1

Introduction

In adjudication, facts matter. Cases are often disputes over conflicting versions of the facts. In both civil and criminal cases, certain important, so-called ultimate facts are specified as determinative by the applicable substantive law.[1] When there is uncertainty about these facts, and trials become necessary to resolve the dispute, burdens of proof structure the tribunal's factual assessments. In American civil cases, for example, the ultimate facts that define a cause of action or defense usually must be shown by the plaintiff to be true "by a preponderance of the evidence," and in criminal cases, the ultimate facts must be shown by the prosecution to be true "beyond reasonable doubt."[2] The epistemic components of these requirements reflect the fact that they do not involve a surrender to some kind of pure proceduralism, in which the quest for accuracy is ignored in favor of whatever results from fair procedures.[3] Instead, they reflect the necessity of judgment under uncertainty and the need to exercise that judgment in a way that makes the best use possible of our unavoidably fallible assessments of the facts. This, at any rate, is the premise on which the following account will build.

But what exactly does it mean to prove a civil case "by a preponderance of the evidence"? Or to prove a criminal case "beyond reasonable doubt"? Much appellate ink has been spilled, and many issues settled, on how to formulate these standards verbally, on which standard applies in which kinds of cases, and on the applicability of yet other intermediate standards to some classes of cases.[4] Nevertheless, fundamental questions

[1] More precisely, ultimate facts are the facts constituting necessary components of a set of facts sufficient to instantiate a cause of action or affirmative defense. They are distinguishable from other facts that, once evidenced, can be the basis for inferring an ultimate fact. Cf. Wigmore (1937) at 9–11 (distinguishing between a "factum probans" and a "factum probandum," the former used to draw an inference to the latter).

[2] 2 McCormick (2013) §§ 339, 341.

[3] See Rawls (1971) at 85–6. For a discussion of the challenges of pure proceduralism in the adjudicative context and its associated fact skepticism, see Kaptein (2009).

[4] Most American jurisdictions have an intermediate standard applicable to certain civil cases or certain issues in civil cases, a standard often characterized as requiring "clear and

about what these standards mean remain deeply controversial.[5] In this book I address a set of issues critically important to answering these questions. Specifically, I explore the relationships among three ideas that infuse modern scholarship regarding the burdens of proof. Clarification of these ideas and their relationships promises significant advances in our understanding of the proof process.

Before stating these ideas, a few preliminary comments are in order. The present discourse concerns proof of "adjudicative" facts, facts concerning the conduct of parties to litigation that trigger the applicability of substantive legal rules – such as the fact that the defendant's conduct caused the injury to the plaintiff or the fact that the accused had the intent to kill the deceased. Courts sometimes must assess a different category of facts, so-called legislative facts, facts that are pertinent to determining the content of the legal rules (whether substantive or procedural) that are to be applied in the litigation. For example, if a court is choosing between two possible common-law rules (or two possible interpretations of a statutory rule), one factual matter that it might consider is a comparison of the consequences that can be expected to flow from one or the other rule candidate. These are matters not submitted to the trier of fact at all. There is little regulation in terms of the burdens of proof in the latter context, and I shall not address this important question here.[6]

As is well known, the "burden of proof" on ultimate adjudicative facts has two aspects, at least in Anglo-American procedure. The "burden of persuasion" is that aspect of the burden of proof that, while specified by law, is applied by the trier of fact. It poses the question whether the standard of proof (such as "beyond reasonable doubt") has been satisfied. This is to be distinguished from the "burden of production," that aspect of the burden of proof that is applied (almost exclusively) by the court.[7]

convincing evidence." See 2 McCormick (2013) § 340. In other common-law jurisdictions, the matter is more complicated and ambiguous, but standards exist that are also clearly considered intermediate between that applied in the ordinary civil case and that applicable to the prosecution in criminal cases. See Redmayne (1999).

[5] See, e.g., Stein (2005); Laudan (2006); Ho (2008); and Clermont (2013).

[6] See 2 McCormick (2013) § 331 (discussing standards for taking judicial notice of legislative facts). I also do not address the standards involved in circumventing the usual proof process by the taking of judicial notice of adjudicative facts. See Id. §§ 329–30.

[7] Id. §§ 336, 338. The main reason for the qualifier ("almost exclusively") is that the assistance of the jury is sometimes enlisted in applying presumptions that shift the burden of production by requiring the jury to determine whether or not the "basic facts," those that trigger the presumption, are true. Id. § 344. As we will see, there is another complication. See § 4.2, infra.

Under various circumstances, that burden requires the court to determine whether the evidence that has been produced at trial is good enough to warrant a determination by the trier of fact.[8] Our initial puzzle obviously relates to the burden of persuasion, but we will see that the burden of production is inevitably brought into the discussion.

One component of the burden of persuasion is the *allocation* question: which party is to bear the burden, however heavy that burden may be? This is functionally equivalent to the question of which ultimately determinative facts will be assigned to the complainant as part of that party's affirmative case and which will be assigned to the defendant as elements of an affirmative defense. It also determines the "default" rule; that is, it specifies which party is to be awarded the verdict if the fact-finder is unable to determine whether or not the burden of persuasion with respect to those ultimate facts has been satisfied. The case law and commentary have identified a variety of factors that supposedly control the decision about this allocation question. Some of them are more coherent than others.[9] They include

1. The policy of placing the burden on the party who seeks to change the status quo (unhelpful without an explanation of why, in cases that are litigated, the status quo is presumptively just or otherwise desirable);
2. The policy of placing the burden on the party whose actions necessitate engaging the public machinery of adjudication and the associated litigation costs (also unhelpful, in light of the fact that *either* party could act to avoid such engagement – the complainant by not filing a complaint and the defendant by settling without litigation);
3. The policy of placing the burden on the party asserting an occurrence rather than on the party asserting a nonoccurrence (an arbitrary dictum that explains nothing because it depends entirely on how the matter is stated – for example, stated affirmatively as a party's "breach of duty" or stated negatively as the party's "failure to exercise ordinary care" or "failure to honor a contractual commitment");
4. The policy of placing the burden on the party to whose case the element is essential (an unhelpful tautology);[10]

[8] Unfortunately, the terminology used to mark this distinction is not uniform, even among common-law countries. What in the United States is usually called the "burden of persuasion" is elsewhere sometimes called the "legal burden"; what in the United States is usually called the "burden of production" is elsewhere sometimes called the "evidential burden." See, e.g., Tapper (1995) at 119–25.

[9] See 2 McCormick (2013) § 337; see also Bolding (1960) at 23–7 and Nance (1994) at 659–69.

[10] The seeming plausibility of this principle arises when and because it is offered not as an explanation or justification for the allocation but as guidance to someone, such as a student

5. The policy of handicapping a substantively disfavored claim, one the very making of which is to be discouraged (sometimes plausible and often overlapping with rationale 8);
6. The policy of placing the burden on the party with presumptively superior access to evidence (seemingly plausible but inconsistent with many conventional allocations because plaintiffs often must prove what defendants did, and with what mental state, and vice versa);
7. The policy of placing the burden on the party whose allegations are *a priori* less likely to be true (also seemingly plausible but ultimately unnecessary unless understood as imposing a judicial evaluation of the strength of evidence on a jury that might be otherwise disposed); and
8. The policy of placing the burden on a party who asserts that an opponent has violated a serious social norm (a useful criterion when disputes involve such allegations, but not in no-fault disputes).

For present purposes, it is unnecessary to develop a unified solution to this question, if indeed any exists.[11] Part of the difficulty in doing so is that there are really two issues being decided with reference to such factors, at least for Anglo-American adversarial trials. The allocation of the burden of persuasion generally determines the allocation of the initial burden of production with respect to the same ultimate facts.[12] This is understandable because inevitably *one* of the parties must start the process, and requiring the other party to present evidence would be wasteful if the party bearing the burden of persuasion cannot establish a prima facie case.[13] From this perspective, policies 6 and 7 make more sense than they otherwise would because they speak more to the allocation of the burden of producing evidence. In fact, allocating the initial burden of production may be the more

or practitioner, who is trying to discern where the burden has been placed, after it has been so placed for other reasons. See, e.g., Park et al. (2011) at 87 n. 23 and accompanying text.

[11] See Wigmore (1981) § 2486 at 291 ("The truth is that there is not and cannot be any one general solvent for all cases. It is merely a question of policy and fairness based on experience in the different situations.").

[12] In the unusual case where the burden of raising an issue is separated from the risk of nonpersuasion on that issue, it can be difficult to say whether the matter is part of the claimant's affirmative case (conditional on its being raised by the opponent) or, rather, a true affirmative defense. Neither characterization seems entirely satisfactory. For example, in a minority of American jurisdictions, the burden to plead (and to present some evidence on) the question of insanity rests on the accused, but the burden to prove sanity then rests on the prosecution. See 1 LaFave (2003) § 8.2.

[13] See Posner (2001) at 1502–4 and Lempert (2001) at 1662–3. Elaborate explorations from a game-theoretic perspective, using simplifying assumptions, can be found in Hay (1997) and Sanchirico (2008).

important of the two tasks. In any event, my focus here is on the question of what it takes to satisfy these burdens once they have been allocated to a party. So the allocation of material facts as between the plaintiff's (or prosecution's) case and affirmative defenses will be taken as given.

The same considerations generally apply in thinking about the burdens placed on the defense – affirmative defenses – as apply in thinking about the burdens for the plaintiff's or prosecution's case. Therefore, to simplify matters, unless otherwise indicated, discussion will proceed in terms of the context of a burden of persuasion placed on the plaintiff or prosecution. Hereafter, when it is unnecessary to distinguish between civil or criminal cases, I will use the generic term *claimant* to refer to either the plaintiff or prosecution.

With these preliminaries out of the way, it is time to articulate the three ideas around which the present discourse is organized. First is the idea of balancing the evidence favoring one side of a case against the evidence favoring the other side. The metaphor of "weighing" the evidence is a powerful one. Thus, in explaining the preponderance-of-the-evidence standard, Pennsylvania's pattern jury instruction advises as follows:

> Think about an ordinary scale pan on each side to hold objects. Imagine using a scale as you deliberate in the jury room. Place all the evidence favorable to the plaintiff in one pan. Place all the evidence favorable to the defendant in the other. If the scales tip, even slightly, to the plaintiff's side, then, you must find for the plaintiff.[14]

Of course, the metaphor can be questioned. For example, it would be more accurate to speak of comparing the degree to which the evidence *as a whole* and the inferences rationally drawn therefrom support one side as compared to the other. Evidence does not always neatly divide into that which favors one side and that which favors the other. In particular, whereas adversarial trials divide evidence into that which is *presented* by one side and that which is *presented* by the other, the evidence presented by one side may, of course, actually favor the other, and it is the net effect of all the evidence that matters.[15]

[14] Pennsylvania Suggested Standard Civil Jury Instructions § 1.42 (3d ed., 2005).

[15] A standard jury instruction for civil cases reads

> In determining whether any fact in issue has been proved by a preponderance of the evidence, unless otherwise instructed you may consider the testimony of all witnesses, regardless of who may have called them, and all exhibits received in evidence, regardless of who may have produced them.

3 O'Malley et al. (2006) § 104.01. See also United States v. Keuylian, 602 F.2d 1033, 1040–41 (2d Cir. 1979) (applying the widely recognized rule, adapted from civil cases, that evidence

Still, the metaphor has undeniable appeal. And it suggests something important, especially when considering cases not governed by the preponderance-of-the-evidence standard.[16] Though a simple scale may not be able to measure it, we know that the physical weight that is involved in the use of ordinary scales can be measured; with appropriate measuring devices, one can say just how much more one object weighs than another, and *a fortiori*, one can measure the ratio of the weights of the objects in the two pans. Does such a strategy carry over to adjudication? By some accounts, the balance of evidence is quantifiable as a ratio of probabilities, for example, "Given the evidence, claimant's case is twice as likely as defendant's." By some accounts, no such quantification is possible: whereas we might say, at least in some contexts, that claimant's case is stronger than defendant's, we cannot quantify *how much* stronger.[17]

In using the phrases "claimant's case" and "defendant's case," I have deliberately abstracted to avoid certain difficulties in defining what it is, precisely, that the claimant or the defendant tries to prove or is required to prove. In particular, one needs to distinguish between an alleged cause of action (or defense) and the specific factual story (or "theory of the case") advanced by a party to show that a cause of action (or defense) is instantiated or that it is not instantiated. This distinction is important in understanding any theory of the burdens of proof, and we will pay attention to it in what follows. And both these ideas must be distinguished from whatever pleading requirements are in place. In civil and criminal cases, most jurisdictions follow pleading requirements that are demanding enough to do somewhat more than just identify the cause of action or affirmative defense relied on but still permissive enough not to require the articulation of a detailed theory of the case.[18]

Quantifiable or not – and abstracting for the moment from exactly what it is that must be compared – this sort of comparative assessment is surely what most lawyers in the common-law tradition regard as the weight of the evidence.[19] Thus, "[i]n civil actions, the burden of persuasion

presented by the defense may cure what otherwise would be a defect in the state's evidentiary case, so a motion for directed verdict by the defense must be decided taking into consideration not only evidence presented by the state but also any evidence presented by the defense that might fill any gap in the state's evidence).

[16] See Walton (2002) at 13–14 (using the image of the "teeter-totter" to illustrate the difference between the criminal and civil standards).

[17] See, e.g., Haack (2014) at 47–77.

[18] See, e.g., Fed. R. Civ. P. 8 & 9 and Fed. R. Crim. P. 7(c)(1).

[19] "[W]eight of the evidence: the persuasiveness of some evidence in comparison with other evidence" declares the standard legal dictionary. Black's (1999) at 1588.

is usually described as a requirement that there must be a *preponderance of the evidence*, or that the *greater weight of the evidence* must exist in favor of the party having the burden of persuasion."[20] Indeed, the two modes of expression are often used interchangeably within the same jurisdiction. For example, one standard jury instruction for federal civil cases reads

> The party who has the burden of proving a fact must prove it by the [(greater weight) or (preponderance)] of the evidence. To prove something by the [(greater weight) or (preponderance)] of the evidence is to prove that it is more likely true than not true.[21]

The committee comments that accompany this instruction indicate that the choice presented in the brackets is a matter of the preference of the individual judge, confirming that the authors see no material difference between the two wordings; to prove by the preponderance of the evidence is just to prove by the greater weight of the evidence.[22] Moreover, the instruction makes clear that what both these ideas mean is proving that a fact is more likely than not, invoking the concept of probability. Several American jurisdictions state this "more probable than not" standard for the typical civil case.[23]

This notion of case weight involves an "apportionment" between the parties. The more the evidence favors the claimant, the less it favors the defendant. What matters to the decision is which of the contending hypotheses is more favored by the evidence and by what margin or to what degree. The margin required to support a verdict depends on the nature of the case. In particular, it is generally thought that a wider margin is required to warrant a conviction in a criminal case than is required to warrant a verdict for the plaintiff in a civil case.[24] I will refer to this sense of

[20] Teply & Whitten (1994) at 828 (emphasis in original).

[21] Manual of Model Civil Jury Instructions for the District Courts of the Eighth Circuit, Instruction No. 3.04 (1999). This model instruction has since been simplified to eliminate the phrase "preponderance of the evidence"; now it equates proving by "the greater weight of the evidence" with proving "more likely true than not true." Manual of Model Civil Jury Instructions for the District Courts of the Eighth Circuit No. 3.04 (2013).

[22] See also Connecticut Civil Jury Instructions 3.2–1 (2011) ("The party who asserts a claim has the burden of proving it by a fair preponderance of the evidence, that is, the better or weightier evidence must establish that, more probably than not, the assertion is true.")

[23] See, e.g., State Bar of Arizona, RAJI (Civil) 4th, Std. 2 (2005) and Judicial Council of California Civil Jury Instructions (CACI) No. 200 (2011). See also 3 O'Malley et al. (2006) § 101.41 (presenting pattern instruction for federal civil cases). On the evolution of the various versions of the usual standard of proof for civil cases, see Leubsdorf (2013).

[24] 2 McCormick (2013) § 341. Continental European countries in the "civil law" tradition often do not explicitly distinguish criminal from noncriminal cases in terms of the standard

weight as the *discriminatory power* of the evidence because it reflects the extent to which the evidence, placed in the context of background information that the trier of fact is entitled to use, discriminates between the two sides of the dispute.

The second idea of concern here is, roughly speaking, the amount of evidence relating to the contending hypotheses that has been developed and taken into consideration. In *very* rough terms, if discriminatory power entails an apportionment, it is plausible to focus on and attempt to measure or otherwise assess that which is being apportioned. Returning to the metaphor of the scales, one can imagine placing a measurable amount X of evidence in the pan favoring the claimant and a measurable amount Y in the pan favoring the defendant. In these terms, discriminatory power compares X to Y by looking, for example, to the ratio of X to Y. But one can also consider combining X and Y, representing the total amount of evidence that is considered by determining, for example, the sum $X + Y$. This idea is also commonly called the "weight" of the evidence by several theorists, following the terminology of John Maynard Keynes.[25] I would follow this simple usage, but for the fact that the term *weight* by itself is so commonly associated by lawyers with what I have here called "discriminatory power." To avoid confusion, this second conception of weight will be referred to as the "Keynesian weight" of evidence.[26]

This notion of weight will be refined in Chapter 3, but even with the relatively crude articulation just provided, an important observation can already be made. Keynesian weight is not comparative in the sense described earlier. It is incoherent to ask how the weight of evidence, in Keynes's sense of the term, differs as between contending hypotheses. This is so because the weight of the evidence with respect to a hypothesis is always the same as the weight of the same evidence with respect to the competing hypothesis. Keynesian weight, however, is comparative in a different sense (one that it

of proof to be applied, a curious fact (at least to the Anglo-American legal mind) that calls for explanation. See, e.g., Clermont & Sherwin (2002).

[25] Keynes (1921) at 71.

[26] In a previous paper I distinguished these ideas by using the term Δ-*weight* for discriminatory power and the term Σ-*weight* for Keynesian weight. See Nance (2008) at 269. The reasons for this nomenclature may already be obvious. The verbal characterizations used here may be easier for readers to keep straight. It comes, however, with its own potential for confusion: what I call the discriminatory power *of the evidence in a particular case* ought not be confused with what forensic scientists call the discriminating power *of a particular test*. The latter measures the relative usefulness of various forensic tests across a group of cases, as distinct from the value of the result of any test in any particular case. See Aitken & Taroni (2004) at 129–41.

shares with discriminatory power): one can, at least sometimes, meaningfully compare Keynesian weight (or, for that matter, discriminatory power) *for different states of the evidence*. To illustrate in terms of our admittedly clumsy metaphor of evidence in the pans of the scales of justice, suppose that one were to double the amount of evidence favorable to the claimant and likewise double the amount of evidence favorable to the defense. Then the ratio of the amount of evidence favoring the claimant to the amount of evidence favoring the defendant would remain unchanged, but the total amount of evidence allocated between them would double. If we use the indicated ratio to measure discriminatory power and the indicated sum to measure Keynesian weight, then discriminatory power is unchanged, whereas Keynesian weight is doubled.

The third idea to be addressed is the *tenacity* with which one holds one's belief that a hypothesis is correct, that is, how difficult it is for one to be persuaded to abandon that belief.[27] The belief in question may be either categorical or partial. (Categorical belief is dichotomous: if p is a proposition of fact, then one believes that p or one does not. Partial belief is gradational: one can believe more or less that p from an extreme of certainty that p is false to an extreme of certainty that p is true.) Especially in the context of a partial belief that is graded in terms of probability, tenacity is sometimes referred to as "resilience." Thus theorists have referred to the resilience of probabilities.[28] As explained later, these ideas are, or at least can be, related to Keynesian weight, but the relationship is complex.

In modern legal scholarship, the question has been posed whether the burden of persuasion requires anything more than, or something completely different from, a comparison of the assessed discriminatory power of the evidence to the standard of proof appropriate to a particular kind of case. For example, if the civil standard is articulated as requiring that a fact be proved more likely than not – as it often is – then the question becomes, what more (if anything) is required for a verdict favoring the burdened party (usually the plaintiff) than a determination by the trier of fact that the odds favor the burdened party, that is, that the probability that the plaintiff's claim is true is greater than the probability that it is not? Most lawyers, certainly most lawyers practicing before common-law courts, would reflexively say that nothing more is required, and there is considerable support for this intuition in recognized authority.[29]

[27] See, e.g., Harman (1986) at 22 and Owens (2000) at 144.
[28] See, e.g., Logue (1995) at 78–95 and Stein (2005) at 82.
[29] See, e.g., 3 O'Malley et al. (2006) § 104.01 (providing a standard jury instruction): "'Establish by a preponderance of the evidence' means to prove that something is more

Nevertheless, there have been dissenting voices. Occasional judicial opinions seem to balk, although it is often difficult to discern in such cases whether the court is simply opposed to probabilistic judgments or is rather concerned that, in the context of the particular case, a probabilistic judgment cannot by itself encapsulate the totality of the standard of proof, or rather that the probability threshold simply has not been set high enough.[30] Picking up on these qualms, several scholars have maintained that an assessment of discriminatory power must be *supplemented with* an assessment of something else, that satisfying the burden of persuasion requires both attainment of the necessary margin of discriminatory power in the evidence and attainment of a certain degree of (what I have called) Keynesian weight.[31] For at least one important scholar, the argument goes so far as to suggest that an assessment in terms of Keynesian weight should be *substituted for* any assessment in terms of conventional probabilities.[32]

Similarly, there are those who have maintained that the assessment of discriminatory power, when conceived in probabilistic terms, must be supplemented with an assessment of, and test for, the resilience of those probabilities.[33] For the most part, those who make this argument do not distinguish it from, and appear to believe that it is the same as, the preceding argument, namely, that the assessment of discriminatory power should be coupled with a separate assessment of Keynesian weight. In some cases, once again, this argument goes so far as to dispense (seemingly) with discriminatory power entirely so that the burden of persuasion consists of the

likely so than not so"; 2 McCormick (2013) § 339 at 661 ("The most acceptable meaning to be given the expression, proof by a preponderance, seems to be proof which leads the jury to find that the existence of the contested fact is more probable than its nonexistence. Thus the preponderance of the evidence becomes the trier's belief in the preponderance of the probability.")

[30] See, e.g., Sargent v. Massachusetts Accident Co., 29 N.E.2d 825 (Mass. 1940) and Lampe v. Franklin American Trust Co., 96 S.W. 2d 710 (Mo. 1936). Not all the cases are older ones. See, e.g., Spencer v. Baxter Int'l, Inc., 163 F. Supp. 2d 74, 80 n.7 (D. Mass. 2001) (noting Massachusetts law). See also Sienkiewicz v. Greif (UK), Ltd. [2011] UKSC 10 (9 March 2011) (expressing reservations about overtly and exclusively statistical proof that meets such a quantitative standard).

[31] See, e.g., Brilmayer (1986) at 681–5; Davidson & Pargetter (1987) at 183–5; Friedman (1996) at 1819–20; and Stein (2005) at 120.

[32] Cohen, J. (1977). By "conventional probabilities" here, I refer to probabilities that are numbers conforming to the familiar Kolmogorov axioms, what Cohen refers to as "Pascalian" or "mathematical" probabilities. Cohen constructed a different kind of probability, which he called "Baconian" or "inductive," which follows a different logic and which Cohen thought cohered better with standards of proof in adjudication.

[33] Davidson & Pargetter (1987) at 183–5 and Stein (2005) at 47–8, 120. See also Cohen, N. (1985).