DALE A. NANCE

THE WEIGHTS OF EVIDENCE

ABSTRACT

Interest in the Keynesian concept of evidential weight has led to divergent views concerning the burden of proof in adjudication. It is argued that Keynes's concept is properly engaged only in the context of one special kind of decision, the decision whether or not the evidence is ripe for a decision on the underlying merits, whether the latter decision is based on probability, relative plausibility, coherence or otherwise. As a general matter, this question of ripeness is appropriately assigned to the judiciary for resolution as part of the burden of production, rather than to the jury or other factfinder as part of the burden of persuasion.

1. TWO SENSES OF WEIGHT

For lawyers, the most familiar sense of evidential weight can be described as the "convincing force" or "persuasive value" of a mass of evidence. A verdict is rendered by comparing it with the applicable standard of proof. In a civil case, for the party bearing the burden of persuasion to be entitled to a favorable verdict, ordinarily the weight of the evidence must satisfy the standard of proof "by a preponderance of the evidence" in favor of that party's claim, while in a criminal case, for the prosecution to be entitled to a favorable verdict, the weight of the evidence ordinarily must be so strong as to render the prosecution's claim true "beyond reasonable doubt" (Strong 2006, 483).

Another sense of weight is less familiar to lawyers, at least under the name of "weight." In recent years, it has become known to legal academicians as a result of the work of several scholars, most notably Jonathan Cohen (Cohen, L. 1977). Cohen took inspiration from Keynes, whose early work on probability included the following insight:

> As the relevant evidence at our disposal increases, the magnitude of the probability of the argument may decrease or increase, according as the new knowledge strengthens the unfavourable or the favourable evidence; but something seems to have increased in either case,–we have a more substantial basis upon which to rest our conclusion. I express this by saying that an accession of new evidence increases the weight of an argument. New evidence will sometimes decrease the probability of an argument, but

DOI: 10.3366/E1742360008000385

https://doi.org/10.3366/E1742360008000385 Published online by Cambridge University Press

> it will always increase its "weight." … [W]eight, to speak metaphorically, measures the sum of the favourable and unfavourable evidence, probability measures the difference (Keynes 1921, 71, 77).

Keynes's point readily generalizes to any theory about the convincing force of evidence, whether or not articulated in terms of probabilities and, if so, whether or not based on the particular "logical" theory of probability that Keynes endorsed.[1] Weight in the first sense is the degree to which a rational decision-maker is convinced of the truth of a proposition as compared to some competing hypothesis (which could be simply that the proposition is false). It is a function of the extent to which available evidence interlocks to form a sound explanatory account favoring the proposition over the competing hypothesis (Haack 1993, ch. 4). But this degree of weight can be based on a relatively thin mass of evidence or a relatively rich one, the difference between which concerns weight in the second sense. Weight in the latter sense depends on the relative completeness of evidence, the extent to which the evidence addresses the important inferential questions that are raised by the competing hypotheses, including questions of the reliability of the items of evidence considered (Stein 2005, 91–106). Again, an increase in weight in the second sense can increase, decrease, or leave unchanged the weight of the evidence in the first sense.

An abstract example will add precision. From a baseline of given evidence with regard to a proposition and a competing hypothesis, consider the effect of conducting a test for which a positive result supports the proposition and a negative result supports the competing hypothesis. Obtaining *some* result (without regard to its value) is what increases weight in the Keynesian sense, but it has an indeterminate effect on convincing force. Conversely, *which* value is obtained (positive or negative) is what affects convincing force, but it has no effect on Keynesian weight. Notice that Keynesian weight is subject to the decision-maker's control in a sense that convincing force is not. Within practical limits, Keynesian weight can be *chosen*, provided the decision-maker has the capacity to cause such tests to be conducted, whereas for a given level of Keynesian weight, convincing force can only be *assessed* (assuming of course that the decision-maker does not independently control the test result).

These two dimensions of evidence are fundamental to decision-making under conditions of uncertainty. Someone contemplating action faces two decisions that correspond to these two senses of weight. When a certain action is desirable if and only if some proposition is true, then one's choice depends on the convincing force of the available evidence regarding its truth. (For example, the truth or falsity of the proposition might affect the expected total utility of the action.) But that decision can be displaced by a second: the decision whether to decide about the contemplated act based on the available information or, in the alternative, to postpone that "primary" decision in order to collect additional information relevant to the predicate proposition. That "decision-about-deciding" relates to

https://doi.org/10.3366/E1742360008000385 Published online by Cambridge University Press

Keynesian weight. A rational choice requires two determinations, however implicit, corresponding to these two concepts of weight (Keynes 1921, 76–7).[2]

Following Keynes's metaphor, I refer to these two kinds of weight – these two dimensions of the total mass of evidence – as "Δ-weight" and "Σ-weight," respectively. Both are matters of degree (Walker 1996, 1095, as to Δ-weight; Cohen, L. 1985, as to Σ-weight). In contrast, many practical decisions that are taken based on the weight of the evidence, certainly decisions about legal liability, are binary: either liability is imposed or it is not. Consequently, degrees of weight must be *compared* with something in order to reach decisions – compared either to degrees of weight favoring an alternative decision or to a standard that is independently specified, some critical level of weight, which I will call W_{Δ}^* or W_{Σ}^*, as the context requires.

Legal adjudication is more complicated than a private decision, most notably because of the detailed procedures involved and because different aspects of the decision can be, and often are, assigned to different decision-makers. In fact, a number of attempts have been made in recent years to incorporate the Keynesian insight into a theory of legal burdens of proof, but these efforts have generally gone astray in failing to recognize the implications of the distinction between these two decisions. The tendency of legal theorists has been to try to incorporate a consideration of Σ-weight into the primary decision, thus becoming part of the burden of persuasion. This, I contend, is a mistake, and correcting this mistake will give us a clearer picture of the roles of weight in the theory of burdens of proof.

2. Δ-WEIGHT AND THE BURDEN OF PERSUASION

Before addressing the role of Σ-weight, it is useful to review one well-developed decision-theoretic model of the burden of persuasion. It assumes a meaningful probability or odds can be assigned to the truth of the allegations of the person bearing the burden of persuasion, whom I will call the "claimant." If O(C|E) denotes the odds in favor of the claim, C, given the evidence, E, then the standard (if a bit oversimplified) result is that minimizing the expected error costs requires a decision for the claimant if and only if

$$O(C|E) > D(+)/D(-) =_{\text{def}} O^* \tag{1}$$

where D(+) is the utility loss attributable to a false positive decision, i.e., an erroneous decision for the claimant, and D(−) is the utility loss attributable to a false negative decision, i.e., an erroneous decision against the claimant (Kaplan 1968).[3] (The odds form of the result is stated above because assessment in terms of odds is likely a more natural framework for decision-makers than assessment in terms of probabilities. The translation is: probability = odds/(odds +1). Thus, the decision criterion in probability terms is: decide for the claimant just when $P(C|E) > P^* = O^*/(O^*+1)$.) When the disutility of a false positive is equal to the disutility of a false negative, as is generally thought to be the case in civil trials, this

https://doi.org/10.3366/E1742360008000385 Published online by Cambridge University Press

critical odds, O*, is 1 : 1, which corresponds to a critical probability, P*, of 0.5. But in criminal cases, the disutility of a mistaken conviction, D(+), is generally regarded as much larger than the disutility of an erroneous acquittal, D(−), and this critical probability approaches 1. For example, if D(+) = 10 × D(−), then the critical odds are 10 : 1, or a probability of 0.91.

By most accounts, the specification of O* (or the corresponding P*) – which constitutes the critical level of Δ-weight, W_Δ* – is generally a policy issue to be determined by law-making authority. This critical value is communicated to the factfinder in appropriate language that, for reasons noted below, generally avoids explicit quantification. The factfinder's assessed odds on the claim constitute a measure of the Δ-weight of the evidence, which the factfinder compares to the critical value. Importantly, the theory does not speak to the question of how such odds are to be determined. In this it reflects the fact that the law – in both Anglo-American and Continental courts – generally does not instruct factfinders in any great detail about how they are to go about assessing the weight of the evidence, a principle known as "free evaluation of evidence" (Stein 2005, 108–16).[4]

Several amplifications of this model are useful here. First, the sense of probability involved here can be neither frequentist nor (merely) subjectivist. Frequentist probability works well for modeling stochastic phenomena that exist quite apart from any uncertainty that human evaluators may have about them – phenomena such as radioactive decay of atoms and (in a more subtle way) games of chance – but they do not work well for the unique, non-repeatable events upon which the decision-maker typically must make factual findings in litigation. In contrast, subjective probabilities, which are measures of a person's level of confidence about an event or proposition, are well-designed for the decision context and well-adapted for handling non-repeatable events. Nonetheless, the burden of persuasion cannot be interpreted as solely a question of subjective probabilities. The reason is that the law rightly imposes certain ancillary conditions designed to assure that the decision-maker's assessment of probability is both well-considered and productive of accurate verdicts. For example, institutional arrangements such as jury deliberation or the requirement that judges in bench trials explain their findings of fact are intended to make the assessments well-considered, and many of our rules of admissibility are designed to assure that particular items of evidence presented are veritistic, or "truth-conducive," in attempting to avoid predictable errors of inference by factfinders. It is, perhaps, arbitrary whether we conceive of these practices as *part* of the burden of persuasion or rather as ancillary institutions, but it cannot be denied that they attempt to impose on the factfinder's assessment of probability something stronger than mere coherence plus good faith belief.[5] The probability that is compared with P* is more like what Ian Hacking enticingly calls "interpersonal/evidential" probability (Hacking 2001, 130). My view is that the pertinent idea is subjective or (in the context of a jury) intersubjective probability reached through a defined, veritistic procedure, the combination of

https://doi.org/10.3366/E1742360008000385 Published online by Cambridge University Press

which is designed to function as the best available proxy for knowledge (cf. Walker 1996).

Second, there are many reasons to think that the critical value cannot be fixed by law-making authorities at a specific value that is appropriate for all cases, or even all cases of a broad category, such as "all civil cases" or "all criminal cases." The most obvious reason is that the exact ratio of specified disutilities will depend on the situation of the individual case and may be reasonably controverted. Retail adjustment by the factfinder may be necessary in some cases (Walker 1996, 1124; Lillquist 2002). The operative language of legal rules generally allows for this.

Finally, the model presented above assumes that a cardinal measure of the Δ-weight of evidence is practically meaningful. All that is necessary is that the factfinder be willing to commit to odds on the truth of the claim. And these odds need not be terribly precise. The factfinder must be able to conclude that $O(C|E)$ is greater than O^*, but there is no need to say by how much.

Although the foregoing model will figure in the following discussion, the basic thrust of the arguments presented does not depend upon being able to model Δ-weight successfully in such cardinal terms. It is enough to note that those who insist on limiting measures of Δ-weight to an ordinal form generally accept that there is a tipping point in the level or degree of certainty (or the level of "warrant") required to impose legal liability and that the appropriate tipping point depends on factors such as the relative utilities of false positives and false negatives (e.g., Walker 1996, 1116–20). With or without a cardinal measure, what is involved is an assessment of Δ-weight, not Σ-weight (1095 n.45).

3. Σ-WEIGHT AND THE STRUCTURE OF LITIGATION

The case for cardinal measures of Σ-weight is less persuasive. For Δ-weight, there are two natural extreme points, certainty that the claim is true and certainty that the claim is false. Placing these at the ends of a scale of credal confidence permits one to divide the "space" in between in accordance with the ratio constituting the odds, thus obtaining a cardinal measure. But trying the same thing with Σ-weight leads to severe difficulties. What is to constitute the complete absence of evidence? No decision-maker approaches a decision task on a *tabula rasa*. Inference is a socially conditioned process that proceeds from a background of more or less shared information, information that is as much evidence as is information freshly acquired. Without the "null evidence" endpoint, a ratio is beyond our reach, even if a common metric for Σ-weight could be developed (cf. Keynes 1921, 72).

Still, Σ-weight can sometimes be compared or ordinally ranked (Cohen, L. 1985). Most importantly, with respect to two competing hypotheses, any mass of evidence E_2 will have greater Σ-weight than mass E_1 if E_2 contains all evidence in E_1 that is relevant to the competing hypotheses plus some additional evidence relevant thereto. The relationship between such increases in Σ-weight and accuracy

https://doi.org/10.3366/E1742360008000385 Published online by Cambridge University Press

of decision is complex. Adding relevant evidence does not guarantee a more accurate decision or even a reduction in error costs. Yet our practices, born of experience, presuppose that augmentation of Σ-weight is epistemically desirable as a general matter, especially when there is no bias in the augmentation process that systematically selects evidence that favors (or disfavors) the proposition in question. Under certain conditions, it can be shown that taking into account additional costless information increases the expected utility of the primary decision (Skyrms 1990, ch.4).[6]

In vague terms, Σ-weight can be maximized by taking account of all relevant evidence, and this might seem to provide the critical level, W_{Σ}*, required to proceed with a decision on the underlying act. There are, however, several important difficulties in giving precise content to the idea of "all relevant evidence," which affect its serviceability as W_{Σ}*.

Consider the possibility of defining all relevant evidence as "all evidence that might be relevant, the sources of which are extant when the primary decision is made." Call this "maximal" weight. This articulation differentiates between sources that are in existence and conceivable sources that never existed. It makes no sense to think in terms of taking account of evidence the sources of which never existed. But such an articulation is still too broad to function as the critical level of Σ-weight, because it also makes no sense to postpone the primary decision in order to acquire evidence that cannot be considered by the decision-maker at reasonable cost in the time frame during which a decision must be made. This leads to an articulation of W_{Σ}* in terms of the information practically derivable from all extant sources that can reasonably be made available to and considered by the decision-maker (Cohen, L. 1985, 265). This more restricted notion is a first cut at the idea of "optimal" weight.

At the same time, the specified sense of maximal weight is too *narrow* to function as optimal weight because it excludes evidence from sources that once were reasonably available but by the time of deliberation and decision are no longer. To be sure, such information cannot be taken into account and in many inferential contexts rightly plays no role. But in the legal context, interested parties who are not the decision-maker can affect what information is lost. A preemptive decision about the adjudication may be required to create appropriate *ex ante* incentives for interested persons not to destroy relevant evidence. To allow for this, I suggest that optimizing Σ-weight requires all (admissible) evidence from sources that reasonably can be made available to the factfinder or (in the case of lost evidence) should have been made available by a party. As thus characterized and enforced, optimality would not always be obtainable, but it would always be a condition precedent for making a decision on Δ-weight.

Of course, determining whether it is cost-justified to try to obtain a given piece of missing evidence is not always an easy matter, but then neither is deciding whether the standard for Δ-weight has been met. The critical points remains: for a given mass of evidence, the Σ-weight of the evidence in relation to the one

https://doi.org/10.3366/E1742360008000385 Published online by Cambridge University Press

hypothesis is the same as the Σ-weight of the evidence in relation to the competing hypothesis (Cohen, L. 1985, 268–70). Consequently, measures of Σ-weight cannot determine the choice between the two. When Σ-weight is optimized, it no longer plays a role in the decision, even if it is not maximized in some sense like that stated above; Δ-weight assessment is all that remains.[7] Conversely, when Σ-weight is not optimized, the decision priority is to eliminate the Σ-weight deficiency, when that is possible, and to take appropriate preemptive action when it is not. The two weight assessments ordinarily do not interact, except to the extent that a decision-maker's assessment of whether an increase in Σ-weight is cost-justified depends on a tentative assessment of Δ-weight.

The goal of optimizing Σ-weight in this sense is undeniably a part of the law and practices of adjudication. All modern systems of adjudication attempt to create appropriate incentives for actors in the legal system to develop relevant evidence for use by the tribunal. In the context of Anglo-American adversarial trials, this means primarily reliance upon the inherent incentives of the parties to discover and present evidence favorable to each side, augmented (especially in civil cases) with extensive mechanisms for compulsory discovery of evidence in the possession of adversaries and third parties. These pretrial devices are supplemented by a variety of "best evidence" exclusionary rules the purpose of which is to encourage the presentation of evidence of greater Σ-weight (Nance 1991).

Such incentive structures, however, are not invariably successful at optimizing evidence production. Aside from evidence that is subject to a valid claim of privilege, sometimes a witness, document, or other tangible object of potential value to the factfinder and reasonably available (at some point) to one or both parties is nonetheless not presented at trial. The causes include: lack of diligence of counsel in conducting investigation and discovery; gaps in the rules of discovery that deny parties' compulsory process for obtaining information in the control of an adversary (especially, but not only, in criminal cases); destruction of evidence by a party; differences of opinion between the factfinder and counsel on the importance of evidence; the reluctance of counsel to present evidence of ambiguous import; and asymmetries in exclusionary rules that prevent only the party favored by the evidence from presenting it (Nance 1998, 631). And the question then becomes: What, if anything, should be done about this? In particular, how should residual deficiencies in Σ-weight be addressed in the doctrines governing the burden of proof?

4. Σ-WEIGHT AND THE BURDEN OF PERSUASION: AN UNHAPPY MERGER

A number of scholars have suggested that considerations of Σ-weight be incorporated somehow into the burden of persuasion. The following paragraphs address the principal strategies that have been explored. These suggestions represent thoughtful and creative efforts to solve an important problem, and at

https://doi.org/10.3366/E1742360008000385 Published online by Cambridge University Press

least arguable legal precedent can usually be found to support each strategy. Still, they are ultimately unconvincing, especially as general solutions to the problem of Σ-weight deficiencies.

Adjusting the Measure of Δ*-Weight.* One strategy is to allow, indeed encourage, factfinders to adjust their assessment of Δ-weight to reflect their concerns about deficiencies in Σ-weight. Thus, if Δ-weight is measured by O(C|E), then these odds could be adjusted to reflect the incompleteness of the evidence (Kaye 1986; Friedman 1997, 1963–5). There is doctrinal support for such a practice in the use of "missing evidence" jury instructions, which tell jurors that they may infer that evidence not presented is adverse to the interests of the party with exclusive (or, sometimes, superior) access to that evidence or the party whom the missing evidence would ordinarily be presumed to favor (Strong 2006, 220–2).

The main difficulty with this idea is that the factfinders do not know the content of the missing information. One can say, crudely, that evidence effectively repressed by one side is likely to be unfavorable to them, but that gives little hint of an answer to the question, "By how much?", which is needed to make optimal adjustments to O(C|E) or some ordinal measure of Δ-weight. The factfinder is put in the position of making highly speculative adjustments, as commentators have noticed (220–4). Which poses the real question: why not avoid this situation entirely by insisting that Σ-weight be optimized as a condition on submitting the case to the factfinder for a verdict on the merits? Furthermore, the suggested remedy is not suitable at all for a great many cases in which a Σ-weight deficiency is problematic, such as the negligent destruction of evidence by a party without any intention of affecting the outcome of the litigation, or the failure of the parties to present evidence that is equally available to each side (Nance 1998, 633–9).

Adjusting the Critical Level of Δ*-Weight Required.* It has also been argued that a deficiency in Σ-weight is one of the factors by virtue of which the factfinder (or perhaps the judge) ought to adjust W_Δ*; a factfinder might consider litigant misconduct in suppressing evidence as a reason to reduce the disutility associated with a false verdict against that party (Walker 1996, 1124–6; Friedman 1997, 1971–2). Indeed, McCormick noted that, in some cases, nominal adjustments to the measure of Δ-weight are probably really adjustments of W_Δ* to reflect a change in the pertinent utilities:

> A question may well be raised whether the relatively modest probative value of this species of evidence is not often outweighed by its prejudicial aspects. The litigant who would not like to have a stronger case must indeed be a rarity. It may well be that the real underpinning of the rule of admissibility is a desire to impose swift punishment, with a certain poetic justice, rather than concern over the niceties of proof (Strong 2006, 228).

It has already been noted that the factfinder should have some discretion to tailor W_Δ* to the specific case. The question is whether a deficiency in Σ-weight is an appropriate basis for such adjustment. In some contexts, as when the deficiency reflects badly on one party (or rather, reflects worse on one party than the other),

https://doi.org/10.3366/E1742360008000385 Published online by Cambridge University Press

this is at least possible. In other contexts, however, the Σ-weight deficiency will not reflect badly on either party or their levels of fault may be indistinguishable, yet some kind of legal response may be appropriate.

Even in the former contexts, McCormick's comment about "prejudice" suggests caution: using evidence about bad acts of a party not integrally a part of the events of the litigated claim presents a risk of what the law has generally considered improper prejudice, and one of the commonly accepted reasons for considering it improper is that it may involve the *sub rosa* adjustment of W_{Δ}* (Lempert 1977, 1038–9). Why should that be considered improper? The main reason is posed by this rhetorical question: if the law deems it appropriate to sanction litigants for such behavior, ought we not be explicit about that, rather than hiding it in a declaration by the factfinder that the party lost on the merits because the evidence did not support the truth of that party's claim?

Formulating the Δ*-Weight Standard in Terms of* Σ*-Weight Criteria.* The principal champion of this approach is Jonathan Cohen, who has argued that the decision criteria for adjudication should be understood in terms of whether a measure of Σ-weight exceeds an appropriate standard (Cohen, L. 1977). Cohen's development of this idea is quite complex, but it rests on the idea that Σ-weight can be ranked on the basis of how many tests of an hypothesis have been conducted. For each adjudication, there will be a maximal number of such tests, and the tests can be ordered in terms of importance, or so it is claimed. The larger the number of tests to which a hypothesis is subjected, the greater the Σ-weight of the evidence. Because Σ-weight does not differentiate between the hypothesis and its negation (or a competing hypothesis), this ranking is supplemented by a determination of the number of tests passed in a manner favorable to each side. For criminal cases, for proof to be "beyond reasonable doubt," all tests must be conducted, thus producing maximal Σ-weight, and the test results must be uniformly favorable to the prosecution. For civil cases, the preponderance of the evidence favors a party if that party's hypothesis survives a larger number of the ordered tests than the opponent's.

There are several difficulties for Cohen's theory. Most importantly, he assumes that each test can be unequivocally resolved in favor of one party or the other. In a surprisingly brief treatment, he models factfinding, in both civil and criminal cases, as a form of presumptive reasoning that proceeds in binary (yes/no) fashion (1977, 247–56). A generalization favoring one side is tentatively defeated by the other side's evidencing an exception to that generalization, to which the first party responds by presenting evidence that defeats the exception raised, and so forth, until one of the parties can proceed no further. (Thus, a "test" seems to be construed as an identifiable series of such exchanges.) Such presumptive reasoning may be useful in *informing* a determination of Δ-weight, but in practice the question of whether a given test is "passed" will almost always be a matter of degree. Consequently, such tests do not yield a series of binary results that can simply be counted, but rather a series of judgments about matters of degree that can only be

https://doi.org/10.3366/E1742360008000385 Published online by Cambridge University Press

synthesized by resort to probability or some other measure of Δ-weight (Schum 1994, 260–1).

Further, to *equate* the results of this method of presumptive reasoning with the Δ-weight assessment is to conflate the method of analysis with the standard that the method is used to inform. This is made clear in Cohen's arguments against the articulation of the Δ-weight standard in terms of subjective probabilities. Cohen insists that the latter fails because it does not "elucidate the reasonings appropriate to jurors and advocates," providing the means with which judges can give "guidance" to jurors about how to assess a case (Cohen, L. 1986, 646, 648). Providing such guidance, however, contravenes the principle of free evaluation of evidence and presupposes that the instruction will be more helpful than confusing. Cohen's ideas are far too complex to satisfy that test. Moreover, if such instruction is meant to stand on its own, allowing the courts to forego Condition (1) or some comparable standard, the instruction must provide something in the nature of a "recipe" for factfinding. For reasons already noted, Cohen's explanation falls short of that as well. If, on the other hand, the contemplated "guidance" is to operate only in conjunction with a distinct Δ-weight standard, then the latter must be doing work for which Cohen does not allow.

Even if Cohen's theory could be developed into a workable system of guidance for factfinders, one that obviates anything like Condition (1), the question would remain whether that system would be desirable. Without fuller development of his theory, this is difficult to discern. As already noted, if we accept Cohen's analysis of the practical implications of his theory, we encounter certain clearly unacceptable prescriptions.[8] Beyond that, we can pose this general question: if (as Cohen himself seems to believe) this developed system would produce results substantially different from the prescriptions of Condition (1), then what reason can be given for deciding cases in ways that we expect will not minimize costs? Finding none in Cohen's work, we are right to reject Cohen's reconceptualization of proof burdens (Stein 2005, 105–6, 144–5).

Requiring Confidence Intervals for the Measure of Δ-*Weight.* Neil Cohen has argued that uncertainty about the factfinder's determination of the posterior probability of the claim should count against the party bearing the burden of persuasion (Cohen, N. 1985). If P* is the critical probability specified in Condition (1), then Cohen's theory would replace Condition (1) with the standard, $P(C|E) > P^* + \varepsilon$, where ε is determined by the degree of confidence that the law expects factfinders to have in their estimate of P(C|E). The greater the required confidence, the larger the ε. Cohen sees his theory as expanding the idea of "equipoise" so that it handles more than just those cases in which the factfinder's best assessment puts P(C|E) at P*, that is, the value that equalizes the expected error costs of a decision for either party (where institutional inertia rightly favors a denial of the claim). Equipoise would embrace cases "too close to call" with the necessary confidence.

Now a more expansive theory of equipoise may be unavoidable from an economic or psychological perspective, but Cohen's particular argument depends

https://doi.org/10.3366/E1742360008000385 Published online by Cambridge University Press

on deficiencies in what I have called Σ-weight, and his argument is unsuccessful, or at least incomplete. When applied to the subjective probabilities that determine the decision under his modified rule, his argument is built on distinguishing between a factfinder's assessment of P(C|E) and the "true" value of the operative posterior probability. Cohen seems to tell us that this true value is given by the factfinder's hypothetical assessment of P(C|T), where T is "the totality of all available evidence" (85). Cohen's T involves essentially what I have called optimized Σ-weight, and the difference between E and T is the Σ-weight deficiency that gives rise to the uncertainty about the "true" posterior probability.

But Cohen has fallen into a trap: if the missing evidence is available, why not take the steps necessary to have it presented and let the factfinder decide whether P(C|T) is greater than P*, thus obviating the need for a confidence interval approach? Instead, Cohen simply assumes that cases should be allowed to go to verdict without optimizing Σ-weight and that the risk associated with failure to do so should fall on the party with the burden of persuasion (86 n.19). In the context of adjudication, where "not deciding is deciding," the only significance of a less than optimal Σ-weight is that the deficiency should be eliminated, not that the verdict should go against the party with the burden of persuasion. If losing the case is the penalty for not eliminating that deficiency (which need not be the case), then this risk should fall on the party properly assigned responsibility to eliminate the Σ-weight deficiency. This implicates a process and standards quite separate from the factfinder's application of the burden of persuasion.

Cohen might try to modify his theory by identifying T with what I have called *maximal* Σ-weight (85 n.18). In that case, merely *optimizing* Σ-weight, with evidence E*, might leave uncertainty surrounding the assessment of P(C|E*) that could be addressed using confidence intervals. Even so, Cohen's theory would not fully address problems of Σ-weight deficiency (and he makes no claim that it would). Moving from E to E* is what does the real work of optimizing Σ-weight. Once Σ-weight is optimized, there remains the question of why the law should choose a decision rule, $P(C|E^*) > P^* + \varepsilon$, that would, by hypothesis, not minimize expected error costs (Kaye 1987; Cohen, N. 1987, 93–5).

Adding Σ-Weight Criteria to Δ-Weight Criteria as Part of the Burden of Persuasion. Alex Stein has correctly recognized that adjudication cannot do without separate assessments of Δ-weight and Σ-weight (Stein 2005, 120). Furthermore, he argues that questions of the adequacy of Σ-weight should not be left to factfinders but should be resolved by judges and legislatures. Stein's reason for this latter conclusion is that Σ-weight inadequacy necessitates risk-allocation, which falls within the domain of political morality and needs to be resolved as a matter of law (12–6, 48–9). This argument is less than fully convincing – juries (and judges acting as factfinders in bench trials) do have at least some legitimate, if residual role in making judgments of political morality – but it does carry some force and it does lead him to what I think is the correct conclusion about the locus of responsibility for assessing Σ-weight.

https://doi.org/10.3366/E1742360008000385 Published online by Cambridge University Press

Nevertheless, Stein's theory contains some very problematic components. Most importantly, he claims that there is risk of error associated with a supposed gap between evidence that optimizes Σ-weight and evidence of "ideal composition" (43) and that this risk should also fall on the party bearing the burden of persuasion (e.g., 78, 85). But he does not explicate what such an "ideal composition" would be, beyond saying that it is more than what I have called optimal, and – despite considerable effort – he does not succeed in making the case that the difference between optimal weight and ideal weight should matter to decision-making at all, let alone why it should count against the party with the burden of persuasion.[9]

5. CONCLUSION

These difficulties of incorporating Σ-weight into the burden of persuasion are particular manifestations of an underlying problem: aside from the question of making the best decision between competing hypotheses, when a decision must be made, there is the preemptive question of whether that decision should be made on the present state of the evidence (cf. Kaye 1987, 76). Conflating these two decisions only confuses matters. Once we determine to keep these decisions analytically distinct, the question arises as to how institutional arrangements, beyond those discussed above in section 3, should reflect the distinction.

The most important principle is that optimizing Σ-weight is generally a judicial function. It should be handled by courts, not juries or judges acting as factfinders. The rules of discovery and admissibility noted in section 3 are all judge-administered, and for good reason. Answering the question of whether a given item of evidence should be presented (or should have been presented) requires a detailed understanding of the practices and limitations of pre-trial discovery and of the strategic considerations of the litigants. This understanding will almost always favor judges over juries, and the same point applies to further measures taken to optimize Σ-weight. Moreover, it is important that such measures be integrated with pre-trial responses to Σ-weight deficiencies in order to avoid inconsistent treatment or "double counting" by factfinders, who are, moreover, in no position to demand the presentation of additional evidence. Finally, as Stein argues, resolving such issues as matters of law permits, though it does not necessitate, the development of appellate case law (and even legislation) on important questions of political morality.

When the judge determines that Σ-weight is sub-optimal, the proper response is to impose a burden of production on one of the parties.[10] In civil cases, to help contain litigation costs, this burden should be assigned to the party with superior access to an extant source of information; in criminal cases, the burden should be assigned to the prosecution unless the accused has sole access to the missing source. For sources that are irretrievably lost, the burden should be placed according to comparative fault principles. Finally, the penalty for failure to meet the obligation imposed by this burden should depend on the seriousness of the breach and the anticipated response of the sanctioned party, but it should often

https://doi.org/10.3366/E1742360008000385 Published online by Cambridge University Press

entail summary determination against that party of any facts likely to be affected by the missing evidence.[11] An elaboration of this set of principles can be found elsewhere (Nance 1998).

This approach puts the matter of Σ-weight in the proper framework: the judicial obligation to assure that a case is ripe for decision pursuant to an appropriate standard of Δ-weight.

REFERENCES

Cohen, L. Jonathan. 1977. *The Probable and the Provable*. Oxford: Oxford University Press.

Cohen, L. Jonathan. 1985. "Twelve Questions about Keynes's Concept of Weight." *British Journal of the Philosophy of Science* 37: 263–78.

Cohen, L. Jonathan. 1986. "The Role of Evidential Weight in Criminal Proof." *Boston University Law Review* 66: 635–49.

Cohen, Neil B. 1985. "Confidence in Probability: Burdens of Persuasion in a World of Imperfect Knowledge." *New York University Law Review* 60: 385–422.

Cohen, Neil B. 1987. "Conceptualizing Proof and Calculating Probabilities: A Response to Professor Kaye." *Cornell Law Review* 73: 78–95.

Friedman, Richard D. 1997. "Dealing with Evidentiary Deficiency." *Cardozo Law Review* 18: 1961–86.

Gillies, Donald. 2000. *Philosophical Theories of Probability*. London: Routledge.

Haack, Susan. 1993. *Evidence and Inquiry: Towards Reconstruction in Epistemology*. Oxford: Blackwell.

Hacking, Ian. 2001. *An Introduction to Probability and Inductive Logic*. Cambridge: Cambridge University Press.

Josephson, John R. 2001. "On the Proof Dynamics of Inference to the Best Explanation." *Cardozo Law Review* 22: 1621–43.

Kaplan, John. 1968. "Decision Theory and the Factfinding Process." *Stanford Law Review* 20: 1065–92.

Kaye, David H. 1986. "Do We Need a Calculus of Weight to Understand Proof Beyond a Reasonable Doubt?" *Boston University Law Review* 66: 657–72.

Kaye, David H. 1987. "Apples and Oranges: Confidence Coefficients and the Burden of Persuasion." *Cornell Law Review* 73: 54–77.

Keynes, John Maynard. 1921. *A Treatise on Probability*. London: Macmillan Press.

Lempert, Richard O. 1977. "Modeling Relevance." *Michigan Law Review* 75: 1021–57.

Lillquist, Erik. 2002. "Recasting Reasonable Doubt: Decision Theory and the Virtues of Variability." *U.C. Davis Law Review* 36: 85–197.

Nance, Dale A. 1991. "Missing Evidence." *Cardozo Law Review* 13: 831–82.

Nance, Dale A. 1998. "Evidential Completeness and the Burden of Proof." *Hastings Law Journal* 49: 621–62.

Nance, Dale A. 2005. "Allocating the Risk of Error: Its Role in the Theory of Evidence Law." *Legal Theory* 13: 129–64.

Pennington, Nancy & Reid Hastie. 1991. "A Cognitive Model of Juror Decision Making: The Story Model." *Cardozo Law Review* 13: 519–57.

https://doi.org/10.3366/E1742360008000385 Published online by Cambridge University Press

Schum, David A. 1994. *Evidential Foundations of Probabilistic Reasoning*. New York: John Wiley & Sons.

Skyrms, Brian. 1990. *The Dynamics of Rational Deliberation*. Cambridge, MA: Harvard University Press.

Stein, Alex. 2005. *Foundations of Evidence Law*. Oxford: Oxford University Press.

Strong, John W. (gen. ed.) 2006. *McCormick on Evidence*. 6th ed. Vol. II. St. Paul, Minnesota: West Group.

Walker, Vern R. 1996. "Preponderance, Probability, and Warranted Factfinding." *Brooklyn Law Review* 62: 1075–136.

NOTES

1 Here, "probability" refers to an attribute satisfying the familiar Kolmogorov axioms, what Cohen refers to as "mathematical" or "Pascalian" probability (1977, 3), and for which Keynes's theory provides one of several competing interpretations (Gillies 2000).

2 Efforts to increase the available information might involve decisions among alternative courses of information gathering activities, presenting the need to make new Δ-weight assessments, which in turn poses new problems of Σ-weight. *Et cetera*.

3 A more complete analysis incorporates the positive utilities of correct decisions (Nance 1998, 623).

4 Considerable thought has been directed to the question of how factfinders do or should go about assessing Δ-weight. Recent theorizing has focused on the insights of probability theory (e.g., Lempert 1977) or on the relative plausibility of the parties' stories (e.g., Pennington & Hastie 1991) and the related idea of "inference to the best explanation" (e.g., Josephson 2001).

5 Here, by "coherence" I mean the property of a set of odds that prevents the person holding them, and who is willing or required to bet on them, from being subjected to sure-loss betting contracts. A person's betting odds satisfy the Kolmogorov axioms if and only if they are coherent in this sense (Hacking 2001, 163–9).

6 If not already well selected, altering W_{Δ}* can reduce total expected error costs, but this familiar improvement is subject to an equally familiar trade-off: decreasing the risk of false positives entails increasing the risk of false negatives, and *vice versa*. The analogous point is not true of the decision-maker's control of Σ-weight: increases in Σ-weight can reduce both kinds of error simultaneously.

7 Thus, Cohen errs when he argues that a criminal prosecution must fail for want of "maximal" weight (which he characterizes as "the totality of relevant evidence" and *not* "the totality of discoverable relevant evidence") if a "vital" eye-witness dies (without the fault of any party) before testifying, that is, even if the testimony of the witness is *not* available and that unavailability is not the fault of either party, and regardless of how high O(C|E) might be (Cohen, L. 1986, 642). Either the word "vital" creates a tautology here or else that position is both normatively implausible and strikingly inaccurate as a description of prevailing practice.

8 See supra note 7.

9 An extended analysis of this point is presented elsewhere (Nance 2005, 138–54).

https://doi.org/10.3366/E1742360008000385 Published online by Cambridge University Press

10 In some contexts, the court may assume the responsibility for adding evidence, although this practice is understandably limited in a system of adversary trials.

11 The exception is the case in which the burdened party is the accused in a criminal case. In order to protect the accused's right to a jury trial, it will often be necessary to adopt the second-best solution of assigning to the factfinder final responsibility for assessing Σ-weight and determining the consequences of a deficiency.

Dale Nance, Professor of Law at Case Western Reserve University, teaches and writes about the law of evidence and general jurisprudence. He is the author of *Law and Justice: Cases and Readings on the American Legal System* (Carolina Academic Press, 2nd ed. 1999) and *Medical Injustice: The Case Against Health Courts* (with Maxwell Mehlman) (American Association for Justice, 2007). In addition to those attending the conference at which this paper was presented, Professor Nance thanks the following individuals for comments on drafts: David Kaye, David Schum, Rosalind Simson, Walter Sinnott-Armstrong, and Vern Walker.

https://doi.org/10.3366/E1742360008000385 Published online by Cambridge University Press