

Thus the Met Office defence is that *decision making* requires uncertainty reports. This seems essentially right. Arguably, Bayesian decision theory currently provides the best general guide to decision making. Bayesian decision theory requires making the range of possible acts explicit, as well as their utilities and the probabilities of the different outcomes pertinent to the decision problem. Acts and utilities vary according to the interests of the decision maker and there is not much advice the Met Office can offer in that regard. Therefore, if Met Office reports are to be relevant to decision making, they need to report probabilities.

Thus a *prima facie* case *can* be made for experts publicly quoting their probabilities to express the uncertainty that attaches to their forecasts. However, as we shall now see, probabilistic forecasts are the thin end of the wedge.

§2

The uncertainty escalator

Having established a need for explicit uncertainty reports, the next question that arises is: *how uncertain do we need to be?*

The Met Office could simply give single probability reports such as, ‘a 40% probability of rain’. Such an expression of *first order probability* is required to get Bayesian decision theory off the ground. It might mean one of two different things. The standard Bayesian take is that this figure is the Met Office’s degree of belief in rain; the infamous Dutch book argument, which develops Ramsey’s connection between beliefs and bets mentioned above, is used to conclude that such degrees of belief ought to obey the laws of probability. But this 40% figure can alternatively be taken as the Met Office’s best estimate of the physical probability of rain; the argument

⁵<http://www.metoffice.gov.uk/news/in-depth/science-behind-probability-of-precipitation>.

being that decision makers are interested in the objective prospects of rain, not in the subjective opinions of those who work at the Met Office.

In fact, the Met Office provides some of its customers with a confidence range, rather than a point-valued probability. Thus one might report ‘a 34–46% probability of rain’. This expression of *imprecise first order probability* admits of two interpretations. An epistemological interpretation takes this to be an expression of partial belief; this position can be motivated by a Dutch book argument where the bettor is allowed to buy and sell bets at different rates. Or this expression can be viewed as providing bounds on the physical probability of rain, the thought being that a point estimate is almost always wrong, so it makes more sense to provide an interval estimate, which might be right on say 95% of occasions. Either way, the idea is that typically there is some uncertainty attached to a precise (point-valued) first order probability report, and providing an imprecise first order probability instead can allow one to quantify this uncertainty. This is important in a decision-making context if the decision to be taken depends on where in the interval a precise probability lies. One may decide to hang out one’s laundry at 40% probability of rain but not at 45%, and knowing that the decision taken is not robust under changes of the probability within the confidence range 34–46% may actually help one take the decision: it may lead one to be more cautious in this case, for instance, and not hang the washing out.⁶

Of course, not all values within the confidence range are equal. Normally, some will be more plausible than others. Thus a 40% probability of rain might be a much more plausible estimate than a 45% probability of rain, and if the difference between one’s confidence in the two values is substantial enough, one might decide to hang out one’s washing after all. The way to provide information for this nuance in the decision-making process is to report a probability distribution over the first order probability of rain. This *second order probability* would normally be construed as epistemological: as a belief about one’s degree of belief in rain, or as a belief about the physical probability of rain. A second reason for moving from imprecise first order probability to second order probability is that the end-points of an imprecise first order probability interval are often somewhat arbitrary; a second-order probability distribution can give positive weight to all first order probabilities and so is not subject to arbitrary cut-offs (Good, 2003).

There are also reasons to move directly from first order probability to second order probability. In fact, de Finetti’s representation theorem shows that under certain circumstances one’s first order degrees of belief can be thought of as having been produced by a second order probability distribution over first order physical probabilities: if X_i is a sequence of binary random variables that take possible values 0 or 1 and that are exchangeable with respect to one’s first order belief function P ,

⁶Some Bayesian decision theorists argue that this uncertainty about the probability of rain should not influence decision making: all decisions should be made on the basis of precise first order probabilities. But this is a controversial line to take. In practice our decisions *are* influenced by considerations to do with their robustness under plausible alternative first order probabilities, and this is apparently reasonable. If one’s threshold for not hanging out one’s laundry is a 45% probability of rain and one is very confident that the probability of rain is less than 45% then one can be confident that one’s decision is the right one in the circumstances. But if one knows that the first order probability is to some extent an arbitrary representation of first order uncertainty, or a poor estimate of physical probability, and that the value could well be higher than the threshold for the decision, then one cannot be confident that the decision is the right one. Thus there is at least a *prima facie* case for taking imprecision into account. The Bayesian position will be revisited in §3.

then

$$P(X_1 = x_1, \dots, X_n = x_n) = \int_0^1 z^{r_n} (1-z)^{n-r_n} dF(z)$$

where $r_n = \sum_{i=1}^n x_i$, where z may be thought of as ranging over possible values of first order physical probability $P^*(X_i = 1)$ for X_i IID with respect to physical probability P^* , and where F may be thought of as a probability distribution over these physical probabilities. Moreover, this second order probability distribution is determined by the first order belief distribution: $F(z) = P(\bar{X}_\infty \leq z)$ where $\bar{X}_\infty = \lim_{n \rightarrow \infty} \sum_{i=1}^n X_i/n$ is the limiting relative frequency of X_i (see, e.g., Schervish, 1995, §1.4; Williamson, 2011, §7.4.2). Hence there is a sense in which a first order belief distribution is equivalent to a second order probability distribution over first order physical probability, in which case moving from first order probability to second order probability is a trivial step.⁷

Note that second order probabilities can be rather speculative. Confidence interval estimation methods might make the Met Office 95% confident that the probability of rain is between 34% and 46%, and 60% confident that there is a 39–41% probability of rain, leading it to announce a second order probability distribution that fits these two constraints. But they might just as well have announced a different second order distribution that fits these constraints; there is often substantial leeway in determining higher order probabilities. So it is more sensible to announce, instead of a single second order distribution, the whole set of second order distributions that fit the available constraints. Moreover, announcing this *imprecise second order probability* can be beneficial for decision making. If some distributions in the set of second order probability functions make it quite probable that the chance of rain is at least 45% then one might justifiably be cautious and refrain from putting one's washing out to dry after all. Thus, announcing an imprecise second order probability can allow one to perform a useful robustness analysis that determines the sensitivity of the decision taken to the precise second order probability measure chosen. Such a robustness analysis can, and often does, influence decision making.

Of course, not all second order probability distributions within this set of distributions are equal. It tends to be the case that some probability functions from within a set of functions will be more plausible or natural than others: a normal distribution, for example, often stands out as being particularly probable in the context of certain convergence results, and so-called concentration results show that there is a high probability that a physical probability distribution is close to the maximum entropy distribution (see, e.g., Jaynes, 1982, §3).

It should be clear by now where this is leading. The Met Office should advertise a *third order probability* function, since that might show that the caution resulting from the imprecise second order probability was too hasty. But it is hardly plausible to suggest that a single such third order distribution stands out to the exclusion of all others—if we consider *imprecise third order probability*, other plausible distributions can influence our decision. But how plausible? Plausibility is a matter of degree and some of these third order probability functions will be more plausible than others. What we need is *fourth order probability* ...

⁷Thus formulated, de Finetti's theorem invokes an exchangeability condition—degrees of belief about a sequence of events are independent of the order of those events. This condition would not normally be satisfied in the context of beliefs about rain on consecutive days, since the order matters here (rain on two consecutive days followed by a dry day may be more likely than rain followed by no rain followed by rain). It is however possible to weaken the exchangeability requirement to yield an analogous result that is more applicable to weather forecasting (Diaconis and Freedman, 1980).

★

We end up with an escalation in levels of uncertainty. If we admit the need for reports of uncertainty for decision making, then there are compelling reasons for progressing to the next level—moving to the next level might, after all, warrant choosing a different act in the decision problem that motivated the need for uncertainty in the first place.

This uncertainty escalator is a problem. As we have seen, there may be no stable answer to the question of whether we should put our washing out to dry, if at each level of uncertainty some higher level leads to a different decision. Moreover, whether or not there is a stable answer, decision making gets harder as we ascend the escalator. First, it becomes pragmatically harder, as there are more kinds of uncertainty that need to be taken into account in order to make a decision. Second, every other level up the escalator we reach imprecision, and this poses its own difficulties. In standard Bayesian decision theory, a precise probability function will normally trigger one or other of the acts available, as there will normally be only one act that maximises expected utility. But when we admit imprecision, i.e., when we consider a *range* of probability functions, it will often be the case that some of these functions trigger one act while others trigger other acts. There are then two options. One response is to say that we do not have the tools to make a decision in such cases. Then we may end up in a situation in which we alternate between being able to make a decision and not being able to make a decision as we progress up the escalator. The second option is to appeal to some further rule in order to allow one to make a decision in such cases. For example, one might choose the act that maximises worst-case expected utility, where one considers the full range of probability functions to determine this worst case. The problem is that a great many such rules have been put forward in the literature, and no doubt many more will be put forward, all of them plausible in one way or another (see, e.g., Troffaes, 2007). This introduces a new kind of uncertainty—uncertainty as to which rule to apply to make a decision. Should one consider robustness of the decision made under a range of such rules? Of course, not all the rules will be equal with regard to the desiderata in play or the particular context of the decision problem, so perhaps one should favour some rules over others. What this amounts to is weighing more highly those rules that fare better with respect to the chosen desiderata so as to provide to a linear ranking of the available acts, in order that one might embark upon the act that is most highly ranked. Formally, such a weighting would correspond to a function that maps a decision problem to a linear ordering of acts. Now, it would be astonishing to think that the constraints would uniquely determine a single such weighting, so one had better consider a range of weightings. But not all weightings are equal ... In short, imprecision yields its own set of difficulties for decision making, and arguably even generates an escalator analogous to the uncertainty escalator introduced above.

The moral of this story is that we need some way to halt the uncertainty escalator, to stop us from indefinitely ascending levels of uncertainty. Any theory of uncertainty, if it is to be relevant to decision making, had better take a principled stance as to which level is the right place to stop. Otherwise we will never get to the dinner tonight for want of being able to decide whether to dry our shirts indoors or outdoors.

Those who think that one should avoid any use of epistemological probabilities and only consider physical probabilities will want to say that we should stop at the first level: feed in only (first order) physical probabilities to the decision problem.

The difficulty with that view is that it requires that all the relevant outcomes be the sorts of things to which physical probabilities attach, and that the decision maker has access to all these physical probabilities. In most cases in which there is limited evidence of physical probabilities it will be hard to make a decision at all. This makes the epistemological view, which we shall examine next, apparently indispensable for decision making.

§3

Bayesian epistemology

One response to the uncertainty escalator is to take the epistemological view of Bayesianism seriously. According to the epistemological view, probabilities measure the strengths of an agent's beliefs. (More precisely, the core tenet of Bayesian epistemology is that if the strengths of the agent's beliefs are apportioned in an *appropriate* way, they are probabilities.) The object of a belief is a proposition. Hence Bayesian degrees of belief are defined over the propositions that the agent can express. It is the agent's language that determines which propositions she can express, so Bayesian probability can be thought of as defined over those propositions expressible via sentences of the agent's language. Setting aside those grammatical sentences that fail to express propositions—sentences like 'Get up!', 'Colourless green ideas sleep furiously.', and perhaps 'This sentence is false.'⁸—, we can think of Bayesian probability as defined over those sentences $\theta_1, \theta_2, \dots$ of the agent's language that succeed in expressing propositions when used by that agent.

This fixes the domain of the probability function, and gives a clear answer to precisely where on the uncertainty escalator the Bayesian comes to a stop. If the agent's language cannot express propositions about first order probabilities, then Bayesian probability is first order. Thus in practice many Bayesian decision-support systems invoke only first order uncertainty, because it is often the case that an artificial agent can only express propositions about, e.g., possible financial states and transactions or possible medical symptoms and prognoses, and not propositions about the probabilities of those states or transactions. On the other hand, if the agent's language can express propositions about at most n -th order probability (precise or imprecise), then Bayesian probability is $(n + 1)$ -st order. Human agents, for instance, can typically express high order uncertainty.

The advantage of this epistemological stance is that it gives a principled response to the problem outlined in §2, and thus renders decision making possible. But there are also several apparent disadvantages. First, it seems set to render decision making rather intractable for agents like us who can express very high order uncertainty. Second, it makes the decision taken depend on the agent's language, and it is by no means obvious that whether one should hang out the washing should depend on one's linguistic capabilities. Third, the epistemological approach to Bayesianism does not fit well with the uses of Bayesianism in statistics—where Bayesian probability is normally second order, where the domain of the probability function is tailored very much to the individual application, and where Bayesian probability is increasingly treated as a pragmatic tool rather than as a particular interpretation of uncertainty (Kass, 2011; Gelman and Shalizi, 2013). Fourth, the dominant epistemological view of Bayesianism takes Bayesian probabilities to quantify *subjective* opinion, and takes satisfying the axioms of probability to be the only synchronic normative constraint on

⁸See, e.g., Goldstein and Blum (2008) on this last kind of sentence.

degrees of belief. This strict subjectivist position would deem rational the forecaster who announced a probability of 99% for a barbecue summer in the total absence of any evidence in favour of a sunny summer—a view that would act as a red rag to a bull to arbiters of public standards like the UK *Daily Mail*, which was quick to condemn Met Office probabilities that *were* based on evidence. Fifth, the strict subjective view of Bayesianism is in danger of making the very act of announcing forecast probabilities incomprehensible: why should Met Office probabilities be of relevance to my decision problems, if I am supposed to be making my decisions on the basis of my own subjective probabilities and there is no normative imperative to let my own degrees of belief be guided by someone else's?

The first three worries can perhaps be alleviated by taking the agent's language to vary with her operating context. An agent who is an obstetrician by day and a carpenter by night might be taken to move between two languages, one of which contains terms such as 'cardiotocograph' and the other of which contains terms such as 'gimlet'. Similarly, in the context of reading this paper one will be operating with a language in which high order uncertainties are expressible, but in the context of hanging out the washing, one operates in a context in which domestic and meteorological possibilities are entertained, but not—normally—higher order uncertainties. It is of course hard to say exactly what the limits are of a language which is contextual in this sense. But this move does bring Bayesian epistemology a step closer to the practice of Bayesian statistics, where considerable effort is required to determine a language in which to express a particular problem. Regarding dependence on language, the statistician can admit that her inferences do depend somewhat on the way in which the problem is formulated, but point out that some ways of formulating a problem are better than others—ways in which, for example, one can describe the phenomena of interest and the other variables that are explanatorily or predictively relevant and exclude as much as possible that is irrelevant to the problem solution. If higher order uncertainties are relevant to a decision, one may then decide that they should be expressible in the problem's 'language'.

One might worry that this way of looking at things makes choice of language a decision problem that needs to be solved before we can start thinking about the original decision problem of whether to hang out our washing, a move which could lead to regress. But one can, on the contrary, put language first: start off with one's default language—the language in which one would normally express the propositions required to state and solve the problem—and test the sensitivity of one's results to small changes in that language. If conclusions and decisions are not robust under such changes then the sensitivity analysis provides grounds to move to a new language—perhaps a language in which one can express a higher level of uncertainty.

We shall now focus on the remaining two problems outlined above, which are concerned with subjectivity and the motivation behind paying attention to expert forecasters' probabilities.

§4

Objectivity and Calibration

Why should we use Met Office probabilities for decision making? This only makes sense if its probabilities are *better* than our own. There are two main ways in which one might try to say how another agent's probabilities could be better than our own. One might think that they could be better because we have made some *mistake* in

the way we've apportioned our probabilities—perhaps we have given θ probability 0.6 instead of 1, not realising that θ is implied by our evidence. But then it would not make sense to use Met Office probabilities for decision making: once we realise we have made a mistake, we ought to correct the mistake, not borrow the probabilities of someone else, since the other agent's probabilities are likely to be grounded on different evidence and not qualify as rational given our own evidence. The second and more promising avenue is to say that Met Office probabilities could be better than our own because they know things that we don't about the weather: they have better evidence and this better evidence makes their probabilities better calibrated to the world.

Bayesian epistemology tends to cash out the idea of calibration in two very different ways. One kind of calibration is indirect, long-run calibration: a strict subjectivist Bayesian such as de Finetti would argue that, if prior probabilities are exchangeable, updating degrees of belief by Bayesian conditionalisation leads to asymptotic calibration with empirical frequencies (see, e.g., de Finetti, 1937). Another kind of calibration is direct, short-run calibration: some Bayesians impose a further norm on degrees of belief—often called the Straight Rule (Miller, 1966, p. 60), Miller's Principle (Miller, 1966) or the Principal Principle (Lewis, 1980)—which says that if an agent knows just the physical probability of a proposition, then she should directly set her degree of belief in that proposition to the physical probability.

Indirect calibration is not applicable to our problem, for two reasons. First, Met Office probabilities are not themselves calibrated in this way—they are not produced by long-run conditionalisation on exchangeable priors. Rather, this sort of forecast is produced by considering the predictions of a collection of plausible models, produced by simulations run over a variety of plausible initial conditions (see, e.g., Slingso and Palmer, 2011). To the extent that a high proportion of these simulations predict rain, a forecast of rain will be given a high probability.⁹

The second reason why indirect calibration is not suited to the context of this paper is that it offers slender grounds for the decision maker to use Met Office probabilities for her decision problem. If she follows the norms of strict subjective Bayesianism, then her probabilities—modulo exchangeability—will automatically calibrate in the long run; she should update her probabilities by conditionalising on her new evidence, *not* by directly adopting Met Office probabilities. Under strict subjective Bayesianism, the only way decision maker D can match forecaster F 's probabilities is if she has made such a commitment in her prior probabilities: $P_D(\theta \mid \varphi \wedge P_F(\theta) = x) = x$, for all relevant propositions θ and all suitable evidence bases φ . But under strict subjective Bayesianism there is no normative imperative to adopt these prior conditional probabilities, given that the agent is free to choose whichever probabilities she fancies for her prior. So this particular coincidence of probabilities should be thought of as a measure-zero possibility. Worse, if F is a forecaster like the Met Office that is known not to abide by the norms of strict subjective Bayesianism, agent D might be (subjectively) inclined to avoid F 's probabilities rather than match them.

While indirect calibration is not applicable to our problem, direct calibration

⁹Moreover, as noted in footnote 7, exchangeability is rarely plausible in the context of this sort of forecasting: the probability of rain on any particular day is higher if there is rain the day before. However, long run calibration is possible under weaker conditions than exchangeability (see, e.g., Gaifman and Snir, 1982, §2). Note that where long-run calibration cannot be guaranteed through conditions such as exchangeability, it can be forced by adopting further norms on Bayesian probability; this is sometimes the strategy of statistical approaches to forecasting (Skouras and Dawid, 1999).

appears to be more promising. First, expert forecasters validate their models by ensuring that they are directly calibrated to physical probabilities: if it is known to rain on 30% of days and a model forecasts rain on 70% of days, then the model will be thrown out or its parameters will be changed to ensure better calibration. Indeed, to the extent that Met Office probabilities are calibrated, they are directly calibrated. Second, direct calibration appears to offer some scope for setting one's degrees of belief to those of the forecaster. If one should directly calibrate to physical probabilities, then it seems plausible that one should directly calibrate to a forecaster's probabilities that are themselves directly calibrated to physical probabilities.

So far we have seen that there are two kinds of Bayesianism. One kind—*strict subjectivism*—maintains that the only norms on degrees of belief are that they should be probabilities and they should be updated by Bayesian conditionalisation. This kind of Bayesianism pins its hopes on indirect calibration, but it does not seem to have the wherewithal to explain why a decision maker should need to make use of expert probabilities. A second kind of Bayesianism admits a further norm, direct calibration, which does seem to offer more scope in this regard.

It should be noted, though, that this second kind of Bayesianism itself admits two versions. One, which is sometimes called *empirically-based subjective Bayesianism*, says that direct calibration is the only further norm on strength of belief. This kind of Bayesianism maintains that, to the extent that the agent knows empirical probabilities, her degrees of belief should be directly set to those probabilities, but otherwise she is rational to adopt any probabilities that she pleases. Another sort of Bayesianism—*objective Bayesianism*—adopts a further norm on strength of belief, which deems an agent's belief function to be irrational if it attaches strong belief or disbelief to a proposition in the absence of evidence that forces such strong commitment. Since entropy is a natural measure of the lack of commitment of a probability function, this norm is sometimes explicated using the Maximum Entropy Principle: an agent's belief function should be one, from all those that are directly calibrated with evidence of chances, that has maximum entropy (Jaynes, 1957; Williamson, 2010).

It turns out that empirically-based subjective Bayesianism has mixed success in accounting for the use of expert probabilities in decision making. If decision maker D learns that $P_F(\theta) = x$ and that x is a good estimate of the corresponding physical probability, then x might become D 's best estimate of the physical probability and it can be rational to set $P_D(\theta)$ to x . However, this requires learning not just the expert's probability but also the fact that that probability is closely calibrated to empirical probability. This is very demanding. More typically, D will learn just that $P_F(\theta) = x$. If D assumes that F follows the norms of empirically-based subjectivism, she cannot infer that x is a close estimate of physical probability, for that would suppose that F knows the physical probability, which may not be true. The most D can infer is that x has been arbitrarily chosen by the expert from some *set* of probabilities compatible with the data that is available to F . Perhaps F only knows another physical probability $P^*(\theta \wedge \varphi) = w$, where $w \leq x \leq 1$, in which case $P_F(\theta)$ is only constrained to lie in the interval $[w, 1]$. Perhaps F only has a confidence interval $[w, y]$ for $P^*(\theta)$, where $w \leq x \leq y$; again, this would lead to a constraint $P_F(\theta) \in [w, y]$. Then learning just that $P_F(\theta) = x$ may tell D next to nothing about $P^*(\theta)$, because D does not know how wide the interval is from which $P_F(\theta)$ has been chosen. To be sure, D does know that F is an expert, and can assume that F has *more* evidence than she has, in which case F 's intervals are perhaps narrower than her own. But that alone does not provide grounds for D to switch her probabilities

to those of F , because she may yet have some important evidence that F does not have. The Met Office, for example, only has evidence about the current weather obtained from certain weather stations at certain times of the day, and often makes false claims about the *current* weather, let alone the future weather. So, if you can see a shower about to blow over your garden then that evidence will be more pertinent in the short term than any Met Office forecast. Overall, then, while empirically-based subjectivism appears more promising than strict subjectivism in accounting for the use of expert probabilities for decision making, the grounds that it offers for matching expert probabilities are typically rather paltry.

In sum, one would want to say that there are some circumstances under which one ought to pay attention to an expert forecaster’s probabilities (e.g., when that expert is likely to have more substantial evidence than one does oneself). Unfortunately, strict subjectivism doesn’t seem able to say that. For the strict subjectivist, anything goes: while the principle of total evidence says that expert forecasts should be added to one’s stock of evidence, as far as strict subjectivism is concerned there is no normative imperative to have them influence one’s degrees of belief. Even empirically-based subjectivism, which does admit direct calibration to physical probabilities, struggles to make this claim: a forecaster’s reported probability does not on its own provide enough information about the physical probability to warrant direct calibration.

Interestingly, however, in certain well-defined circumstances objective Bayesianism does offer more scope in this regard, as we shall now see.

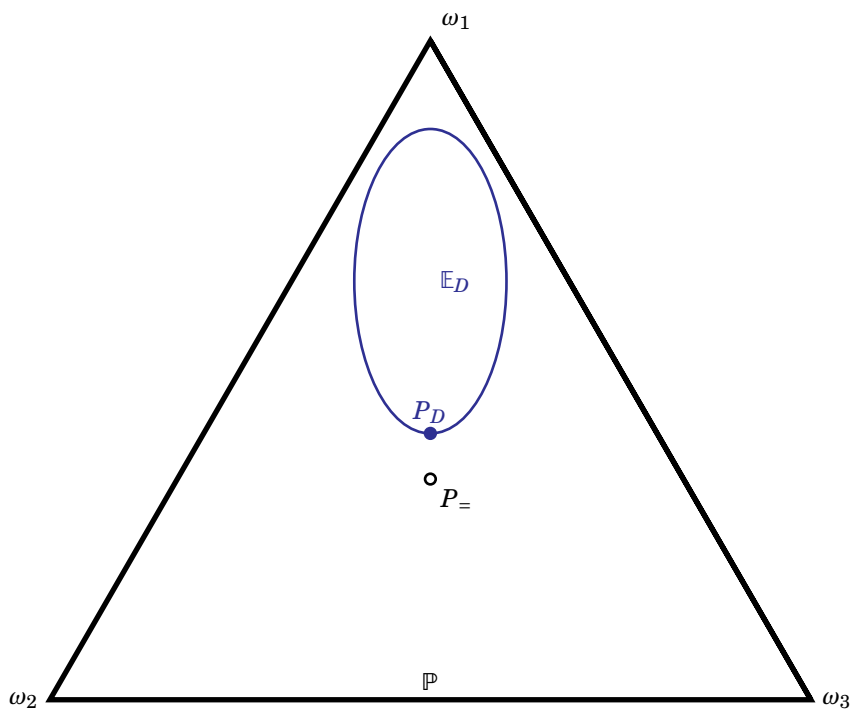
§5

Matching objective Bayesian probabilities

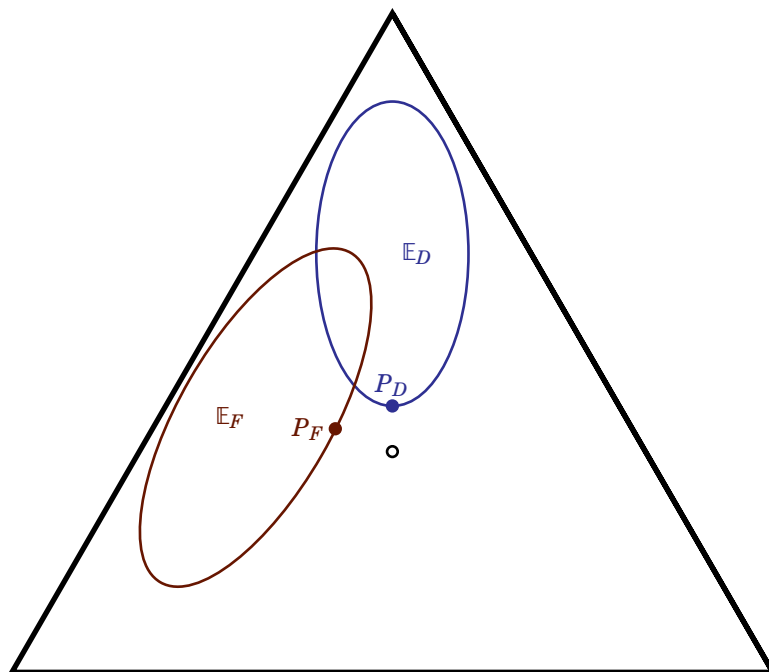
In this section we focus on the specific question as to when, if two agents follow the norms of objective Bayesianism, it can be rational for one to adopt the probabilities of the other.

¶ *An objective Bayesian analysis.* We saw that for the objective Bayesian there are three norms on rational belief. These norms can be explicated as follows (Williamson, 2010). The Probability norm says that a belief function should be a probability function: D ’s belief function $P_D \in \mathbb{P}$, the set of all probability functions defined on the sentences of the agent’s language of the moment. The Calibration norm says that a belief function should be directly calibrated with evidence of physical probabilities: $P_D \in \mathbb{E}_D$, the convex hull of the set of probability functions that satisfy constraints imposed by evidence.¹⁰ The Equivocation norm says that a belief function should otherwise be equivocal: P_D should be as close as possible to the *equivocator* probability function $P_{=}$ which gives equal probability to each possible state of the world that the agent can express, where closeness of probability function to the equivocator is measured using Kullback-Leibler (KL) divergence. (A probability function in \mathbb{E}_D that is maximally equivocal in this sense is the one that has maximum entropy.) In the case in which the agent can only represent three possible states $\omega_1, \omega_2, \omega_3$ of the world (e.g., signifying respectively rain, shine and overcast weather at a specific location and a specific point in time), this procedure is depicted as follows:

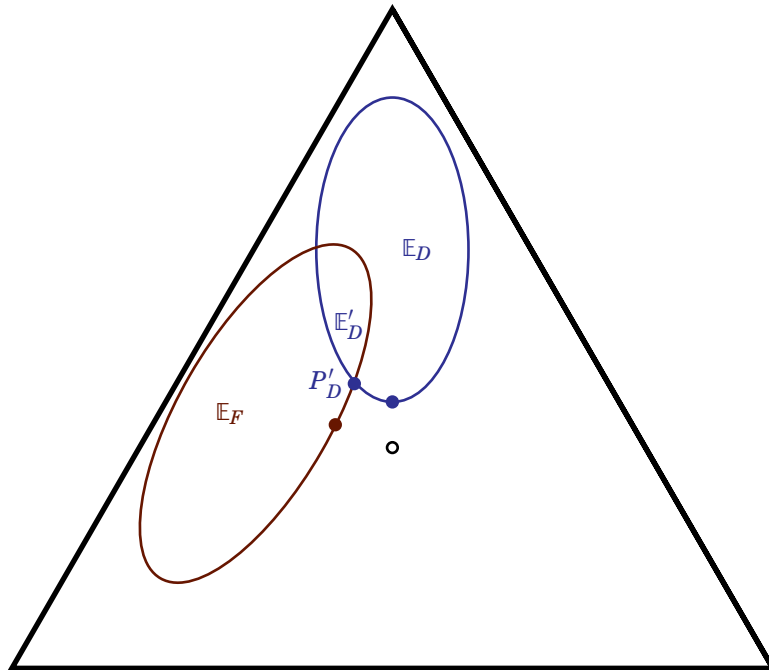
¹⁰Throughout this section, we shall assume for ease of exposition that constraint sets such as \mathbb{E}_D are closed as well as convex. This is the case, for example, with equality constraints or non-strict inequality constraints.



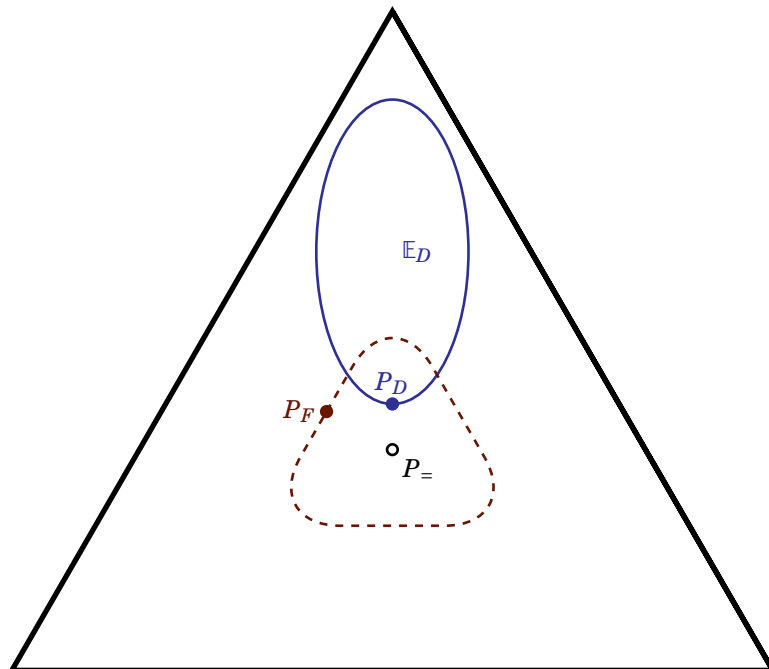
With the forecaster F following the same procedure we have the following picture:



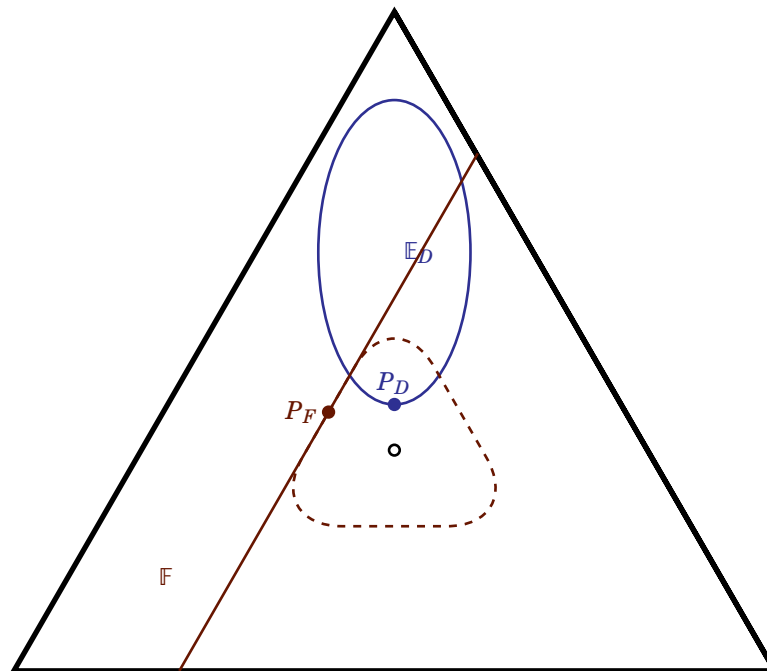
Were decision maker D to learn the forecaster's evidence, she would be able to derive \mathbb{E}_F and then revise her belief function to point $P'_D \in \mathbb{E}_D \cap \mathbb{E}_F$ that is closest to the equivocator function $P_=$:



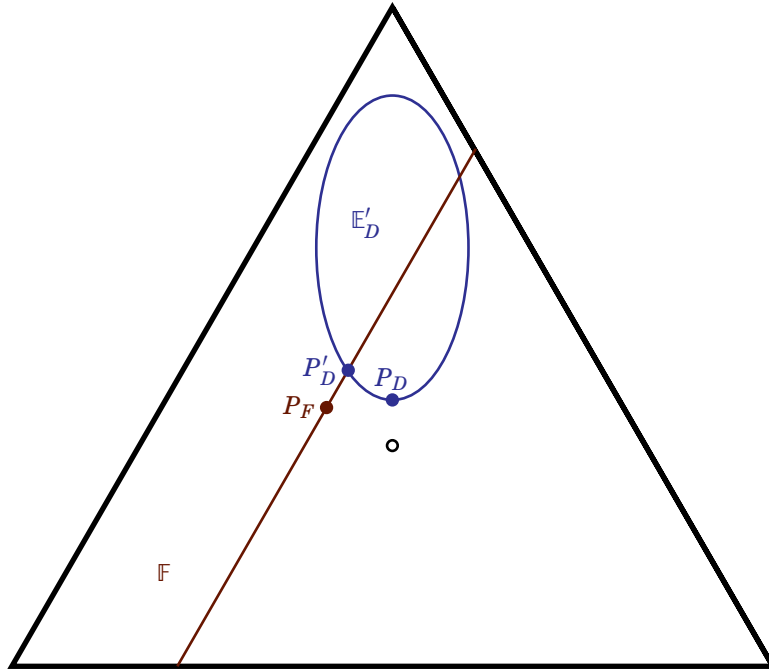
But in the case of probabilistic forecasting considered in this paper, the decision maker learns only the forecaster's probability function P_F defined over the partition $\{\omega_1, \omega_2, \omega_3\}$ of interest, or some values of that function. Suppose that the agent learns the forecaster's probability function P_F . In this case D can infer that there is some convex set \mathbb{E}_F whose most equivocal member is P_F . Since P_F is the closest member of \mathbb{E}_F to the equivocator $P_=$, the set \mathbb{E}_F must be in the region of \mathbb{P} that is at least as far from the equivocator as P_F . I.e., \mathbb{E}_F cannot lie inside the dashed curve below:



In fact, because \mathbb{E}_F is convex, it must lie beyond the tangent at P_F to that curve, i.e., in region \mathbb{F} which may be defined as the largest convex set of probability functions that contains P_F but contains no probability function closer to the equivocator than P_F :



Learning P_F , therefore, is tantamount to learning that there exists evidence that constrains degree of belief to lie in region \mathbb{F} . D 's new evidence is thus such as to constrain her belief function to lie in $\mathbb{E}'_D = \mathbb{E}_D \cap \mathbb{F}$, as long as this set is non-empty. Her new belief function P'_D is then the maximally equivocal member of \mathbb{E}'_D :



Thus when $\mathbb{E}_D \cap \mathbb{F}$ is non-empty, $\mathbb{E}'_D = \mathbb{E}_D \cap \mathbb{F}$. On the other hand, if the new evidence is inconsistent with the old, $\mathbb{E}_D \cap \mathbb{F} = \emptyset$, then some consistency maintenance procedure must be invoked. One simple procedure is to take $\mathbb{E}'_D = \langle \mathbb{E}_D \cup \mathbb{F} \rangle$, the convex hull of the functions that fit either the new evidence or the old (Williamson, 2010, §3.3.1). We shall adopt that procedure here.

In the diagram above, $P'_D \neq P_F$ and the decision maker does not simply adopt the forecaster's probabilities. The question arises as whether there are circumstances in which the decision maker ought to switch to the forecaster's belief function. The answer is that the switch is justified precisely when the forecaster's belief function is consistent with the decision maker's prior evidence:

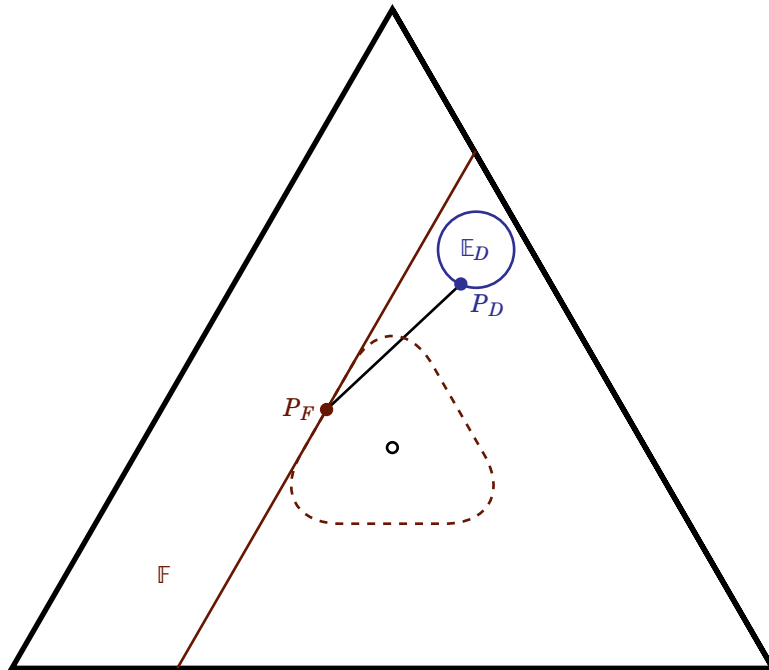
Theorem 1 $P'_D = P_F$ iff $P_F \in \mathbb{E}_D$.

Proof: First we shall see that $P_F \in \mathbb{E}_D$ implies $P'_D = P_F$. Now $P_F \in \mathbb{E}_D$ implies that \mathbb{F} is consistent with \mathbb{E}_D , so $\mathbb{E}'_D = \mathbb{E}_D \cap \mathbb{F}$. Note that KL-divergence from the equivocator has a unique minimiser in a closed convex set of probability functions, and \mathbb{E}_D, \mathbb{F} and \mathbb{E}'_D are all closed and convex. P_F is the unique function in \mathbb{F} that is closest to the equivocator P_* , so there is no other function in $\mathbb{F} \cap \mathbb{E}_D$ that is as close to the equivocator as P_F . Hence $P'_D = P_F$.

Next we shall see that $P'_D = P_F$ implies $P_F \in \mathbb{E}_D$. There are two cases.

(i) \mathbb{F} is consistent with \mathbb{E}_D , so $\mathbb{E}'_D = \mathbb{E}_D \cap \mathbb{F}$. Suppose that $P'_D = P_F$. Since $P_F = P'_D \in \mathbb{E}'_D = \mathbb{E}_D \cap \mathbb{F}$, we have that $P_F \in \mathbb{E}_D$.

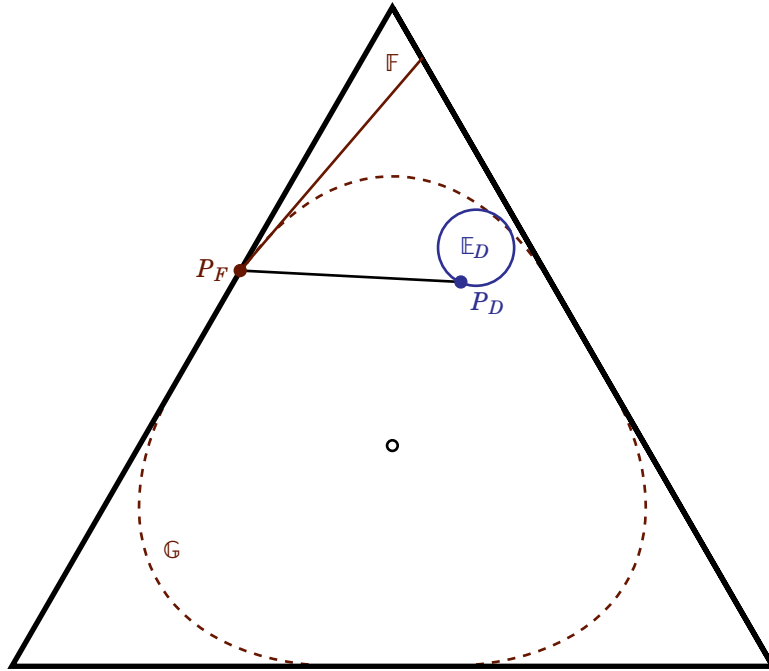
(ii) \mathbb{F} is inconsistent with \mathbb{E}_D , so $\mathbb{E}'_D = \langle \mathbb{E}_D \cup \mathbb{F} \rangle$. In this case $P_F \notin \mathbb{E}_D$ and we need to show that $P'_D \neq P_F$. Geometrically this can be seen as follows:



Since \mathbb{E}_D lies outside \mathbb{F} , some part of the line segment between P_F and P_D will lie inside the contour of probability functions that are equally far from P_* as P_F (the dashed curve above). This part of the line segment contains points in $\mathbb{E}'_D = \langle \mathbb{E}_D \cup \mathbb{F} \rangle$ that are closer to the equivocator than P_F , so it is not possible that $P'_D = P_F$.

This picture generalises to higher dimensions. As long as P_F is not on the boundary of the probability simplex, it lies on a locally smooth convex surface of points that are the same distance from P_* as P_F , and region \mathbb{F} is bounded by the hyperplane tangential to this surface at P_F . Call this case (a). Since the line segment between P_F and P_D does not lie on that tangent plane (nor indeed in region \mathbb{F}), and the tangent plane is the only supporting plane at P_F of $\mathbb{G} \stackrel{\text{def}}{=} \{P \in \mathbb{P} : P \text{ is at least as close to the equivocator as } P_F\}$, the line segment must intersect the contour surface and contain points that are closer to the equivocator than P_F , so $P'_D \neq P_F$.

On the other hand, if P_F is on the boundary of the probability simplex then there is no tangent plane in the usual sense. Case (b): if P_F is a vertex of the probability simplex then \mathbb{G} is the simplex itself and any non-vertex is more equivocal than P_F . In particular, any point sufficiently close to P_F on the line segment from P_F to P_D is more equivocal than P_F , so $P'_D \neq P_F$. Case (c): otherwise P_F is on the boundary but not a vertex and the contour surface meets the boundary at P_F . This case is pictured in two dimensions below:



Here we have a combination of the previous two cases, as there are two or more supporting hyperplanes to \mathbb{G} : the faces of the simplex on which P_F lies, and the ‘one-sided’ tangent plane of the contour surface at P_F defined as the limit of a sequence of tangent planes to the contour surface at points on the surface (but not on the boundary of the simplex) that approach P_F . These planes also bound \mathbb{F} . Then, depending on the location of the line segment between P_F and P_D , it either crosses the contour surface as in (a), or else passes through a region of points closer to the equivocator as in (b). Either way, $P'_D \neq P_F$, as required. \square

Thus objective Bayesianism does provide grounds for adopting a forecaster’s probability function when that function is consistent with the decision maker’s evidence.¹¹

¶ Presuppositions. We shall now examine how three of the presuppositions behind the above analysis can be relaxed.

While we have thus far supposed that the forecaster reveals her whole probability distribution P_F over the partition of outcomes of interest, this supposition is not essential. Suppose instead, for example, that outcomes $\omega_1, \dots, \omega_{2^n}$ are the atomic states $\pm A_1 \wedge \dots \wedge \pm A_n$ of a propositional language with propositional variables A_1, \dots, A_n , and that forecaster F reveals only that $P_F(\theta_1) = x_1, \dots, P_F(\theta_k) = x_k$, for sentences $\theta_1, \dots, \theta_k$ of the propositional language. So the forecaster reveals that $P_F \in \mathbb{P}_F$,

¹¹One might worry whether Theorem 1 is consistent with the Obstinacy Principle of Paris (1999), which says, ‘receiving evidence supporting what one already thinks should not cause one to alter one’s views’ (Paris, 1999, p. 80). If $\mathbb{F} \cap \mathbb{E}_D \neq \emptyset$ then the decision maker’s new evidence is compatible with her prior evidence and one might worry that Obstinacy demands that her belief function should not change, i.e., that P'_D should be P_D rather than P_F . (I am grateful to Hykel Hosni for raising this concern.) But in fact there is no inconsistency here. In the set-up of this paper, Obstinacy implies that $P'_D = P_D$ if $P_D \in \mathbb{F}$. On the other hand, Theorem 1 implies that $P'_D = P_F$ if $P_F \in \mathbb{E}_D$. These are consistent because if $P_D \in \mathbb{F}$ and $P_F \in \mathbb{E}_D$ then $P_F = P_D$.

where $\mathbb{P}_F \stackrel{\text{def}}{=} \{P \in \mathbb{P} : P(\theta_1) = x_1, \dots, P(\theta_k) = x_k\}$ is a closed convex set of probability functions (in fact it is an affine subspace of the simplex). Decision maker D can infer that there is evidence that constrains the forecaster's probability to lie in region \mathbb{E}_F that is itself contained in the region \mathbb{F} determined by the tangent to the contour line at the point $\hat{P}_F \stackrel{\text{def}}{=} \text{the function in } \mathbb{P}_F \text{ that is closest to the equivocator } P_{\perp}$. Theorem 1 then implies that $P'_D = \hat{P}_F$ iff $\hat{P}_F \in \mathbb{E}_D$. In particular, if D has no evidence relevant to $\theta_1, \dots, \theta_k$ then $\hat{P}_F \in \mathbb{E}_D$ so $P'_D = \hat{P}_F$ and $P'_D(\theta_1) = x_1, \dots, P'_D(\theta_k) = x_k$, i.e., the decision maker adopts the forecaster's announced probabilities:

Corollary 2 *If F reveals that $P_F(\theta_1) = x_1, \dots, P_F(\theta_k) = x_k$ and D has no prior evidence bearing on $\theta_1, \dots, \theta_k$, then $P'_D(\theta_1) = x_1, \dots, P'_D(\theta_k) = x_k$.*

A second presupposition underlying the above discussion is that closeness to the equivocator is to be measured using KL-divergence. KL-divergence is the natural measure of closeness when the loss function in the decision scenario is not known in advance; however, a case can be made for using other measures given particular loss functions (Williamson, 2010, §3.4.4; Grünwald and Dawid, 2004). It should be apparent that the key features of the closeness measure that were required in the proof of Theorem 1 were that a closed convex set of probability functions should have a unique member that is closest to the equivocator, and that tangent planes should be definable at each point on a contour surface (a set of probability functions that are equally close to the equivocator), which normally requires differentiability of the contour surface. Other divergence functions that have these properties can be substituted for KL-divergence in the above analysis.

A third presupposition is that the decision maker's prior evidence still counts for something after the forecaster's probabilities are revealed. Thus \mathbb{E}'_D is $\mathbb{E}_D \cap \mathbb{F}$ or $(\mathbb{E}_D \cup \mathbb{F})$, which are symmetrical with respect to the agent's prior evidence region \mathbb{E}_D and the inferred region \mathbb{F} containing the forecaster's evidence. If, on the other hand, D defers to F 's evidence by abandoning her own previous evidence and setting $\mathbb{E}'_D = \mathbb{F}$, then switching to the forecaster's probabilities will always be justified: $P'_D = P_F$ simply because P_F is the most equivocal function in \mathbb{F} . (The difference between these two policies corresponds to the distinction between *belief merging* and *belief revision* in the theory of qualitative belief change.)

¶ *Caveats.* Admittedly it is debatable whether the Met Office follows the norms of objective Bayesianism sufficiently closely for an objective Bayesian decision maker to apply the above justification and adopt Met Office probabilities. While the Met Office certainly does not follow the norms precisely as explicated above, it does directly calibrate its models to frequency data, and it does equivocate insofar as it doesn't announce extreme probabilities in the absence of evidence that forces them to be extreme. (The Met Office does announce 'weather warnings' in situations in which extreme weather is consistent with evidence.) In any case, this concern is not exclusive to objective Bayesianism. Whatever one's theory of uncertainty, the decision maker needs to be confident that the forecaster is forecasting in a rational way in order to justify using the forecaster's probabilities for decision making. While strict subjective Bayesianism and empirically-based subjective Bayesianism have difficulty providing such a justification even where the forecaster is following the appropriate norms, we have seen that the objective Bayesian can provide such a justification.

A second caveat should be borne in mind. The discussion of this paper is based on the situation in which a single forecaster gives a single forecast. The picture

obviously becomes more complicated when multiple forecasters are introduced, and also when the decision maker has some grounds (such as the long-term forecasting history of each forecaster) for assessing the reliability of forecasts. There is already a substantial literature offering treatments of the more complicated situation from the perspectives of statistics and machine learning (see, e.g., Skouras and Dawid, 1999; Cesa-Bianchi and Lugosi, 2006). How these complications relate to the discussion of this paper is left as an interesting question for further research. The focus is on the single-forecast case here because there remains a need for a satisfactory interpretation of this more simple situation. From the philosophical perspective, at least, one needs to walk before one can run.

§6

Conclusions

The main conclusions of the paper are as follows. While probability forecasts are indeed useful for decision making, there is a natural escalation in the level of uncertainty at which those forecasts should be pitched: however one expresses the uncertainty that attaches to a forecast, there are good reasons for expressing further uncertainty about that uncertainty, either by invoking imprecision or higher-order uncertainty. Bayesian epistemology provides the means to avoid rising endlessly up this escalator, by tying expressions of uncertainty to the propositions expressible in an agent's language. But Bayesian epistemology comes in three main varieties. Strict subjective Bayesianism and empirically-based subjective Bayesianism have difficulty in justifying the use of a forecaster's probabilities for decision making. On the other hand, objective Bayesianism can justify the use of such probabilities, at least when they are consistent with the agent's evidence. Hence objective Bayesianism offers the most promise overall for explaining how testimony of uncertainty can be useful for decision making.

Some of the lessons of this paper carry over to a rather different concern: that of justifying the Principle of Reflection, which says roughly that one ought to set one's current degrees of belief to one's future degrees of belief, should one know them. Here one's future self stands in the place of the forecaster. Subjective Bayesians often adopt this principle by fiat, as it is closely connected with Bayesian conditionalisation which plays a central role in subjective Bayesianism.¹² But Bayesian conditionalisation plays a less central role in objective Bayesianism: in many circumstances new objective Bayesian probabilities obtained by the procedure outlined in §5 concur with what might have been obtained by Bayesian conditionalisation, but in certain cases they do not (Williamson, 2010, Chapter 4). The question then arises as to whether a Principle of Reflection should be adopted and, if so, what form it should take. Corollary 2 shows that objective Bayesians *should* endorse a Principle of Reflection: in the absence of any prior evidence bearing on the relevant propositions, if an agent learns some of her future degrees of beliefs then she ought to set her current degrees of belief to those future degrees of belief right away. Moreover, one can drop the qualification about prior evidence if the agent is prepared to defer to her future evidence.

The Principle of Reflection is usually formulated using conditional probabilities: $P(\theta|P'(\theta) = x) = x$. As Howson (2012, §5) notes, this formulation is inconsistent with

¹²There is some debate as to what the precise connection between the two principles is. While van Fraassen (1995, §§5–6) argues that conditionalisation implies reflection, this is disputed by Weisberg (2007).

the axioms of probability: if $P(\theta) = 1$, $P(P'(\theta) = x) > 0$ and $x < 1$, the axioms of probability imply that $P(\theta|P'(\theta) = x) = 1 \neq x$. Mirroring the less central role played by Bayesian conditionalisation in objective Bayesianism, conditional probabilities also play a less central role, and the objective Bayesian would avoid such a formulation (Williamson, 2010, §4.4.2). If the agent had no prior evidence concerning θ , then learning that $P'(\theta) = x$ would tell her that she will receive evidence that the physical probability $P^*(\theta)$ is no more equivocal than x , so she can set $P(\theta) = x$ since that is the most equivocal value compatible with her new evidence (Corollary 2). If, however, her prior evidence forced $P(\theta) = 1$ then the new value $P'(\theta) = x < 1$ does not satisfy prior evidence. Everything then depends on how the prior evidence is treated. If the prior evidence were retained, then the method of §5 would apply and Theorem 1 would ensure that $P(\theta) \neq x$. On the other hand, if she were to defer to her later evidence, she would switch to $P(\theta) = x$.

While the objective Bayesian approach to reflection provides a principled response to Howson's concern, it also overcomes other problems that beset reflection. For example, I believe to degree 1 that I ate a sandwich for lunch today, θ , but I realise that in a year's time my degree of belief in θ will be much lower, because I will no doubt forget what I ate today long before then (Talbot, 1991, §2). Given that I eat a sandwich for lunch about four times a week, I am confident that my future degree of belief that I ate a sandwich for lunch today will be $\frac{4}{7}$. Rather dubiously, the usual principle of reflection would seem to require that I should now adopt degree of belief $\frac{4}{7}$ in θ . But objective Bayesianism would not require this. Let us apply the analysis of §5. My current evidence forces $P(\theta) = 1$, so $\mathbb{E} \subseteq \{Q \in \mathbb{P} : Q(\theta) = 1\}$. I then realise that my evidence in a year's time will, by the Calibration norm, force $P'(\theta) = 4/7$, so $\mathbb{F} \subseteq \{Q \in \mathbb{P} : Q(\theta) \geq 4/7\}$. Setting aside changes of evidence in the intervening year that are unrelated to θ , it is plausible that $\mathbb{E} \cap \mathbb{F} = \mathbb{E}$. Hence I should continue to adopt degree of belief 1 in θ for the time being, as seems rational.

We see, then, the although objective Bayesianism does admit a version of the principle reflection, it is not the usual version. In fact, neither does the objective Bayesian version coincide with van Fraassen's General Reflection Principle (van Fraassen, 1995, §4), which can be formulated in our framework as follows. Consider what one's current evidence tells one about a future belief function P' . Clearly, as far as current evidence is concerned, $P' \in \mathbb{P}$, the set of all probability functions. Suppose that \mathbb{Q} is the smallest convex set of probability functions such that, as far as one's current evidence is concerned, $P' \in \mathbb{Q}$. Then one's current belief function should lie in that set, i.e., $P \in \mathbb{Q}$.

Again, the approach of §5 reveals the extent to which objective Bayesianism conforms to such a principle. Since current evidence implies $P' \in \mathbb{Q}$, it implies something about one's future evidence, namely that that $\mathbb{E}' \subseteq \mathbb{F}$, where \mathbb{F} is determined by the most equivocal function $Q \in \mathbb{Q}$ as in §5. Now $\mathbb{E} \cap \mathbb{F} \neq \emptyset$ for otherwise $\mathbb{E} = \langle \mathbb{E} \cup \mathbb{F} \rangle$ which implies that $\mathbb{F} \subseteq \mathbb{E}$, a contradiction. So $\mathbb{E} = \mathbb{E} \cap \mathbb{F}$. Hence $\mathbb{E} \subseteq \mathbb{F}$. What does this imply about one's current belief function P , the most equivocal function in \mathbb{E} ? There is clearly no general guarantee that $P \in \mathbb{Q}$. Nor does there seem to be good reason to insist that $P \in \mathbb{Q}$ as a new norm on objective Bayesianism.

On the other hand consider this slightly different formulation of the General Reflection Principle: suppose that \mathbb{Q} is the smallest convex set of probability functions that, as far as one's current evidence is concerned, contains the set \mathbb{E}' of probability functions compatible with one's future evidence; then one's current belief function should lie in that set, i.e., $P \in \mathbb{Q}$. In this case, current evidence implies more than $\mathbb{E}' \subseteq \mathbb{F}$: it implies that $\mathbb{E}' \subseteq \mathbb{Q}$. Applying similar reasoning to that used above, $\mathbb{E} \cap \mathbb{Q} \neq \emptyset$

so $\mathbb{E} = \mathbb{E} \cap \mathbb{Q}$ and hence $\mathbb{E} \subseteq \mathbb{Q}$. In particular, then, $P \in \mathbb{Q}$. So this version of the General Reflection Principle does hold for objective Bayesian degree of belief.

Acknowledgements

I am very grateful to the UK Arts and Humanities Research Council for funding this research, and to Kristoffer Ahlstrom-Vij, David Corfield, Hykel Hosni, Colin Howson and two anonymous referees for helpful discussion and comments.

References

- Cesa-Bianchi, N. and Lugosi, G. (2006). *Prediction, Learning, and Games*. Cambridge University Press, Cambridge.
- de Finetti, B. (1937). Foresight. its logical laws, its subjective sources. In Kyburg, H. E. and Smokler, H. E., editors, *Studies in subjective probability*, pages 53–118. Robert E. Krieger Publishing Company, Huntington, New York, second (1980) edition.
- Diaconis, P. and Freedman, D. (1980). De Finetti's theorem for Markov chains. *Annals of Probability*, 8(1):115–130.
- Gaifman, H. and Snir, M. (1982). Probabilities over rich languages. *Journal of Symbolic Logic*, 47(3):495–548.
- Gelman, A. and Shalizi, C. R. (2013). Philosophy and the practice of Bayesian statistics. *British Journal of Mathematical and Statistical Psychology*, 66:8–38.
- Goldstein, L. and Blum, A. (2008). When is a statement not a statement?—When it's a Liar. *The Reasoner*, 2(2):4–6.
- Good, I. J. (2003). The accumulation of imprecise weights of evidence. <http://www.sipta.org/isipta03/jack.pdf>. Contributed to the 3rd International Symposium on Imprecise Probabilities and their Applications.
- Grünwald, P. and Dawid, A. P. (2004). Game theory, maximum entropy, minimum discrepancy, and robust Bayesian decision theory. *Annals of Statistics*, 32(4):1367–1433.
- Howson, C. (2012). Modelling uncertain inference. *Synthese*, 186(2):475–492.
- Jaynes, E. T. (1957). Information theory and statistical mechanics. *The Physical Review*, 106(4):620–630.
- Jaynes, E. T. (1982). On the rationale of maximum-entropy methods. *Proceedings of the IEEE*, 70(9):939–952.
- Kass, R. E. (2011). Statistical inference: the big picture. *Statistical Science*, 26(1):1–9.
- 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.
- Miller, D. (1966). A paradox of information. *British Journal for the Philosophy of Science*, 17:59–61.
- Palmer, T. N. and Hardaker, P. J. (2011). Handling uncertainty in science. *Philosophical Transactions of the Royal Society A*, 369:4681–4684.
- Paris, J. B. (1999). Common sense and maximum entropy. *Synthese*, 117:73–93.
- Ramsey, F. P. (1926). Truth and probability. In Kyburg, H. E. and Smokler, H. E., editors, *Studies in subjective probability*, pages 23–52. Robert E. Krieger Publishing Company, Huntington, New York, second (1980) edition.
- Schervish, M. J. (1995). *Theory of statistics*. Springer, New York.
- Skouras, K. and Dawid, A. P. (1999). On efficient probability forecasting systems. *Biometrika*, 86(4):765–784.

- Slingo, J. and Palmer, T. (2011). Uncertainty in weather and climate prediction. *Philosophical Transactions of the Royal Society A*, 369:4751–4767.
- Talbott, W. J. (1991). Two principles of Bayesian epistemology. *Philosophical Studies*, 62:135–150.
- Toulmin, S. E. (1958). *The uses of argument*. Cambridge University Press, Cambridge, updated 2003 edition.
- Troffaes, M. C. M. (2007). Decision making under uncertainty using imprecise probabilities. *International Journal of Approximate Reasoning*, 45:17–29.
- van Fraassen, B. C. (1995). Belief and the problem of Ulysses and the sirens. *Philosophical Studies*, 77:7–37.
- Weisberg, J. (2007). Conditionalisation, reflection and self-knowledge. *Philosophical Studies*, 135:179–197.
- Williamson, J. (2010). *In defence of objective Bayesianism*. Oxford University Press, Oxford.
- Williamson, J. (2011). Why frequentists and Bayesians need each other. *Erkenntnis*, 78(2):293–318.