Making Fit Fit

Mike T. Hicks

Sept. 9, 2016

Contents

1 The Best System and the Principal Principle **3**
1.1 The Best Systems Account 3
1.2 The Chance-Credence Link 5

2 Fit and Accuracy **7**
2.1 Measuring Accuracy 7
2.2 Fit, Accuracy, and the BSA 8
2.2.1 Simplicity: Conservatively Modified 8
2.2.2 Simplicity: Radically Modified 9
2.2.3 Simplicity and Evidence 10

3 Chance and Epistemic Utility **11**
3.1 An Epistemic Utility Argument for the Principal Principle 11
3.2 Vindicating Chance 12

4 Conclusion **13**

Abstract

Reductionist accounts of objective chance rely on a notion of *fit*, which ties the chances at a world to the frequencies at that world. Here, I criticize extant measures of the fit of a chance system and draw on recent literature in epistemic utility theory to propose a new model: chances *fit* a world insofar as they are *accurate* at that world. I show how this model of *fit* does a better job of explaining the normative features of chance, its role in the laws of nature, and its status as expert function than do previous accounts.

words:4987

Introduction

Objective chances have two roles. They are part and parcel with the laws of nature. With the laws they play a role in determining–probabilistically–how the world unfolds. But they also have a normative role. Chances ought to impact our partial beliefs, or credences. Our credences should match the chances.

Humeans about objective chance have followed Lewis (1980) in understanding objective chances through their relationship with the laws of nature and in taking principles linking chance to credence to be central to our understanding of objective chance. The Humean theory of laws is well-suited to connect the two roles of chance: laws, according to a Humean, are true generalizations that together maximize simplicity and informativeness. Objective chances, then, are just like laws, except that informativeness is measured by their impact on credences rather than on full belief.

Lewis (1980, 1994) calls the measure of informativeness for chancy laws 'fit'. Laws *fit* a world in proportion to how likely they make that world. This way of understanding the informativeness of probabilisitic laws is flawed: it offers too mediated a connection between the chances and our credences. In this paper I employ notions from epistemic utility theory to advance a better measure of the informativeness of probabilistic laws. The view developed has two primary advantages over the traditional Lewisian view: first, it provides a more direct explanation of the fact that chances are probabilities; second, it provides a simpler explanation of the normative force of chances on credences. On the view advanced here, the chances are better fit to play their normative role than on the Lewisian proposal.

The view put forward in this paper is inspired by the following guiding principle: Humean laws are those generalizations which sit between induction (from experience) and deduction (of predictions and explanations). Traditional Humean views–those of Hume, Mill, Ramsey, and Lewis–have focused on the deductive aspect of laws[1]. The laws, on this view, are the best way of organizing all facts, whether or not those facts or the laws are accessible to agents in the world or scientists.

These criticisms of and refinements to that orthodox view aim to bring the laws down to us. The chief advantage of the ideal Humean view is that it shows why embedded agents have a use for laws, causation, and counterfactuals. The role laws have to play here is in extending our knowledge from the observed to the unobserved. To do this, they must be epistemically accessible to agents embedded in the world. Such a system is better, not just if it contains more information, but also if it is more conducive to empirical discovery. This requires us to look for generalizations which can be easily discovered through observation as well as easily extended to prediction.

The paper is structured as follows. In §1 I provide a critical overview of the Humean theory of laws and chance. I'll then discuss the form of chance-credence linking principles. The chance-credence link I will focus on for the majority of the paper is Ned Hall's New Principle (Hall (1994)). Taken together, these provide an opinionated introduction to the two aspects of chance mentioned earlier. In §2 I develop an accuracy-based account of informativeness for probabilistic

[1]So Ramsey says "Even if we knew everything, we should still want to organize our knowledge in a deductive system [...] what we do know we tend to organize in a deductive system and call its axioms laws" (Ramsey, 1927).

laws and improve on the Humean notion of simplicity for chance-functions. In §3 I show how the metaphysical picture of chance defended here fits with one accuracy-based argument for the New Principle.

1 The Best System and the Principal Principle

The role of chances in explanation is similar to that of laws. The trajectory of the cue ball and the eight ball at t_0 explains their trajectories at t_1 *in virtue of* the law of conservation of momentum. The fact that the cards in the deck are half red explains that half of my draws are red *in virtue of* the fact that I was equally likely to draw each card in the deck[2].

Because of the similarities between law and chance in underwriting explanations and causal facts, we should expect our accounts of law and chance to be closely connected. In §1.1 I'll outline an account of chance and law that connects them: the regularity theory of law, which fits neatly with modified frequentist accounts of chance.

The second face of chance looks to our partial beliefs. If the chance of an atom's decay in the next twenty minutes is .5, I should be just as confident that it will decay as I am that it will fail to decay. But if I know that the chance of drawing a blue ball from an urn is .25, I would be irrational to take a draw of blue to be more likely than not. In §1.2 I'll outline some proposed principles linking chance to credence, and show how they fit with this account of laws.

1.1 The Best Systems Account

The regularity theory of law holds that laws are true generalizations; the frequentist account of chance holds that chances are actual frequencies. A naïve regularity account holds that all true generalizations are laws; a naïve frequentist account holds that all chances are exactly equal to actual frequencies.

Both naïve views are clearly false. Not all true generalizations are laws: some, like the fact that all of the eggs in my refrigerator are brown, are merely accidental. Laws support counterfactuals, but this generalization about my refrigerator does not: were I to buy white eggs, they would not become brown when I put them in the fridge. Similarly, not all chances precisely match their frequencies. For suppose the actual number of coin flips were odd; then it would be impossible for exactly half of the flips to be heads. But it is ridiculous to suppose that a fair coin could not be flipped an odd number of times. Similarly, the quantum mechanical probability of some events is an irrational number, but as actual relative frequencies are ratios of integers they cannot be irrational.

Refined frequency and regularity accounts circumvent these counterexamples by restricting the generalizations which are laws and loosening the link between chances and frequencies. The most developed such account is David Lewis'

[2]This explanation does not go by way of a principle of indifference; rather, the fact that I'm equally likely to draw each card is a result of my shuffling, which is a chancy dynamic process.

(1980, 1994) Best System Account (BSA) of laws and chances. Lewis's account holds that the laws and chances are those generalizations and chance-statements which form deductive system which maximizes three virtues: *strength*, *simplicity*, and *fit*.

We'll call a set of generalizations and chance statements a '*lawbook*'. Each lawbook receives a score for each virtue. Lawbooks get higher marks in strength for implying many true propositions (or, equivalently, ruling out many worlds). Lawbooks get higher marks in simplicity for being syntactically shorter: by having fewer and shorter sentences[3]. Lawbooks get high marks in fit for having chances which are close to the frequencies of the world. The points in each category are tallied; the lawbook with the most overall points comprises the scientific laws of the world.

The BSA escapes the counterexamples to the naïve regularity and frequentist accounts by evaluating lawbooks holistically and invoking a simplicity consideration. Because lawbooks get higher marks for simplicity when they have fewer sentences, not every true generalization will be a law–only those whose contribution to the inferential power of the lawbook outweighs their cost in syntactic length. Similarly, Lewis suggests that a lawbook may contain probabilities that diverge from the actual frequencies for simplicity considerations. For example, a lawbook may assign 0.5 to a coinflip coming up heads despite the actual frequency of heads being 0.4999999.

A brief advertisement of this approach: if we take the laws to be the theorems of the best system, defined in Lewis' sense, we can give a boilerplate justification of their counterfactual robustness. The laws underwrite counterfactuals because they are organizationally central to our belief set: the laws are amongst a small set of beliefs (because of their simplicity) from which many other propositions in our belief set follow (because of their strength). If we remove the laws from our belief set, this will generate a large change in the set.

When evaluating counterfactual or indicative conditionals, we employ the Ramsey test: we first add the antecedent of the conditional to our belief set and make as few changes as possible[4]. Because of the organizational centrality of the laws, removing them requires large changes in our belief set. So when evaluating conditionals, we hold the laws fixed.

[3]Understanding simplicity in terms of syntactic complexity makes our simplicity requirement language dependent. This means either we must hold that there is exactly one preferred language of evaluation or that generalizations are laws only relative to a choice of language. Lewis held the former; for an exploration of the latter view see Callendar and Cohen (2010, 2011). For a mixed view, on which the simplicity of a lawbook is evaluated both by its syntactic length and the complexity of the translation between the lawbook's preferred language and our language, see Loewer (2007)

[4]Which beliefs we change depends in interesting ways on the type of conditional we're evaluating. When we evaluate "if Oswald didn't kill Kennedy, someone else did" we change fewer beliefs than when we evaluate "if Oswald hadn't killed Kennedy, someone else would have." In both cases, however, we hold the laws fixed. One task of a metaphysics of laws is to explain this constancy; the proponent of the BSA can do so by appealing to our conservatism in changing our beliefs, the epistemic centrality of laws, and our use of the Ramsey test in evaluating conditionals. Proponents of more metaphysically robust accounts of law must do so, apparently, by making brute stipulations about their metaphysical danglers.

Now for some brief criticism: Lewis' account of fit, like his account of strength, applies to worlds as a whole, rather than local events. Just as a system may be strong but allow inferences only from global states of the world to other global states, a system may fit well but provide little information about the likelihood of local events. Such systems are almost completely useless to embedded agents, who can neither discover nor apply them. This should worry us; our laws should be *ours*, both available and useful to agents like us.

Furthermore, it's not clear how simplicity should lead us away from the actual frequencies of the world. In what sense is .4999999 more complex than .5? Both are rational numbers. Perhaps we should count number of decimal places. But if this is correct we have a problem: quantum mechanics often assigns events irrational probabilities. These have infinite decimal expansions and so are by this measure maximally unsimple. But since irrational numbers are not ratios of integers, they cannot be the actual value of frequencies. So simplicity measured in this way cannot explain why quantum chances diverge from frequencies.

I'll explore an alternative constraint in §2.2. The alternative constraint I offer is more closely tied to our epistemic access to the laws. Strength is counterbalanced in laws, not by a virtue measuring the ease by which they can be expressed, but instead by a virtue measured by how easily they can be discovered[5]. First, however, I'll provide an overview of the orthodox view of the second face of chance: its relationship to our partial beliefs.

1.2 The Chance-Credence Link

The connection between chance and credence is codified by chance-credence norms. Three are currently active in the literature: the Principal Principle (Lewis (1980)), the New Principle (Hall (1994), Thau (1994), Lewis (1994), and the General Principal Principle (Ismael (2008)). I'll discuss each of these in turn, and then give reason to prefer the New Principle.

PP: $b(A|T_i\&E) = ch_i(A)$

As a constraint on initial credences, the Principal Principle tells us to match our credences to the chances. That is, if we think that the correct chance theory assigns a chance of x to A, we should have a credence of x in A, whatever other (admissible) evidence we have.

Unfortunately there is a problem combining the Principal Principle with the reductionist account of chances outlined in §1.1. For recall that *fit* requires the chances to match the frequencies as much as simplicity will allow. Hence arbitrary mismatches between the chances and the frequencies are metaphysically impossible. As is customary I will call such mismatches *undermining futures* because they provide evidence against the chance function, and call chance

[5]It's not obvious that this is best called 'simplicity', as that term seems tied to syntactic measures of complexity. But there is plenty of research on the curve-fitting problem and in Bayesian confirmation theory which argues that simpler theories are more quickly arrived at or provide more accurate predictions (see, for example Rosenkrantz (1979) and Henderson (2013); for a Humean account which understands simplicity in this way see Hoefer (2007)).

functions which allow them *modest* chance functions. A modest chance function is so-called because it does not assign itself a chance of one–it is not certain that it is the right chance function.

As an illustration, suppose T is a chance theory that that assigns .5 to coinflips landing heads. If I know that the coin will be flipped exactly 400 times, what credence should I assign to *every* flip coming up heads? Because T will only be the chance theory if the frequency of heads is near 0.5, if T is the chance theory this implies that *all heads* is false. We should give no credence to the metaphysical impossibilities[6], so we should set $b(allheads|T) = 0$. But if $ch(heads) = .5$, then $ch(allheads|400flips) = 3.87 \times 10^{-121} > 0$. So by PP, $b(allheads|T) = ch(allheads) > 0$. Hence the Principal Principle leads to inconsistent constraints on our credences: either we should have nonzero credence in the metaphysically impossible, or we should diverge our credences from the chances.

In response to this worry, reductionists about chance have embraced the New Principal (Hall (1994), Thau (1994), Lewis (1994)):

NP: $b(A|T\&E) = ch(A|T\&E)$

NP tells us to set our credences to the chances conditional on all of our evidence, *including the fact that they are the chances*. In the special case of non-self-undermining chance functions, $ch(T) = 1$, $ch(A|T) = ch(A)$, and NP agrees with PP. But if a chance function assigns positive chance to propositions with which it is incompatible, $ch(T) < 1$, and so for undermining propositions U $ch(U|T) = 0 \neq x = ch(U)$. Thus $b(U|T) = ch(U|T) = 0$ without contradiction.

Finally, Ismael's (2008) General Principal Principle:

GPP: $b(A) = \Sigma_{ch_i:chi(E)>0} b(T_i)ch(A|E)$

Because GPP places no direct constraints on an agent's conditional credences, it's compatible with GPP that an agent have conditional credence $b(A|T_i) = 0$ when $ch_i(A) \neq 0$. However GPP has other unpalatable consequences: Pettigrew (2013a) shows that GPP is inconsistent with Bayesian conditionalization, and that GPP requires us to assign credence zero to chance functions which are themselves modest but which assign a nonzero probability to immodest chance functions.

Because NP is the least internally problematic principle, and because it fits best with other expert principles, I will focus on it for the remainder of the paper.

[6]It's open to the defender of chance reductionism to avoid the counterexample to the Principal Principle by rejecting this claim; surely one should not be dogmatically certain of one's metaphysical views. But we regard chance as a sort of epistemic expert. Whether or not we are certain of the metaphysical truths, chance, as an expert, ought to be. If it is, contradiction follows.

2 Fit and Accuracy

We now have in sight the two sides to the chasm we wish to bridge. On the one side we have an explication of chance's status amongst the laws (§1.1). On the other we have a formal principal linking it to our credences (§1.2). The question we now wish to answer is: why does the chance theory which best balances fit with simplicity deserve our deference in the sense given by NP?

For the same reason, I hold, that we defer to those that we consider to be experts. It's natural to count someone as an expert just in case she is maximally accurate; we should defer to another agent in some domain if and only if we take that agent's beliefs to be better approximations of the truth than ours. Similarly we should defer to the objective chances if and only if they are closer to the truth than our partial beliefs. Consequently, we should take *fit*, the virtue which measures how informative a chance theory is, to be a measure of accuracy: closeness to the truth.

2.1 Measuring Accuracy

There is a growing literature on accuracy measures for credence functions (Joyce (1998), Leitgeb and Pettigrew (2010a)). The idea is simple: we take one credal function to be maximally accurate at a world. Call this the *vindicated function.* If our aim is truth, we will take the vindicated function at a world w to be the truth-function at w, v_w, where the truth function assigns 1 to truths and 0 to falsehoods.

We then devise a measure of *distance* between credal functions. The accuracy of a credal function then is its distance from the vindicated function. The distance measure is a function $D(b_i, b_j)$ which measures the distance between two credal functions. If both functions are the credences of agents, the distance measure will tell us by how much they disagree. The standard distance measure takes *distance* to be the sum of the squared differences between the credences of the agents in each proposition:

Distance: $D(b_i, b_j) = \Sigma_{A \in F}(b_i(A) - b_j(A))^2$,

where we assume for simplicity that the two agents have credences in the same (finite) set of propositions, F.

Now that we have a maximally accurate function v_w and a distance measure, we have all we need to compare the accuracy of partial beliefs at a world w. For we can define the *in*accuracy of a credal function $b(*—*)$ as distance from maximal accuracy:

Inaccuracy: $I(b, w) = \Sigma_{A \in F}(b(A) - v_w(A))^2$

Now that we have in hand a notion which will allow us to compare the accuracy of credal functions, we can state the thesis that the fit of a chance theory varies with respect to its accuracy; chance theory T has a higher *fit* score than T* at a world w if and only if ch_T is less inaccurate than ch_{T^*}.

Fit As Accuracy: $fit(T, w) > fit(T^*, w) \equiv I(ch_T, w) < I(ch_{T^*}, w)$

I will not argue directly for FIT AS ACCURACY. Rather, I will show how FIT AS ACCURACY provides a better explanation of the normative force of the NP, more directly explains the fact that chances are probabilities, and overcomes putative objections. FIT AS ACCURACY does better than traditional measures of fit in some respects and worse in none[7].

2.2 Fit, Accuracy, and the BSA

Minimizing inaccuracy is a good first step for out account of the chance function. But it should be clear that the chance function cannot merely be that ur-credence function which minimizes inaccuracy. For we know what this function is: it's the truth-function. And we know that the truth function is not the chance function.

The BSA, of course, has a response to this: fit trades off against simplicity. But Lewis is unclear about how one chance function can be simpler than another, as I discussed in §1.

Here I'll make two new proposals.

2.2.1 Simplicity: Conservatively Modified

The first involves replacing the history-to-chance conditionals with functions from fundamental quantities to chances. Developing this proposal requires a little groundwork: first, we should think of the Humean mosaic as composed of fundamental quantities; these quantities are determinables, and their values determinates. For example, *mass* is a fundamental quantitative determinable of which *1g* is a determinate. Thinking of the mosaic this way has two advantages: first, it's closer to the way scientists think the world is[8]. We can allow the chances to be a function with a variable for each fundamental quantity, which takes as values determinates of this quantity. The chances, then, will be given by a function from quantity variables at points to probabilities of quantity variables at points. For example, quantum mechanics gives probabilities for some quantities (e.g., position after measurement) as a function of the values of other quantities (the amplitude of the wavefunction).

We can evaluate the simplicity of this function the same way we do that of deterministic laws: by looking at its syntactic complexity when it is stated using variables corresponding to the fundamental quantities. But we might want to add other requirements as well: for example, we may prefer chance functions which are continuous and so are such that small changes in values of our variables

[7]Here may be the best time to address a worry the reader may have: how does fit as accuracy deal with worlds, like ours, which have infinitely many chancy events? At least as well as traditional measures of fit. Like traditional measures of fit, we can trump up a finite set of 'test propositions', and measure the accuracy of the chance function in terms of these, as presented in Elga (2004). If we have a countable, ordered set of events we can also take the inaccuracy of the system to be the limit $lim_{n \to \infty} \Sigma_{A_1, A_n} (b(A_i) - v_w(A_i))^2$.

[8]'Closer' because gauge freedoms at the fundamental level leave open whether the fundamental qualities of the world are more akin to graded relations than quantities. This debate is ongoing, and not one that I wish to engage here.

correspond to small changes in the chances. We may also prefer to rank the simplicity of such theories not (or not only) by their syntactic simplicity, but instead include *symmetry* considerations. A chance theory which respects spacial and temporal symmetries is much more practical than one which does not. If the chance of an outcome depends not just on local features of a chance setup but instead on the setup's location in spacetime, embedded agents (folks like us) would be unable to divine the correct chance setup by repeating qualitatively identical experiments[9].

2.2.2 Simplicity: Radically Modified

My second proposal follows Hoefer (2007) and Ismael (2013) in taking conditional probabilities to be primitive, where '$ch(A|E) = x$' means that the chance that an event of type E is of type A is x. We can then take the simplicity of the chance function to depend only on how many different situation-types it distinguishes between.

We can understand this typing extensionally by taking each type to be a set of events, so that x is of type E just in case $x \in E$. Each chance function will partition the events of the world into types, the set of which, T, is such that $E \in T$; we can then compare the simplicity of chance theories by comparing their typing schemes. If all of the types $E \in T$ are subsets of the types $E^* \in T^*$, then T^* is a simpler typing scheme than T, and a chance theory C based on T^* is simpler than a chance theory C* based on T.

The former proposal §2.2.1 requires appeal to fundamental properties, an appeal which some versions of the BSA eschew. For example, Callender and Cohen (2009, 2010) develop a best system account on which each science has its own stock of fundamentalia; Loewer (2007) argues that we should evaluate the complexity of a system with respect to its informativeness in a macroscopic language, and judge its choice of fundamental properties by the ease in which they can be used to give us information about the middle-sized goods we typically encounter.

This latter proposal has no direct reliance on fundamental properties; instead, we evaluate the simplicity of a chance theory by evaluating the simplicity of its language.

Both proposals allow us to take the simplicity consideration to be tied to our evidence for the chances. For the simpler a chance theory is, on either measure, the more opportunities we have to observe it. The connection between frequency and chance goes in two directions: we gain information about the frequencies from the chances, and we gain information about the chances from the frequencies. In order for the chances to be epistemically accessible to us, we

[9]This requirement is noted by Arntzenius and Hall (2003:179) who note that "Your recipe for how total history determines chances should be sensitive to basic symmetries of time and space–so that if, for example, two processes going on in different regions of spacetime are exactly alike, your recipe assigns to their outcomes the same single-case chances." The difference between the view advocated by Arntzenius and Hall and the view here advocated is that they take respecting spatiotemporal symmetries to be a requirement on chance functions rather than merely a goodmaking feature.

need to be able to infer them from observation. And in order to observe them, we need a broad class of events whose outcomes are assigned the same chance. So simpler chance theories are more *epistemically accessible*.

2.2.3 Simplicity and Evidence

Our chance function should deliver to us the most accurate beliefs available to us given our evidence. I've just argued that this requires the chance function to be invariant over a broad enough set of events for us to observe the frequency of outcomes for those event types; it also requires the chance function to yield the same chances for two situations which are indistinguishable. If, prior to performing observing an outcome of a chance setup, E and E* cannot be distinguished, then $ch(A|E) = ch(A|E^*)$. This gives us an absolute lower bound for the simplicity of our chance theories: they must respect evidential equivalence:

> Evidential Equivalence: For all A, E, and E*, If no evidence can distinguish between E and E*, then $ch(A|E) = ch(A|E^*)$.

Evidential Equivalence may at first seem circular: our notion of what counts as evidence for what is tied to our notion of laws and objective chances. How, then, can we determine whether two setups are evidentially indistinguishable before we know what the laws and objective chances are? The worry, then, is that this requirement is toothless. What it is, we may think, for two setups E and E* to be *distinguishable* is for $ch(A|E) \neq ch(A|E^*)$. This worrisome thought is mistaken.

First, it is false that we have no notion of what counts as evidence for what prior to our account of laws and chances. For we don't yet have the final theory of laws and chances, but we do have a lot of true beliefs as to which situations are evidentially indistinguishable. Two double-slit experimental setups are evidentially indistinguishable, provided they're made of the same materials and are the same size, even if they are in different laboratories. Two shuffled decks of cards are evidentially indistinguishable before the first card is drawn.

Second, even as an *internal* requirement on packages of law and chance, this requirement has teeth. For it requires setups to differ in *more than the chance of their outcomes*. For two setups to be assigned different outcome chances, there must be some *other* difference between them, either in terms of their internal distribution of fundamental properties or their causal history.

To sum up: I hold that chance functions fit the world better when they are more accurate, as measured by the Brier score. But the chances are not maximally accurate because they must respect two evidentialist constraints: first, we must be able to gain evidence for them by observing frequencies. Second, they cannot make distinctions between events which exceed our ability to distinguish between those events. The chance function, then, is the most accurate credal function for which we can gain evidence through observation of frequencies and employ to constrain our credences about future events.

3 Chance and Epistemic Utility

This definition of chance fits neatly into an argument for the New Principle given by Pettigrew (2012). In §3.1 I will present Pettigrew's (2012) proof. In §3.2 I will argue that the account of chance I have provided can be used to underwrite a key premise in Pettigrew's argument.

3.1 An Epistemic Utility Argument for the Principal Principle

Pettigrew's (2012) proof is based on the notion that chance is to credence as truth is to full belief. Our credences, holds Pettigrew, should aim at the chances just as our beliefs should aim at the truth. To represent this goal, Pettigrew introduces the *chance-based Brier score* as a measure of epistemic utility:

CBBS: $I(b, w) = \Sigma_{A \in F}(b(A) - ch_w(A|E))^2$,

where E is the total evidence admissible to an agent at a time. We will say that CBBS measures epistemic utility in the presence of E. Taking *cbbs* as a measure of epistemic utility follows from the claims that (1) the ideal credence function at w is $ch_w(*|*)$ and (2) the *dis*utility of a credal function is proportional to its distance from the ideal function, as measured by the sum of squared differences. Note that the CBBS is just INACCURACY from §2 with the 'vindicated' function taken to be the chance function rather than the truth function.

The remaining premises in Pettigrew's argument are imported from Joyce (1998). First, we assume (3) *dominance*: if credence function b has a higher epistemic utility than $b*$ at all worlds, and there is no other credence function that has a higher epistemic utility than b at all worlds, then it is irrational for an agent to employ $b*$.

Pettigrew shows that (1), (2), and (3) together imply that agents are irrational to adopt credal functions which do not obey the axioms of probability and the Principal Principle–and that by slightly tweaking CBBS we can show that agents must obey the New Principle. Pettigrew's proof relies on taking the chance function to be a probability function.

Pettigrew provides little support for claim (1), that the ideal credence function at w in the presence of E is $ch_w(*|E)$. His argument rests on the claim that chance is to credence as truth is to full belief, which Al Hájek has defended in **unpublished work**. But it is difficult to see how this claim could be defended without a metaphysical account of chance. We take truth to be the aim of belief not as a basic posit but because it comports well with our theories of truth. On the most naïve correspondence theory, the true propositions represent the actual world. This world-dependence is what makes truth an appropriate aim for belief at @; in order to accept the claim that our credences should aim at the chances, we need a similar account of the world-dependence of the chances.

Pettigrew also fails to support the claim that the chances must be probabilities. To my knowledge, there are two defenses of the claim that the chances must be probabilities. The first relies on frequentist or hypothetical frequentist accounts of chance: frequencies, as ratios of outcomes, obey the probability axioms, so the chances must be probabilities[10]. But this justification relies on a false theory of chance, as discussed in §1.1. The second justification goes *via* the Principal Principle–the chances are those things to which we are rationally required to match or credences, it is irrational to have nonprobabilistic credences, thus the chances must be probabilities. But in the context of Pettigrew's proof, this is circular. A new vindication is given below.

3.2 Vindicating Chance

If the chances obey our two constraints, we have an argument for the claim that the chance function is the ideal credence function and a reason to believe that the chance function is a probability function. I'll start by arguing that it is irrational to fail to respect the chance function arrived at by taking the lower bound of simplicity, that is, the least simple chance function compatible with EVIDENTIAL EQUIVALENCE.

Recall that the chances, according to FIT AS ACCURACY and EVIDENTIAL EQUIVALENCE, are the credence function that maximizes accuracy while also maximizing discoverability (§2.2.1, §2.2.2) and treating evidentially indistinguishable setups as equivalent. The chance function, then, is the most accurate credence function *which obeys the same evidential constraints that we do*. In this sense, then, chance is to credence not as truth is to belief, but as *knowledge* is to belief[11].

Now suppose that someone knowingly fails to match their credences to the chances. Then she is either employing a credence function which fails to respect EVIDENTIAL EQUIVALENCE or she is employing a credence function which respects EVIDENTIAL EQUIVALENCE but is accuracy-dominated by another such function. If she fails to respect evidential equivalence, then her credences in the outcomes of chance setups depend on more than her evidence; this is irrational. And if she respects EVIDENTIAL EQUIVALENCE but is accuracy-dominated, she is also irrational–because she is knowingly employing a credence function farther from the truth than she could. Hence if she fails to respect the chances she is irrational.

This gives us reason to believe that our credences should aim, not directly at the truth, but instead at the objective probabilities. For just as our goal for full belief is not merely to have true beliefs, but to have knowledge, our aim for partial belief should not be merely to have accurate beliefs, but to have accurate and well-supported beliefs. Since our credences cannot be more accurate (while retaining evidential support) than the chance function, we should take the objective chances at our world to be the target of our credences.

Finally, this account of accuracy puts the claim that chances are probability functions on more solid ground.

By making accuracy a constitutive feature of chance, we are able to show that the same constraints that require our credences to be probabilities require the

[10]This view is apparent in **Ramsey (19??)**

[11]Or perhaps as *justified belief* is to belief.

chances to be probabilities: namely, any non-probabilistic chance function will be accuracy dominated by a probabilistic one (see Joyce (1998)). But since our chances must be the most accurate credal function at some grain of simplicity, they will not be accuracy dominated. So they will be probabilities.

4 Conclusion

Lewis proposed a Humean account of chance which sought to directly link the chances to the frequencies at a world. I've provided a similar account which ties the probabilities not directly to frequencies, but instead to the *accuracy* of a chance function. I've also provided a refined account of the *simplicity* of a chance function, which more directly links simplicity to our evidence for a chance theory. I've argued that this account of chance underwrites a proof of the PP.

References

[1] Callender, Craig, and Jonathan Cohen (2009). "A Better Best System Account of Lawhood" *Philosophical Studies* 145 (1): 1-34.

[2] Callender, Craig, and Jonathan Cohen (2010). "Special Science, Conspiracy, and the Better Best System Account of Lawhood." *Erkenntnis* 73: 427-447

[3] Hall, Ned (1994). "Correcting the Guide to Objective Chance." *Mind* 103 (412): 505-518.

[4] Hájek, Al (FILL THIS IN)

[5] Hoefer, Carl (2007). "The Third Way on Objective Probability: A Skeptic's Guide to Objective Chance." *Mind* 116 (463): 549-596.

[6] Ismael, Jenann (2008). "Raid! Dissolving the big, bad bug." *Noûs* 42 (2):292–307.

[7] Ismael, Jenann (2009). "Probability in Deterministic Physics." *Journal of Philosophy* 86 (2): 89-109.

[8] Leitgeb, Hannes & Richard Pettigrew (2010). "An Objective Justification of Bayesianism I: Measuring Inaccuracy." *Philosophy of Science* 77 (2):201-235.

[9] Leitgeb, Hannes & Richard Pettigrew (2010). "An Objective Justification of Bayesianism II: The Consequences of Minimizing Inaccuracy." *Philosophy of Science* 77 (2):236-272.

[10] Loewer, Barry "Laws and Natural Properties." *Philosophical Topics* 35 (1/2):313-328.

[11] Lewis, David (1980) "A Subjectivist's Guide to Objective Chance." In Richard C. Jeffrey (ed.), *Studies in Inductive Logic and Probability*. University of California Press: 83–132

[12] Lewis, David (1994). "Humean Supervenience Debugged." *Mind* 103 (412): 473–490.

[13] Meacham, Christopher J. G. (2010). "Two Mistakes Regarding the Principal Principle." *British Journal for the Philosophy of Science* 61 (2):407-431.

[14] Pettigrew, Richard (2012). "Accuracy, Chance, and the Principal Principle." *Philosophical Review* 21:2 241-275.

[15] Pettigrew, Richard (2013a). "Accuracy and Evidence." *Dialectica* 67 (4):579-596.

[16] Pettigrew, Richard (2013b). "A New Epistemic Utility Argument for the Principal Principle." *Episteme* 10 (1):19-35.

[17] Schaffer, Jonathan (2003). "Principled Chances" *British Journal for the Philosophy of Science* 54 (1):27-41.

[18] Strevens, Michael (1999). "Objective Probability as a Guide to the World." *Philosophical Studies* 95 (3):243-275.

[19] Thau, Michael (1994). "Undermining and Admissibility." *Mind* 103 (412): 491-504.

[20] Joyce, James (1998). "A Non-Pragmatic Vindication of Probabilism" *Philosophy of Science* 65 (4):575-603.