Likelihood Ratio Tests and Legal Decision Rules

Louis Kaplow[*]

Abstract

Various legal decision-making criteria can be formulated as likelihood ratio tests, wherein liability, prohibition, or other outcomes are associated with evidence strength exceeding a posited threshold. Stating rules in this manner clarifies their nature, facilitates the comparison of conventional and optimal rules as well as the identification of differences between rules across contexts, and provides further illumination in instances in which a decision standard is not truly a likelihood ratio test.

JEL Classes: D81, K14, K41, K42

---

1. Introduction

The likelihood ratio is a statistical term that refers to the strength of information in supporting one hypothesis over another, often its negation. A likelihood ratio test, deriving from Neyman and Pearson (1933), accepts a designated hypothesis when the likelihood ratio exceeds a critical value and rejects it otherwise. Likelihood ratios have received only modest attention in the legal literature, mainly in the context of expert presentations of scientific evidence such as DNA tests, some discussions of statistical evidence more generally, and expositions of the legal notions of relevance and probative force (e.g., DeKay 1996, Friedman 1986, Johnston 1987, Kaplan 1968, Koehler 1996, Lempert 1977, and Martin and Schum 1987). By contrast, likelihood ratios are much more familiar in medical assessments and in a range of other decision settings in government and business.

This article's thesis is that the likelihood ratio concept and likelihood ratio tests in particular can usefully illuminate a broad range of legal standards, from conventional burden of proof notions to optimal formulations in various contexts to criteria for preliminary dispositions to protocols for agency screening, assessments of probable cause, and investigative prioritization. By stating each such decision rule in the form of a likelihood ratio test, we can better understand what it entails. In addition, this common presentation facilitates comparisons of tests. Finally, any difficulties in articulating legal tests in this manner itself reveals important features.

To begin, what is a likelihood ratio? Throughout this article, we will consider scenarios in which we wish to determine whether a case, investigation, or other legal setting involves a truly harmful act, denoted by H, or a benign (or beneficial) one, denoted by B. At the decision point in question, the totality of information—that is, the available evidence—is designated by e (which can be understood as a vector, summarizing multiple items of information). Finally, $P(e|H)$ indicates the likelihood (frequency) of observing e when the act is H (that is, conditional on H) and $P(e|B)$ the likelihood of observing e when the act is B.

The likelihood ratio for the evidence e, $LR(e)$, is given by $P(e|H) / P(e|B)$. A high likelihood ratio indicates that e is more frequently generated when the true act is H than when it is B; conversely for a low ratio. A ratio of one means that e is equally likely to be generated by either act. The likelihood ratio is a measure of how strongly the evidence e supports the hypothesis that the act is H rather than B.

A likelihood ratio test is one under which the decision associated with H being true is accepted if and only if $LR(e) > LR^*$, where LR^* is some critical value of the likelihood ratio. One can immediately see the intuitive appeal of such a test. Suppose that we choose H when the likelihood ratio is $LR(e)$ for some evidence e. Under a likelihood ratio test, we then necessarily choose H for any higher likelihood ratio, that is, whenever the evidence is stronger than this e. Similarly, if we reject H, adopting the result associated with B, for some other likelihood ratio associated with some other evidence, then we necessarily reach the same negative conclusion

when the likelihood ratio is even lower, that is, when the evidence is even weaker.[1] Any test with this property, it should be apparent, can be stated as a likelihood ratio test. (See Neyman and Pearson 1933, Karlin and Rubin 1956, and Milgrom 1981.) Therefore, the prior suggestion that a broad array of tests used in legal settings (and in many others, for that matter) can usefully be stated in the form of a likelihood ratio test is unsurprising.

This article explores the usefulness of stating legal tests in likelihood ratio form. Section 2 begins by elaborating on evidence, likelihood ratios, and likelihood ratio tests. Importantly, it distinguishes likelihood ratios from Bayesian posterior probabilities, notably, P(H|e) and P(B|e): the probability that H is true given the evidence e and the probability of B given e. In legal settings, we commonly ask—for example, in reaching a decision at the conclusion of a trial—how likely it is that H is true given the evidence e. Note that the previously defined likelihood ratio had the order of the conditional probabilities reversed, examining instead P(e|H) and P(e|B). (This contrast is so central to the understanding of legal decision rules that it is the subject of the concluding section of this article.)

Section 3 uses likelihood ratios, and this contrast with Bayesian posterior probabilities, to examine formulations of the burden of proof at trial: how strong evidence must be for liability. This section considers both conventional notions and optimal proof burdens, in settings involving the regulation of future conduct and in those concerned with the provision of incentives for ex ante behavior—recasting Kaplow (2011a, 2011b, 2012). Formulating all pertinent rules as likelihood ratio tests highlights both the similarities and differences between rules and across contexts.[2] One important difference brought into sharp focus is that conventional burden of proof rules ignore consequences for social welfare that are central in determining the critical likelihood ratio under optimal rules. Another is that the comparison of optimal likelihood ratio tests across contexts makes transparent divergences in pertinent welfare consequences.

Section 4 draws on Kaplow (2013a, 2013b) to consider preliminary or interim decisions in a setting of multistage legal decision making—which may be formal (motions to dismiss, summary judgment, grand jury indictment, probable cause determination) or informal (agency screening, investigative triage). Contrasts with the likelihood ratio tests for the burden of proof at trial are instructive. Moreover, in this context, there are important respects in which the likelihood ratio formulation of the decision rule is not a true likelihood ratio test: Although a higher likelihood ratio, all else equal, still favors the result associated with H, the test is not a proper likelihood ratio test because the critical value of the likelihood ratio is not fixed but itself depends on the evidence e. As a consequence, a case with a high likelihood ratio could warrant result B because the critical likelihood ratio is even higher, whereas another case with a lower likelihood ratio could warrant result H because the critical ratio is even lower.

---

[1]Suppose that the test is not of the form that H is accepted if and only if LR(e) > LR*, for some LR*. Then there must exist two different values of e, call them e_1 and e_2, such that we choose H when the evidence is e_1 and B when the evidence is e_2 even though LR(e_1) < LR(e_2).

[2]By contrast, as elaborated in section 5, one cannot readily state key rules in terms of threshold Bayesian posterior probabilities because a key element of those probabilities (the prior probabilities) do not occupy a central (or any) role in optimal decision criteria in settings involving incentives for ex ante behavior.

The section also discusses another, quite different sense in which common understandings imply that the legal decision rule for preliminary decisions is not, cannot, and should not be formulated as a likelihood ratio test: the view that the object of interest is not a likelihood ratio—which is to say that the rule does not depend on P(e|H) and P(e|B) but rather on something qualitatively different. As will be explained, an implication of this view is that it need not be true that higher likelihoods favors the result associated with act H, and for reasons that seem to be divorced from consequences for social welfare.

Section 5 offers a concluding discussion. As mentioned, it emphasizes the similarities and differences between formulating legal decision rules as likelihood ratio tests versus Bayesian posterior probabilities, the latter providing the basis for conventional (but not optimal) decision criteria and being more familiar in our thinking about decision problems more generally.

2. Basic Concepts

2.1 Likelihood Ratios and Likelihood Ratio Tests

To begin, it is useful to introduce the setting, notation, and basic concepts more thoroughly than was done in the introduction. As mentioned, the discussion throughout this article will consider settings that involve two possible acts, a harmful one (H) and a benign or beneficial one (B). At a criminal trial, H would be associated with the true state of affairs being one in which the accused individual committed the alleged act, and B with others (perhaps that the act never happened, perhaps that the accused was not the individual who committed the act). In a civil case, act H might be a contract breach or tort. In a zoning hearing, H would indicate that the applicant's prospective activity has undesirable traits that make a denial of permission appropriate. Analogues can readily be imagined for other settings, including investigations or preliminary proceedings.

Let P(H) and P(B) denote the initial unconditional (prior) probabilities of acts H and B, respectively. These are the probabilities of each act in the relevant population and, specifically, before consideration of the evidence in a case.[3] They are also referred to as base rates.

The notation employing a vertical bar (|), also introduced in the introduction, indicates a conditional probability: the probability of the first-listed variable, conditional on the value (realization) of the second-listed variable. Hence, P(e|H)—read as "the probability of e given H"—denotes the probability that we observe evidence e when the true act is H. Likewise for P(e|B). It is helpful to think of each of these particular conditional probabilities as the ex ante likelihood that the designated act generates the evidence e. As remarked in the introduction, the variable e should be interpreted broadly (as a vector, where appropriate) as incorporating all pertinent information. Additionally, it is usefully understood as the raw information—for

---

[3]In some settings, they might usefully be understood as probabilities based on evidence through some prior stage. In addition, in many of our settings, since act B is often taken to encompass all possibilities other than H, the two probabilities will sum to one, and we could, for example, express P(B) as $1 - $P($H$). However, for reasons of generality and clarity, this approach will not be taken.

example, the mass of documents and witness statements—and not the inferences or conclusions drawn therefrom. Those resulting (Bayesian posterior) probabilities will be considered in subsection 2.2.

To make more explicit these conditional probabilities and the formulation of the likelihood ratio, consider the situation in which there is only a single piece of evidence, where e denotes the strength of that evidence. That is, a higher e indicates stronger evidence. (The careful reader may note that we are a bit ahead of ourselves, since in a moment we will use the likelihood ratio, LR(e), to *define* evidence strength. Nevertheless, it is helpful to begin with this thought in mind.)

Intuitively, it is natural to understand a higher e as constituting stronger evidence of H being true—specifically, in contrast to B being true—as meaning that higher values of e arise more frequently when H rather than B occurs.[4] A concrete depiction of this portrayal is presented in Figure 1.

Figure 1: Probability Distributions for Strength of Evidence
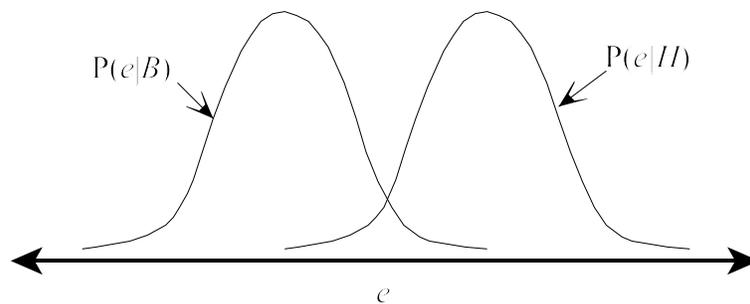
This Figure is meant merely to be suggestive. It is hand-drawn as two (roughly) normal distributions with the same variance and different means: The curve for the distribution for P($e|H$) is to the right of that for P($e|B$).[5] As we can see, higher values of e are more likely when H rather than B is the truth, but, even so, both distributions overlap such that any value of e is possible in either case.

We can use Figure 1 to illuminate the concept of the likelihood ratio: LR(e) = P($e|H$) / P($e|B$). We can see that this ratio tends to be higher toward the right of the Figure. This is most apparent for values of e between the two peaks. As we move toward the right, P($e|B$), the denominator of the likelihood ratio, is falling and P($e|H$), the numerator of the ratio, is rising.

---

[4]As is conventional, this article throughout employs language such as more likely, more probable, or more frequent even when we are considering continuous density functions (such as those in Figure 1), wherein the probability of any specific value of e is zero. As will be apparent from the discussion to follow of the likelihood ratio, all that matters for present purposes is the ratio of the density functions at any given e, and this ratio is a well defined concept (as long as the denominator is nonzero, which we are supposing to be true).

[5]This special case is typically used in presenting basic signal detection theory, the first formally developed system (originally created for radar detection) of this sort. See Luce (1963) and Macmillan and Creelman (2005).

Hence, LR(*e*) is clearly rising. The property that the likelihood ratio rises with *e* may, however, not be so obvious at the extreme left or extreme right of the Figure, as one moves into the tails of the distributions. Nevertheless, for many standard distributions (including the normal distributions employed here), this will be true.[6]

Let us now consider explicitly the likelihood ratio, LR(*e*) = P(*e*|*H*) / P(*e*|*B*). We can derive this ratio by taking the ratio of the heights of the two probability distributions in Figure 1. Again using a hand-drawn figure that takes some license, we obtain Figure 2.

Figure 2: Likelihood Ratio as Function of Strength of Evidence
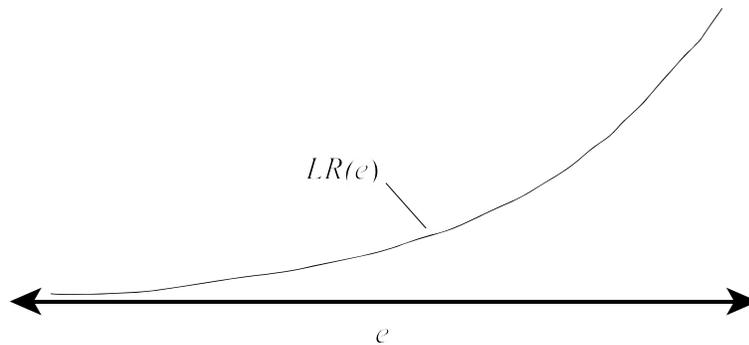
As already explained, the likelihood ratio corresponding to the curves drawn in Figure 1 rises as *e* increases. Although no vertical axis is drawn, we know that, when *e* takes the value where the two curves cross in Figure 1, LR(*e*) = 1. To the far left of Figure 2, LR(*e*) is very close to zero, whereas toward the far right, the value is relatively high, and we can imagine that as *e* increases ever further, LR(*e*) approaches infinity. Note that LR(*e*) = 0 represents a value of *e* that has a probability of zero when the truth is *H*, which is to say that *H* is ruled out entirely; and, when LR(*e*) approaches infinity (think of the case in which P(*e*|*B*) = 0), *B* becomes essentially impossible, making *H* certain.[7]

As also mentioned in the introduction, a likelihood ratio test has the form that the decision associated with *H* being true (often abbreviated here in various forms, such as result or decision *H*) is taken if and only if LR(*e*) > LR*, for some critical value of the likelihood ratio,

---

[6]This trait of distributions is referred to as the (strict) monotone likelihood ratio property (MLRP), and it is assumed to hold in a number of applications in statistics and economics. See Karlin and Rubin (1956) and Milgrom (1981). If the property did not hold for some range of *e*, then a higher value of *e* in that interval would not indicate "stronger" evidence in the relevant sense, where here, the relevant sense is taken to be that indicated by the likelihood ratio. The reasons for this definition of evidence strength, which is in accord with intuition, will become apparent as we proceed.

[7]In most situations, these are conceptual possibilities; in real settings, it is hard to imagine evidence that literally renders one result certain and another impossible, given all manner of sources or error and the potential for deception.

denoted LR*.[8] Note that, in our present setting, this implies that there will be a critical value of evidence strength, which we can denote e^*, such that we take decision H if and only if $e > e^*$. Recall that, in introducing the variable e and in presenting Figure 1, it was stated that e is an indicator of evidence strength, whereas in our more precise derivation, we have really defined evidence strength as $LR(e)$—on which more in subsection 2.2. Given that we are imagining situations such as those depicted in Figure 2, in which $LR(e)$ rises monotonically with e, these two descriptions of a likelihood ratio test are equivalent.

Finally, before proceeding, it is useful to emphasize that our measure of evidence strength and, relatedly, any likelihood ratio test each depend on the likelihood *ratio*—not on the absolute magnitude of any probability, notably, of $P(e|H)$. That is, a high likelihood ratio (strong evidence) is a relative concept. It indicates not "how likely is the evidence e given that we truly have a harmful act H" but rather "how likely is this possibility *relative to* seeing the same evidence e given that the truth is B."

This point can be seen sharply by considering a case in which we observe e and $P(e|H)$ is small, say 1 in 100. One might be tempted to think that the likelihood we are in situation H is quite low, but this conclusion is invalid, for our information is incomplete. Among other things (for a full statement, see subsection 2.2), we also need to know the value of $P(e|B)$, and let us suppose it is 1 in 1000. In that case, our likelihood ratio is 10 (0.01 / 0.001), which is high. This tells us that when the truth is H (or the act under consideration is of type H), evidence e is generated ten times as often as when the truth is B (or the act under consideration is of type B).

Similarly, we could also observe an e where $P(e|H)$ is much larger, say 1 in 5. But it might be that $P(e|B)$ is even greater, say 4 in 5. Then our likelihood ratio is only 0.25 (0.2 / 0.8)—only 1 in 4, which is to say, one fortieth as high as in our prior example. Put another way, even though e is generated twenty times as often when the truth is H than was so in our first example, the likelihood ratio is only one fortieth as high. Hence, it is important as a practical matter to focus on the ratio and not on the absolute magnitudes. The significance of this message is reinforced by the discussion of our next subject, posterior probabilities.

2.2 Bayesian Posterior Probabilities Compared

First, let us introduce some terminology. Probabilities indicate how likely an event is to happen—or, retrospectively, how likely it is that it happened. As is entirely familiar, probabilities range from zero to one, zero corresponding to the event being impossible and one to it being certain. In assessing how initial probability estimates are transformed (updated) in light of additional information (evidence) to form revised probabilities—the subject of Bayes' Theorem (or Rule)—it is useful to introduce the additional language of prior probabilities and posterior probabilities. For example, if historical data indicate that, in a particular locale, it snows 25% of the time on January 1, our prior probability of snow on the upcoming January 1,

---

[8]Throughout, I employ the convention that ties are associated with decision B, as is common in legal and other settings. When dealing with continuous density functions, ties have probability zero, although we can imagine reasons that actual decision rules regarding ties may have nontrivial effects. See Kaplow (2012, p. 758 n. 34).

P(SnowJanuary1), would be 0.25. Now, suppose further that our meteorologist studies weather data available on the immediately preceding December 25 and revises the estimate to 50%; then, the posterior probability P(SnowJanuary1|NewData) is 0.50. In a moment, we will examine a version of Bayes' Theorem that explains how the new data is processed to move from our prior probability estimate to our posterior.

Another useful term, familiar from settings such as gambling (for example, horse races), is the odds ratio, which is given by the probability of our event divided by the probability of the contrary event. For example, the toss of a fair coin has an odds ratio for coming up heads[9] of 50–50, or 1 to 1, or just 1 when stated as a fraction. Whereas probabilities range from zero to one, odds ratios range from zero (impossible events) to infinity (certain events).

Taking our prior probability of snow of 0.25, the prior odds ratio is 1 to 3, or 1/3: There is initially a 25% chance of snow and a 75% chance of no snow, so the ratio of 25% (snow) to 75% (no snow) is the so-called odds (or odds ratio) of snow (to no snow). Likewise, our above posterior probability (after data becomes available on December 25) of 50% means that snow on January 1 is now a 50–50 proposition (just like the coin toss), which is an odds ratio of 1 to 1, or just 1. As a final remark on our example, foreshadowing what comes next, observe that our odds ratio rose from 1/3 to 1, that is, by a factor of 3. As we now shall see, this occurs when the likelihood ratio associated with the additional weather data equals 3—which convenient feature explains why the odds ratio is introduced here and the odds-ratio form of Bayes' Theorem is presented next.[10]

Bayes' Theorem allows us to translate the prior probability of an event into a posterior probability, in light of additional information. Stated in odds ratio form, Bayes' Theorem holds that the prior odds ratio times the likelihood ratio associated with the additional evidence equals the posterior odds ratio. Using our earlier notation, we have:

$$[P(H) / P(B)] \times LR(e) = P(H|e) / P(B|e).$$

Or, if we wish to state the likelihood ratio in terms of its composite conditional probabilities, we can equivalently state:

$$[P(H) / P(B)] \times [P(e|H) / P(e|B)] = P(H|e) / P(B|e).$$

We begin with the initial ratio of probabilities of our two events (the odds ratio of the priors) and multiply by the ratio of the conditional probabilities of seeing the evidence *e* given each of our two events (the likelihood ratio), and the product gives us the ratio of the posterior probabilities

---

[9]Or tails.

[10]Bayes' Theorem has been featured in some legal writing, particularly pertaining to debates about the use of statistical evidence in court and in probability-based discussions of relevance and probative force. See Kaplan (1968), Cullison (1969), Finkelstein and Fairley (1970), Tribe (1971), Lempert (1977), Friedman (1986), Johnston (1987), Martin and Schum (1987), DeKay (1996), and Koehler (1996) . Use of the likelihood ratio and likelihood ratio tests (apart from discussions of scientific evidence, like DNA tests, and relevance or probative force) has been less common and, even then, they do not usually play a central role.

of our two events (the odds ratio of the posteriors). This formulation of Bayes' Theorem is a straightforward transformation of the more traditional presentation, as shown in the margin.[11]

We can now reflect on what these expressions, intuitively, are telling us: We start with the prior odds—a 25% chance of snow in our example, which gave an odds ratio of 1 to 3—and multiply by the likelihood ratio implied by the evidence at hand—asserted to be 3 in our example—to obtain the posterior odds of 50–50, or 1 to 1. Examining either expression, we can see that a likelihood ratio above one raises the prior odds and a likelihood ratio below one lowers them, which is in accord with the explanation of the likelihood ratio in subsection 2.1. When the likelihood ratio is, say, 3, this means that action H generates our evidence three times as often as action B generates the evidence. Accordingly, when we observe this evidence, the prior odds triple. If, instead, the likelihood ratio were, say, 0.5 (one half), then the evidence would cut the prior odds in half; in our example, they would fall from 1 to 3 to a ratio of 1 to 6.[12]

Using these expressions, we can see that there is an important connection between the likelihood ratio associated with evidence, LR(e), and the Bayesian posterior probabilities, which are much more familiar in the legal setting, particularly with regard to standards of proof. As we will see in the remainder of this article, however, extracting the likelihood ratio and placing it front and center in our minds is extremely important, for Bayesian posterior probabilities almost never tell us what we really need to know if our purpose is to advance social welfare: Sometimes they tell us only part of the story (omitting a great deal), and sometimes they tell us none of the story (omitting everything of relevance). In contrast, the likelihood ratio is an important piece of the picture in all of the settings examined here.

---

[11]The traditional statement of Bayes' Theorem, which is in probability form, yields the following two equations for the two separate posterior probabilities:

$$P(H|e) = [P(e|H) \times P(H)] / P(e),$$

$$P(B|e) = [P(e|B) \times P(B)] / P(e).$$

Now, we can divide the left side of the top expression by the left side of the bottom expression, and equate this to the right side of the top expression divided by the right side of the bottom. (We are starting with the first equation and dividing each side by the same values; the second equation tells us that our two divisors, although expressed differently, are equal.) The resulting left side here is identical to the right side in the text, and the resulting right side here is identical to the left side in the text (noting that the P(e)'s cancel).

[12]The reader may note that translating back and forth between odds and probabilities, although a very simple mathematical operation, can be a bit tricky if we are not careful. For example, these prior odds of 1 to 3 correspond to a prior probability of 1/4 or 25%, and the posterior odds of 1 to 6 correspond to a posterior probability of 1/7, or about 14%. When this translation is made, we can see that the likelihood ratio of 1/2 does not cut the prior *probability* in half (14% is somewhat more than half of 25%) even though the prior *odds* are cut precisely in half. (The translation from the likelihood ratio works correctly when the likelihood ratio is one, for in that case both the posterior odds and the posterior probability equal the prior odds and prior probability, respectively.) Because odds ratios interact with the likelihood ratio in this simple manner, many analysts find the odds ratio formulation of Bayes' Theorem more convenient, and this is why the approach is adopted here. As the preceding footnote explains, the odds ratio formulation of Bayes' Theorem follows in a straightforward manner from the formulation in terms of probabilities.

3. Burden of Proof

3.1 Conventional Formulations: The Preponderance Rule and Beyond

We can use the terminology, notation, and above expressions to define and illuminate conventional formulations of the burden of proof at trial, that is, legal rules indicating the threshold probability of H that must be surpassed to assign liability.[13] This exposition is useful not only for understanding proof burdens in final adjudication but also for examining many other legal tests since they similarly are formulated in terms of—or bear family resemblance to—threshold likelihoods understood in terms of Bayesian posterior probabilities.

Begin with the preponderance of the evidence rule, taken to require that H be more likely than B, or, equivalently, that H be more likely than not (H and B are here understood to be exhaustive of the possibilities), that the odds of H exceed 50–50, or that the odds ratio associated with H exceeds 1. At this point, it is helpful to restate the odds ratio version of Bayes' Theorem:

$$[P(H) / P(B)] \times [P(e|H) / P(e|B)] = P(H|e) / P(B|e).$$

On the right side of this expression we have the Bayesian posterior odds ratio, which the preponderance rule demands to exceed one. That is, it requires $P(H|e) / P(B|e) > 1$. Therefore, we can state the preponderance rule as:

$$[P(H) / P(B)] \times [P(e|H) / P(e|B)] > 1.$$

Taking this expression, we can multiply both sides by the inverse of the ratio of the prior probabilities, $P(B) / P(H)$, which gives us the following two equivalent expressions for the preponderance rule in likelihood ratio form:

$$P(e|H) / P(e|B) > P(B) / P(H), \text{ or}$$

$$\text{LR}(e) > P(B) / P(H).$$

Let us now interpret this statement: The likelihood ratio associated with the evidence e must exceed the prior odds that the act is of type B rather than H. To understand this requirement further, consider three possibilities. First, suppose that the prior odds ratio (the value on the right side of these expressions) is one. This means that, before considering the evidence, it is a 50–50 proposition whether we have an act of type B or type H. Now, if the likelihood ratio for the evidence exceeds this ratio—if it is greater than one—the test is satisfied. This result is consistent with what we learned in subsection 2.2 about the likelihood ratio: When it exceeds one, the prior probability is increased because this likelihood ratio is more associated with H than with B. Since we started at 50% and we now have a higher value, the preponderance rule is satisfied. (Similarly, if $\text{LR}(e) < 1$, the prior odds fall and the rule is not

---

[13]Although controversy remains, it has become fairly conventional in recent times to formulate burdens of proof in probabilistic terms. See Kaplow (2012, p. 741 n. 3, pp. 779–80 n. 78) for references.

met.)

    Now suppose that the prior odds ratio on the right side is higher, say 2 to 1. (Recall that this means that the *B* is twice as likely as *H*.) The formula tells us that, to meet the preponderance rule, we require stronger evidence, with a likelihood ratio exceeding 2, for that would be necessary to offset the initial deficit and bring the posterior odds above 50–50. Finally, if the prior odds ratio was only 1 to 2, this means that *H* is twice as likely as *B* to begin with, so our posterior odds will exceed 50–50 as long as the likelihood ratio exceeds 1/2. Note that even a likelihood ratio below one may suffice because, as long as the prior odds are not pulled down too far, they will still exceed 50–50, or a probability of 50%. (Consider situations governed by the doctrine of res ipsa loquitur: Given just the initial description of the situation, the likelihood of *H* is fairly high, so unless the defendant presents evidence that pulls down the odds a good deal, below the proof requirement of 50%, the plaintiff will prevail.)

    Finally, it is useful and easy to extend the foregoing analysis to cover any evidence threshold formulated in terms of Bayesian posterior probabilities. Consider, for example, a requirement of clear and convincing evidence, and suppose further, for concreteness, that this requires that the probability of *H* exceeds 75%. This demand implies a maximum probability of *B* of 25% and therefore an odds ratio of at least 3 to 1, or a fraction exceeding 3. Therefore, this rule can be stated as:

$$[P(H) / P(B)] \times [P(e|H) / P(e|B)] > 3.$$

From this, it follows (repeating our analysis for the preponderance rule) that this rule, in likelihood ratio form, is:

$$LR(e) > 3 \times P(B) / P(H).$$

That is, for any given prior odds, the requisite likelihood ratio is three times as high as under the preponderance rule.[14]

    Suppose instead that we wished to implement a requirement of proof beyond a reasonable doubt and that this required the posterior probability of *H* to exceed, say, 90% (or perhaps 95%). Then the minimum odds ratio would be 9 to 1 (90% to 10%) [or 19 to 1 (95% to 5%)], and we could accordingly insert this higher factor on the right side. More generally, any posterior probability threshold means that the posterior odds must equal the ratio of that threshold probability to one minus that probability, and we can insert a multiplier of *m*, indicating that ratio, on the right side, as follows:

$$LR(e) > m \times P(B) / P(H).$$

---

[14]It is worth recalling the difference between raw probabilities and odds ratios. It may seem counterintuitive the threshold likelihood ratio is three times greater when the threshold probability is 75% rather than 50%: One might have thought that since 75 is only 50% greater than 50, the threshold likelihood ratio would only rise by 50%. However, the odds associated with 75% are 3 to 1 (75% to 25%) rather than 1 to 1 (50% to 50%).

In sum, any such decision rule can be expressed in this simple likelihood ratio form, and all such rules are qualitatively similar: They all depend on the same information—on prior probabilities (prior odds) and on the likelihood ratio associated with the evidence, LR(*e*). They differ only in the multiplicative factor *m*.

3.2 Optimal Threshold: Future Conduct

We now know what a likelihood ratio is, the role it plays in Bayes' Theorem, and, in light of subsection 3.1, how it relates to the preponderance rule and other rules that require a minimum ex post probability that we have *H* rather than *B*. Nothing, however, has yet been said about how one versus another decision rule relates to social welfare and, in particular, what decision rule maximizes welfare.

This subsection starts with a simple case, one that is *not* the central case for most writing on the economics of the legal system or in analyses of most areas of substantive law. Typically, the focus is on the law as providing incentives, mainly, the deterrence of harmful activity, a subject that is taken up in subsection 3.3. Here we begin with the analytically more straightforward setting in which the (only) effect of a legal decision is in directly governing future conduct.[15] This context is analogous to medical decision making, where the problem is to choose a course of treatment in light of available information. Pertinent legal domains include licensing (of drivers, professionals, or nuclear power plants, perhaps), zoning, approving drugs, and permitting mergers. Such is the realm of conventional cost–benefit analysis, specifically, for decisions involving uncertainty over outcomes.

The central idea is that the critical likelihood ratio is not given purely by the ratio of prior probabilities but rather by the product of this ratio and the ratio of the consequences associated with each outcome. In medical decision making, this point is familiar. Should an individual have surgery if, in light of all the evidence, there is a 30% likelihood that the pertinent ailment is present? Obviously we cannot answer this question without knowing the benefit of surgery when it is truly appropriate and the cost when it is not. The same logic is applicable in legal settings that involve the regulation of future conduct.

Let Gain denote the benefit of regulation when the act is actually of the *H* type. Gain may be the social benefit from prohibiting a harmful activity, often the avoided harm net of any benefit that the activity would have produced. Similarly, Loss refers to the social cost of mistakenly prohibiting the *B* type of act. Loss would typically be the forgone benefit of an act that is not truly harmful. (For convenience, each of these are measured by reference to the baseline in which the activity is permitted.)

The optimal decision is to prohibit an act if and only if the probability that the act is of type *H*, in light of all available evidence (i.e., the Bayesian posterior probability), times the Gain

---

[15]Often, actual settings will involve mixed cases in which both ex ante incentives and the regulation of ex post conduct are relevant. For example, drug approval not only affects which developed drugs will be marketed but also the incentives for research and the development of different types of drugs. In such instances, the present analysis would have to be combined with that in subsection 3.3 to derive the optimal rule.

from prohibiting an act of type *H* exceeds the probability that the act is of type *B*, given the evidence, times the Loss from prohibiting an act of type *B*. That is,

$$P(H|e) \times \text{Gain} > P(B|e) \times \text{Loss}.$$

Therefore, instead of the preponderance of the evidence rule or others of that class, our optimal rule is:

$$P(H|e) / P(B|e) > \text{Loss} / \text{Gain}.$$

Next, we can substitute on the left side using Bayes' rule (in odds-ratio form), just as we did when deriving the likelihood ratio version of the preponderance rule, to obtain:

$$[P(H) / P(B)] \times [P(e|H) / P(e|B)] > \text{Loss} / \text{Gain}.$$

Finally, we can rearrange this expression as we did before and then substitute using the definition of the likelihood ratio to produce the result:

$$P(e|H) / P(e|B) > [P(B) / P(H)] \times [\text{Loss} / \text{Gain}], \text{ or}$$

$$\text{LR}(e) > [P(B) / P(H)] \times [\text{Loss} / \text{Gain}].$$

This final representation of the optimal rule for determining whether to prohibit an act that has future consequences of the sort described is in accord with the introductory suggestion: In addition to knowing the ratio of the prior probabilities (the relative initial likelihoods), we also need to know the ratio of the consequences for social welfare associated with prohibition for each type of act. Another way to restate this final expression that may prove helpful is to combine the numerators and to combine the denominators on the right side:

$$\text{LR}(e) > [P(B) \times \text{Loss}] / [P(H) \times \text{Gain}].$$

This equation tells us that the likelihood ratio associated with the evidence needs to exceed not the ratio of the prior probabilities of the two types of acts but the ratio of the expected consequences associated with the two types of acts.

The difference between the optimal rule and, say, the preponderance rule can be striking—which should not be surprising since the latter ignores consequences. Suppose, for example, that the Gain is very large relative to the Loss. With a medical treatment decision, we can imagine that the treatment is lifesaving if one had the condition *H* but that the costs (in expenditures, adverse side effects, whatever) are modest when the true condition is *B*. In that event, treatment is indicated even when the likelihood ratio—the strength of the evidence indicating *H* rather than *B*—is quite low. The prior odds of no problem (*B*) rather than a problem (*H*), that is $P(B) / P(H)$, might be 4 to 1, and the likelihood ratio associated with the test results, LR(*e*) might be only 1 to 1 (that is, a totally uninformative test), but if the ratio of the Loss to the Gain is, say, 1 to 100, treatment is powerfully indicated ($1 > 4 \times 0.01 = 0.04$). And that is precisely what the stated rule tells us. Turning to legal regulation, it may be optimal to

- 12 -

deny an operating license for a nuclear power plant if the available information indicates even a small probability of a disaster.

Similarly, if the Gain is tiny relative to the Loss, then we do not want to prohibit an activity even if the likelihood ratio indicating that it is type H is fairly high. If the prior odds are 1 to 4 (the activity is initially seen as 80% likely to be of the H type) but the ratio of the Loss to the gain is instead 100 to 1, then the value on the right side of our decision rule is 25 (0.25×100). Accordingly, even powerful evidence of H—say, LR(e) = 10—would fall significantly short of that necessary to justify prohibition.

How does this simple lesson relate to legal practice? Interestingly, despite the impression—reinforced by academic writing—that the preponderance of the evidence rule is overwhelmingly dominant outside the criminal context, it appears that this depiction is apt only with regard to ordinary civil *litigation*, usually about whether prior actions were improper, the setting considered next, in subsection 3.3. Most legal decisions that involve the regulation of future conduct, such as those described at the outset of this subsection, are made by regulatory agencies that, one presumes (and, in light of the present analysis, hopes) are not governed by the preponderance rule. The decisions to allow versus prohibit new drugs, the operation of nuclear power plants, and countless other actions presumably do not depend solely on whether the probability of the pertinent trait exceeds 50%, or some other standard, target probability. Rather, they ordinarily depend as well on the consequences with regard to different types of activities.

The preponderance of the evidence rule, it was stated, ignores consequences, and this feature is transparent when both the preponderance rule and the optimal rule are stated in a common form, as likelihood ratio tests. The optimal likelihood ratio test takes as its critical value not the ratio of the prior probabilities but rather that ratio multiplied by the ratio of the Loss to the Gain. Put another way, the preponderance rule applied in the present setting treats situations as if the Gain and Loss were equal to each other (for then the omitted factor would be 1 divided by 1).[16] In some particular cases in various settings, this implicit assumption may be true (or approximately so). But in many cases, perhaps most, this convenient equality of Gain and Loss will not hold. As the foregoing illustrations suggest, they often will differ a great deal, even by orders of magnitude, and in either direction. Hence, the social welfare consequences of operating under the preponderance rule in most legal settings involving the regulation of future conduct would be quite adverse. As we now turn to the domain of litigation, which is often concerned with the provision of appropriate ex ante incentives, a different optimal rule is derived, but it remains true that ignoring consequences, as is done under the preponderance rule, is inappropriate if the objective is to advance social welfare.

---

[16]Alternatively, one could consider the class of preponderance-like rules: those that insert a multiplier m into the formula. If one then sets m equal to Loss / Gain, the rule would be optimal. The problem is that the Loss / Gain ratio tends to vary substantially across situations—just as it does in the medical treatment context. Once one allows the m to differ in each case as a function of the pertinent Gain and Loss, one no longer has a test that targets a posterior probability and thereby ignores these consequences.

3.3 Optimal Threshold: Ex Ante Behavior

Policy analysis of the substantive law as well as much work on the legal system per se is substantially or exclusively concerned with the provision of incentives for ex ante behavior. Deterrence—the provision of disincentives for the commission of harmful acts—is often taken as the central objective. Sometimes, particularly when the subject concerns uncertain evidence and the possibility of error, attention is also given to the prospect that liability may chill beneficial conduct. Chilling, like deterrence, is a form of discouragement of ex ante activity, but it is usefully described with a separate term to distinguish the reduction of desirable acts (those of type *B*) as an incidental effect of attempts to penalize harmful acts (those of type *H*). The social objective is to maximize overall welfare, where deterrence contributes positively and chilling negatively.[17]

To begin, it is necessary to describe this setting in more detail. Suppose that some individuals might commit *H* type acts and others might commit *B* type acts. Each type of act, ex ante, is associated with different distributions of evidence strength, as described in subsection 2.1. The legal system will apply a sanction when the evidence is strong enough to meet the stated threshold. The prospect of such a sanction will generate an expected sanction associated with *H* type acts and another—hopefully much lower—expected sanction associated with *B* type acts.[18]

To determine the optimal evidence threshold, it is useful (as in many optimization problems) to consider how social welfare changes as we adjust the policy instrument. Specifically, let us examine the effects of reducing the evidence threshold slightly. That is, we will imagine a small reduction in the minimum evidence strength required for liability. (The implication is that, for cases with evidence strength just below the old threshold, the outcome is now taken to be liability rather than no liability.)

First, lowering the evidence threshold will increase deterrence somewhat, for those whose acts are of type *H* will expect to be held liable a bit more often. The social benefit of this change will be the product of the reduction in the number who commit *H* acts and the social benefit per deterred *H* act. Again, we will use the term Gain to denote the latter, the pertinent net benefit, which in the present context is the harm avoided minus the forgone value that an

---

[17]In this subsection, we will continue to abstract from the costs of operating the legal system, including any social costs associated with the application of sanctions per se. For an indication of how including such costs would affect the analysis, see note 26 and the accompanying text. In addition, analysis is confined here to the case in which chilling matters as well as deterrence. Some earlier work considered models with only deterrence, wherein the only cost of mistaken imposition of sanctions was to dilute deterrence somewhat (because failure to cause harm is less attractive). In that model, the optimal rule takes a particularly simple form: $LR(e) > 1$. See Kaplow (2011a, 2011b). This rule does differ less, qualitatively, from the preponderance rule, but it nevertheless remains importantly different because the critical likelihood ratio does not depend on the ratio of priors.

[18]In addition to ignoring system costs, as mentioned in the preceding note, the present discussion employs other simplifying assumptions, including taking the rate of inflow of cases into the legal system as given (and thus ignoring how private litigants or public prosecutors' incentives to sue might be influenced by the evidence threshold) and abstracting from the possibility of settlement. Additional dimensions are discussed in Kaplow (2011a, 2012) and some are analyzed in Kaplow (2013c).

individual would have derived from commission of the act.

Second, reducing the evidence threshold will also chill a greater number of acts because those contemplating acts of type *B* also will expect to be held liable more often. The social cost will be the product of the reduction in the number who commit *B* acts and the Loss per chilled *B* act, which is the forgone value that individuals would have received from such acts.

If the incremental benefit (the product of the number of acts deterred and the Gain per deterred act) exceeds the incremental cost (the product of the number of acts chilled and the Loss per chilled act), then reducing the evidence threshold will raise social welfare. If the opposite is true (the cost exceeds the benefit), then increasing the threshold will increase welfare. At the optimum, the marginal benefit will equal the marginal cost.

Where, in this cost–benefit analysis, does the evidence strength come in? The answer is that it influences a subcomponent of the number of acts deterred and the number chilled as a consequence of adjusting the evidence threshold. Suppose, for example, that, for evidence *e* at and near a contemplated evidence threshold, such evidence is generated often by *H* acts but rarely by *B* acts. Then, reducing the evidence threshold slightly—applying sanctions in additional cases with values of *e* in this range—will tend to contribute much more to deterrence than to chilling. The reason is that, in the cases for which sanctions are newly to be applied, many of these will be ones involving *H* type acts and few will be ones with *B* type acts.

To make this more precise, we can employ the notation used throughout. Recall that $LR(e) = P(e|H) / P(e|B)$. Now, what, again, is $P(e|H)$? It is the probability that we observe *e* given an act *H*. As just stated, the greater this is, the more it is true that applying a sanction when such an *e* is observed will deter *H*-type acts. Indeed, all else equal, it is proportional: For example, if $P(e|H)$ is twice as high, then applying the sanction when *e* is observed means we have twice the contribution to the expected sanction for acts of type *H*. The increase in the expected sanction will also depend on other features of the legal system (such as what portion of acts of type *H* enter the legal system and the magnitude of the sanction), but those are held constant in the present analysis. Next, we need to consider how many additional *H* type acts will consequently be deterred. This quantity will depend not only on the magnitude of the increase in the expected sanction but also on the concentration of individuals whose private benefits from *H* type acts make them nearly indifferent, for these will be the individuals deterred as a consequence of a slight reduction in the evidence threshold.

In sum, a number of factors determine how many *H* acts will be deterred. Holding everything constant except $P(e|H)$, we have a proportional effect. That is, we can write the total quantity of acts deterred in the form $P(e|H) \times$ Deterred, where Deterred refers to the combination of all the other factors (which we are holding constant). Therefore, we can write the social benefit from lowering the evidence threshold with regard to deterrence as $P(e|H) \times$ Deterred $\times$ Gain (that is, the number deterred times the benefit per deterred act).

For the chilling of acts of type *B*, the analysis is the same. Reducing the evidence threshold will have a chilling effect that is proportional to $P(e|B)$, where again the pertinent level of *e* is that at the threshold under consideration. We can express the total number of acts of type

B that are chilled as P(*e*|*B*) × Chilled, where Chilled refers to the other factors (held constant) that influence the chilling of desirable behavior. Finally, the total chilling cost (the number of acts of type *B* that are chilled times the social cost per chilled act) is given by P(*e*|*B*) × Chilled × Loss.

At this point, we can express our requirement that the incremental benefits from reducing the evidence threshold exceed the incremental costs from doing so as

$$P(e|H) \times \text{Deterred} \times \text{Gain} > P(e|B) \times \text{Chilled} \times \text{Loss}.$$

Rearranging terms and, in moving to the second expression, substituting with the definition of the likelihood ratio LR(*e*), we have:

$$P(e|H) / P(e|B) > [\text{Chilled} \times \text{Loss}] / [\text{Deterred} \times \text{Gain}], \text{ or}$$

$$LR(e) > [\text{Chilled} \times \text{Loss}] / [\text{Deterred} \times \text{Gain}].$$

Accordingly, it will be optimal to lower the evidence threshold until this inequality no longer holds, which is to say that the final expression is satisfied as an equality. Let LR* denote the likelihood ratio for which this is so. In that case, the optimal rule will involve assigning liability if and only if the evidence is stronger, that is, when the evidence *e* in the case at hand is such that LR(*e*) > LR*, where the later, as just stated, equals the right side of the preceding expression.[19]

Having derived the optimal rule, let us now interpret it. The numerator on the right side tells us that a greater number chilled per unit increase in the likelihood of applying the sanction and a greater social loss per act of type *B* that is chilled both indicate a higher threshold, which is to say that when magnitude and cost of chilling are larger, we require a higher likelihood ratio (stronger evidence) to apply the sanction. The denominator on the right said tells us that a greater number deterred per unit increase in the likelihood of applying the sanction and a greater net social benefit per act of type *H* that is deterred both favor a lower critical likelihood ratio, that is, weaker minimum evidence strength should be required to assign liability. Taken together, the lesson is that, the higher the ratio of the incremental chilling cost to the incremental deterrence gain, the larger the minimum evidence strength we should demand.

Two comparisons prove instructive. First, we can compare this optimal rule, when the social problem concerns incentives for ex ante behavior, to the optimal rule when the problem is

---

[19]The presentation in the text takes some license with notation to ease the exposition. In the foregoing derivation, *e* referred to the evidence just at the evidence threshold under consideration, so in a sense the *e* in the preceding expressions refers to the evidence threshold rather than the evidence in a particular case. Having set the optimal evidence threshold, as just explained, liability is then assigned when the evidence (*e*) in a particular case is stronger than the threshold value of the evidence, that is, when LR(*e*) > LR*. Put another way, LR* is evaluated at the threshold evidence level, which we previously designated as *e**. So the requirement can also be expressed as *e* > *e**. (Relatedly, on the right side of the expression that determines the optimal rule, the terms Chilled, Loss, Deterred, and Gain each implicitly depend on the chosen evidence threshold (*e**), in the manner described in the preceding text, but not on the evidence (*e*) in the particular case.)

the regulation of future conduct. For future conduct, instead of Chilled in the numerator and Deterred in the denominator—both of which calibrate the magnitude of *marginal* effects of *changing* the evidence threshold on *behavior*—we instead had P(*B*) in the numerator and P(*H*) in denominator—both of which refer to the *preexisting* (that is given) *levels* of each of the two types of acts, *B* and *H*.

On one hand, we have important similarities. Both are likelihood ratio tests. Both depend on the consequences. And, in both cases, the consequences depend on gains and losses associated with the two types of acts and frequency-related weighting factors.

On the other hand, there exist strong differences because these weights are qualitatively different. Here, for ex ante behavior, the weighting factors are, as emphasized, marginal impacts—deterrence and chilling—that is, effects due to changes in the levels of the two types of activities. The ratio of the weighting factors refers to the comparison of these deterrence and chilling impacts. Previously, for future conduct, the weighting factors were the levels of the two types of activities that individuals would like to undertake. These levels are taken as given, and the question is what portion will be permitted by the legal system (by contrast to which actions individuals will voluntarily forgo in light of their expectations of the costs that the legal system would impose if they chose otherwise).

Second, let us compare the current optimal rule, involving settings in which ex ante behavior is influenced, with the preponderance of the evidence rule (and similar types of rules), which are the rules ordinarily stated to be applicable in such settings. As mentioned, the preponderance rule ignores consequences. Indeed, if we compare the optimal rule derived in this subsection to the preponderance rule in subsection 3.1—both expressed as likelihood ratio tests—we can see that the determinants of the critical or threshold likelihood ratio are categorically different.

To see this clearly, let us review these two rules. First, the preponderance rule:

$$\text{LR}(e) > \text{P}(B) / \text{P}(H).$$

Now, the optimal rule:

$$\text{LR}(e) > [\text{Chilled} \times \text{Loss}] / [\text{Deterred} \times \text{Gain}].$$

The prior probabilities, the only terms making up the critical likelihood ratio under the preponderance rule, are irrelevant under the optimal rule. And each of the terms constituting the critical likelihood ratio under the optimal rule—all of which are concerned with consequences—are irrelevant under the preponderance rule. There is no overlap. (In a more complete analysis, certain background conditions may affect both P(*B*) and P(*H*), on one hand, and Chilled and Deterred, on the other hand, but the large qualitative differences remain for the

reasons noted momentarily.[20])

Therefore, although both sorts of rules can be viewed as likelihood ratio tests, when we express them in this parallel format it is immediate that they have nothing else in common. Under conventional formulations like the preponderance rule, the critical threshold is stipulated at a level that depends on the prior probabilities (base rates), and on nothing else.[21] Under the optimal rule, it is determined based on the consequences for social welfare, which depend both on how behavior changes and on the social welfare impact of those changes in behavior, neither of which are relevant under the preponderance rule.

Conventional burden of proof notions like the preponderance of the evidence rule—in looking at given prior probabilities rather than at how behavior (and thus the level of each probability) changes as a function of how the evidence threshold is set—implicitly take a static view. That perspective, we saw in subsection 3.2, is appropriate with regard to the regulation of future conduct, where, ironically, it seems that preponderance-like rules are not ordinarily employed, yet it is entirely inapt in the present context with regard to the legal system's creation of incentives for ex ante behavior, where such rules are used. Second, we have seen that, in both legal contexts, these standard rules ignore consequences for social welfare that are central in determining the critical likelihood ratio under optimal formulations. About the only shared feature is that, since all can be expressed as likelihood ratio tests, all are more inclined to prohibit conduct or to assign liability when the evidence is stronger, where evidence strength is understood in the particular manner that is captured by the likelihood ratio, $LR(e)$.[22]

4. Optimal Threshold: Multistage Adjudication

This section extends the analysis in subsection 3.3, involving optimal evidence thresholds when the concern is with ex ante incentives, to a setting in which cases may be removed from the legal system before final adjudication. This setting encompasses formal pretrail decision points—such as when a civil case may be dismissed or a criminal indictment or other prerequisite to continuation may not be issued—as well as myriad actions that may be more informal, such as investigative decisions, agency screening, and other exercises of prosecutorial discretion.

---

[20]Kaplow (2011a, pp. 1110 & n. 7, 1119) explains how this is so with regard to factors bearing on the volume of acts of each type entering the legal system.

[21]If we considered the class of preponderance-like rules discussed at the end of subsection 3.1, allowing the insertion of a constant factor m on the right side, we would still have no elements in common. See also the discussion in note 16, above.

[22]Even this point is subject to the subtle but important caveat that conventional rules can violate basic notions of consistency and coherence due to the fact that, in settings in which legal rules influence behavior and thus prior probabilities, they attempt to target a ratio (of prior probabilities) that is itself endogenous to the rule. See Kaplow (2011a, pp. 1120–21; 2012, pp. 790–95 & n. 93). To suggest the nature of the problem, note that raising the evidence threshold, which might be supposed to raise the likelihood that cases just above the threshold involve acts of type H, reduces the chilling of acts of type B, which increases their flow into the legal system and thus may lower the posterior probability that an act generating evidence at the new (higher) threshold is of type H.

Specifically, we will continue to suppose that there are two types of acts, *H* and *B*, that individuals will commit if and only if their private benefit exceeds the expected costs imposed by the legal system. These expected costs, however, will now be understood in a more complicated manner.

Working backward, suppose for simplicity that we have a two-stage system, and the final stage, adjudication, is as described in subsection 3.3, and in this section we will take that decision rule as given. It may or may not be optimal, but in any event the rule indicates whether the sanction will be imposed for any set of evidence (now understood as that available when a case enters the legal system and the further information that is generated if the case is allowed to continue, on which more in a moment).

To this, we add a preliminary (first) stage at which a decision is made whether to terminate the case (in which event no sanction will be assigned) or to allow it to continue. If the latter course is taken, there will be an interim cost (associated with additional investigation, discovery, or some other process) that yields additional information. For present purposes, it will be convenient to let the evidence *e* refer the information available *at the outset*; this set of evidence, depending on the setting and the particular case, may be quite meager. For purposes of our analysis, it simply is what it is. The preliminary, first-stage decision must be made by considering only this evidence *e*, whereas if the case is continued, at the posited cost, the final decision will reflect the supplemental information as well.

In this setting, consider the expected costs faced by an individual contemplating whether to commit an act of either type. As before, we have an expected sanction, but now this will reflect the likelihood that the case will be continued at the first stage as well as the further probability, conditional on continuation, that liability will be imposed at the second stage. In addition, whenever there is continuation, we are supposing that the individual (who at that point may be thought of as the defendant) will incur some adjudication cost on account of continuation itself. That is, even if ultimately victorious in final adjudication, the individual will have borne this cost along the way if the case is continued rather than terminated at the first stage.[23]

To derive the optimal stage-one termination/continuation rule, consider any particular set of evidence *e* that may be available at this first stage, which will be referred to as a scenario. For each scenario, we wish to assess the consequences for social welfare of continuation relative to a baseline of termination, taking as given how these termination/continuation decisions will be made in other scenarios, that is, for other possible sets of evidence *e*.[24]

---

[23]Of course, even without introducing this additional, preliminary stage, individuals in general incur adjudication costs, which have been set to the side until now. Some of the lessons derived from this extension will be distinctive to the costs associated with continuation at a nonfinal stage of adjudication, whereas others will be more broadly associated with the presence of any sort of legal costs that actors expect to bear as a consequence of the acts they choose to commit.

[24]As will become apparent, elements of the optimal rule depend an aggregates, such as the overall levels of deterrence and chilling, and hence on how decisions are understood to be made in other scenarios. The same was implicitly true when considering final adjudication. The difference is that, there, with a conventional likelihood ratio test, specification of a single decision threshold LR* completely describes the outcomes for all possible evidence *e* because any evidence *e* is associated with a unique likelihood ratio LR(*e*) and LR* indicates the resulting decision.

Continuation will have one benefit and two costs. The benefit is enhanced deterrence. Just as in subsection 3.3, we wish to know how many more acts will be deterred and the net social benefit per deterred act. Regarding the former, instead of certain application of the sanction (which happens when the legal threshold is exceeded in final adjudication), continuation is associated with a probability that the sanction will be applied at stage two. In addition, continuation in and of itself augments prospective actors' expected costs of their acts (they bear additional litigation costs), which also contributes to deterrence. Regarding the latter, note also that, because our model now includes adjudication costs, each act deterred has an additional social benefit: Any act not committed is one that does not have the prospect of entering the legal system and thus generating such costs.

The first cost of continuation is due to chilling, and it is determined analogously to deterrence. Again, even if the ultimate likelihood of sanctions is low—perhaps continuation is highly likely to generate additional evidence demonstrating that the act is of type *B*—the prospect of continuation nevertheless has a chilling effect because adjudication costs will be incurred. In addition, the social cost associated with a benign act being chilled now has an offset due to the fact that a chilled act is not one that has the possibility of generating adjudication costs.

The second cost of continuation is the obvious one, the cost of continuation itself. Note further that continuation costs are likely to be scenario-specific. That is, based on the initially available evidence *e*, it may be apparent that subsequent adjudication will be particularly costly or not very costly at all. For example, it may sometimes appear that only an exhaustive search of a huge organization's documents and depositions of many of its employees will get to the bottom of the matter, whereas in other situations a single witness or a few documents may suffice.

At this point, we can state the optimal decision rule in the form of a likelihood ratio test (which in some respects resembles that in subsection 3.3 for the burden of proof at the final stage of adjudication) as follows:[25]

$$\mathrm{LR}(e) > \frac{[\mathrm{Chilled}(e) \times \mathrm{Loss}] + [\mathrm{P}(B) \times \mathrm{ContinCost}(e)]}{[\mathrm{Deterred}(e) \times \mathrm{Gain}] - [\mathrm{P}(H) \times \mathrm{ContinCost}(e)]} \, .$$

In interpreting this rule, we will initially set to the side the fact that many of the terms on the right side are a function of the evidence *e* in the particular scenario, a very significant feature to which we shall return.

Both the numerator and the denominator on the right side have, as their first elements (in square brackets), roughly what we had before for the optimal burden of proof in final-stage adjudication. As mentioned, however, both the number chilled and the number deterred is

---

As will emerge below, such is not the case in the present setting.

[25]Because the steps that generate this formula are similar to those presented in subsection 3.3, and the modifications are indicated by the foregoing discussion, they are omitted. See Kaplow (2013a, 2013b) for elaboration.

determined somewhat differently. In addition, both the Loss and the Gain now each have an added component of avoided system costs because, for each act chilled or deterred, there is now the associated reduction in expected legal system costs. (However, if our model of final-stage adjudication had been more complete, it too would have included adjudication costs, and there would have been this added component in the Loss and Gain for the optimal burden of proof rule as well.[26])

The most apparent difference is that we now have an additional, second term (in square brackets) in each of the numerator and denominator. These reflect continuation costs. In the numerator, continuation costs are added: The numerator collects all the consequences associated with acts of type B, and they are chilling costs plus system costs (that is, regarding benign acts, we have two costs, so they are added). In the denominator, continuation costs are subtracted: The denominator collects all the consequences associated with acts of type H, and they are a deterrence benefit minus system costs (that is, regarding harmful acts, we have a benefit and a cost, so the latter must be subtracted from the former).[27]

In both instances, the continuation costs are the product of two factors. The latter, ContinCost(e), is the continuation cost per case in that scenario that is actually continued. The former factor in each instance is the prior probability, which is familiar from earlier discussion. It reflects the fact that expected continuation costs depend on the number of each type of act that ends up in the legal system in the scenario in question, and that, in turn, will depend on the frequency with which each type of act is committed.[28]

Having described each of the core elements, let us now return to the point that many of the factors on the right side of this optimal likelihood ratio test are functions of e. First, note that Chilled and Deterred each have this property. Consider Deterred(e): As explained above, deterrence overall depends on the expected costs of acts, which in turn are the sum of expected sanctions and expected system costs. Regarding expected sanctions, the degree to which they are increased as a consequence of continuation in the scenario associated with evidence e depends importantly on the probability that liability will be assigned in the second, final stage of adjudication. Moreover, it should be apparent that this probability will depend on the e that we observe at the first stage. Suppose, for example, that e is very strong; then, we have a high predicted probability of liability at stage two. Of course, there is some chance that we will learn that the case is weak after all, but the meaning of strong evidence at stage one is that this is not very likely. By contrast, if e is weak at stage one, then, even if we continue the case, stage two liability is not very likely, so the contribution to the expected sanction for an act of type H will be small.

---

[26]Introducing direct social costs of sanctions would affect the optimal decision rule in a similar manner.

[27]The denominator could be negative, in which case it would typically be optimal to terminate. (Not necessarily because, as mentioned, the Loss from chilling now has the offset that chilled acts avoid possible system costs; hence, the numerator could also be negative, which would favor continuation.)

[28]Some license is taken in presenting this expression by implicitly normalizing in various respects. What actually matters is the volume of cases generated by the two types of acts, not the percentage that are of each type. For details, see Kaplow (2013b).

Moreover, as mentioned above, continuation costs depend on the scenario, that is, on the evidence *e*. And, as was explained, continuation costs also contribute to deterrence. Therefore, the increase in deterrence as a consequence of continuation in a given scenario depends on the evidence *e* in two ways: through expected sanctions and through expected continuation costs.

For Chilled(*e*), the analysis is the same. Those contemplating acts of type *B* are discouraged both by the possibility of being sanctioned and by the prospect of bearing adjudication costs. The degree to which each of these components is augmented as a consequence of continuation in the scenario at hand depends, just as with deterrence, on the evidence *e*.

Finally, as already mentioned, ContinCost(*e*) depends on the evidence *e* in the given scenario. Taking all of this together, we see that, unlike with the burden of proof in final adjudication, the optimal termination/continuation rule, stated as a likelihood ratio test, depends importantly on the evidence *e*.

This fact raises the conceptual question whether this decision rule should really be understood as a likelihood ratio test. In a superficial sense, the answer is affirmative: After all, we wrote it that way. And we have already seen that there is value in doing so because it is easier to compare and contrast this optimal rule with that for final-stage adjudication, as we already have done. In addition, if we had other proposed rules for the termination/continuation decision, we could put them in the form of a likelihood ratio test (if feasible) and make further comparisons—on which more in a moment.

In an important, substantive sense, however, the foregoing optimal rule, although presented in the *form* of a likelihood ratio test, is not aptly viewed in this manner after all. The reason is that it is *not* true that evidence *e* that is associated with a higher value of LR(*e*) necessarily implies a stronger case for continuation. Remember that previously we said that a likelihood ratio test is one that makes a designated decision (here, continuation) if and only if LR(*e*) > LR*, for some stipulated value LR*. And, if the test is satisfied for some *e*, it is necessarily satisfied for any *e* with a higher likelihood ratio. Similarly, if it fails for some other *e*, it necessarily fails for any *e* with a lower likelihood ratio.

Not so here. The reason is precisely that the right side of this decision rule is itself a function of *e*. To see the implications, suppose we have some *e* that has LR(*e*) = 2.1. Moreover, when we evaluate all the factors on the right side of the above rule for this *e*, it turns out that the value of the right side is 1.8. Because 2.1 > 1.8, the test is satisfied, so continuation is optimal.

Now consider some other *e* that has LR(*e*) = 3.7. However, when we assess all the factors on the right side for this other *e*, it turns out that the value is 5.6. Obviously, 3.7 < 5.6, so the test fails, meaning that termination is optimal. Taken together, for our initial *e* that had LR(*e*) = 2.1 we wish to continue, whereas for this *e* with LR(*e*) = 3.7 we wish to terminate. How might this happen? Perhaps ContinCost(*e*) is much higher in the second case, which raises the numerator and reduces the denominator, both raising the value on the right side. There are additional effects through Chilled(*e*) and Deterred(*e*) which could further raise the right side or lower it. We are simply imagining an example in which the overall effect is as stated.

Combining these two cases, we can see that it is indeed possible—and can arise through a number of channels by which *e* affects the value of the right side of our expression—that the optimal rule will favor continuation when the likelihood ratio is lower than in other instances in which it favors termination. On reflection, this should not be that surprising because we can imagine a case that is slightly more promising in terms of the strength of the initial evidence but involves massively higher continuation costs than is normal.[29] In other (analytically simpler) settings like medical treatment, suppose that one patient, in light of initial evidence, is a bit more likely than another to have an ailment, but the next step is an invasive test that poses an unusually high risk to that particular patient (but not to the other). Hence, we may not conduct the test for the first patient even though the likelihood ratio is higher than the level that might ordinarily trigger this intervention, which may be optimal for the second patient.

Therefore, although we can express this optimal legal decision rule in the *format* of a likelihood ratio test, there is an important sense in which it is not one. To review the logic concerning the latter point, consider the following. Suppose someone were to state that "I hire everyone with a test score above a cutoff, but I must confess that I employ a different cutoff for each applicant, and, moreover, the different cutoff is a function of all manner of information about the applicant, including some that are influenced by the test score itself." Obviously, such pronouncements are not informative about the use of the test score in the same manner as when there is a given, common cutoff for everyone.[30]

Having just discussed the sense in which the optimal decision rule for nonfinal stages in multistage adjudication is not a true likelihood ratio test, let us now turn briefly to conventional formulations. In contrast to the burden of proof (particularly the preponderance of the evidence rule, which is interpreted as requiring a posterior probability exceeding fifty percent), it is less clear what, if any probability is required at earlier stages.

For concreteness, consider motions to dismiss under the U.S. Federal Rules of Civil Procedure. Under one view, which may have reflected blackletter law (although perhaps not actual practice) until recently, one might have viewed the minimum probability as zero or perhaps only slightly above zero: A case should be continued if it was merely logically possible

---

[29]As an alternative to a likelihood ratio test conception, contemplate instead a heuristic that ranks scenarios by the desirability of continuation and then chooses a cutoff, above which cases would be continued and below which they would be terminated. In light of the argument in the text, the ranking would not follow likelihood ratios; instead, it would directly follow the benefit–cost difference, according to the optimal decision rule itself. This heuristic, however, is only helpful to frame thinking, for one must analyze each component of the optimal decision rule in each scenario to do the ranking. Moreover, because of the multidimensionality of the optimal test and the interdependence across scenarios, there does not exist a unique, correct ordering—instead, the ordering depends on the chosen cutoff. For example, in some scenario, the contribution to deterrence may be very high, but the continuation costs may be substantial as well. If the optimal pattern of decisions involved termination in most other scenarios, continuation may be quite favorable, generating a high rank, because the deterrence deficit would be large. But if the opposite were true, incremental deterrence might not be very valuable, in which event the scenario would rank low. More abstractly, as one moves the cutoff, both the absolute and relative desirability of continuation in other scenarios will change, and the latter can alter the ranking.

[30]To take an extreme case, suppose that the cutoff simply rises at twice the rate of the test score; then the person might, after conversion, really be taking all individuals with scores *below* some (fixed) cutoff! For example, if the rule is score(actual) > score*, and score* = 2×score(actual) – 10, then the rule reduces to 10 > score(actual).

that a violation occurred.[31] However, recent Supreme Court decisions in *Twombly* and *Iqbal* seem to require more.[32] In each case, the Court disclaims that the new "plausibility" standard is one gauged by probabilities and then immediately articulates the proper inquiry in probabilistic language.[33] Of course, if the test is to be related to consequences for social welfare, it surely has to depend on probabilities that the information at hand was generated by harmful versus benign acts. More precisely, it would seem that any appealing rule would have to be expressible in the form of a likelihood ratio test. As previously discussed, the logic is that, if some case is to be continued, then another case that is identical in all respects except that its LR(*e*) is higher should be continued as well, and if some other case is to be terminated, then another that is identical except that its LR(*e*) is lower should be terminated too.[34] Hence, it is difficult to understand how a legal rule for nonfinal stages of adjudication that violates this property could sensibly be defended.[35]

5. Likelihood Ratios versus Bayesian Posterior Probabilities in the Formulation of Legal Decision Rules

This article's thesis is that it is useful to formulate legal decision rules as likelihood ratio tests. At this point readers can draw their own conclusions about the extent to which this exercise pays off. An implicit thesis that this section now makes explicit is that it is not as useful to formulate legal decision-making criteria in terms of Bayesian posterior probabilities; indeed, this common and alluring approach can lead us astray.[36]

To begin, it is useful to restate what we are talking about. The likelihood ratio, LR(*e*), is given by P(*e*|*H*) / P(*e*|*B*), where each of these terms refer to the probability of the evidence *e* given the situation, *H* or *B*. Bayesian posterior probabilities, P(*H*|*e*) and P(*B*|*e*), instead indicate the probabilities that *H* and *B*, respectively, are true given the evidence *e*. This distinction, as we have seen, can make a large difference: Whether with regard to *H*, to *B*, or to various of the ratios, either of these sorts of conditional probabilities could be larger or smaller, and to any degree, than the reverse form of the conditional.

This key difference does not per se imply the usefulness of working with one rather than the other. After all, the two are linked—by Bayes' Theorem—so it would seem that the stakes

---

[31]This view is well captured by the language of the U.S. Supreme Court in Conley v. Gibson, 355 U.S. 41, 45–46 (1957), under which a case could not be dismissed unless there existed "no set of facts" consistent with liability.

[32]See Bell Atlantic Corp. v. Twombly, 550 U.S. 544 (2007), and Ashcroft v. Iqbal, 129 S. Ct. 1937 (2009).

[33]For further discussion, see Kaplow (2013a, §IV.A).

[34]Even though, as elaborated just above, the optimal decision is not a true likelihood ratio test because the target likelihood ratio depends on the available information, it remains true that, if the values of all the components on the right side (put another way, if everything relevant to welfare) were indeed the same, then indeed a higher likelihood ratio would favor continuation over termination.

[35]Part of the challenge in examining court opinions and legal commentary is that proffered legal rules are not articulated using sharply defined terms, so it can be quite difficult to ascertain what sort of decision criterion is under discussion, which in turn makes it hard to analyze what its virtues and deficiencies might be.

[36]For a largely different but partially overlapping discussion, see Kaplow (2012, pp. 781–814).

are small (at most). This point turns out to be true in some settings but not in others. It is strongest with regard to conventional burden of proof rules, such as the preponderance of the evidence rule. As we saw in subsection 3.1, it can be equivalently stated in terms of the posterior probabilities $[(P(H|e) / P(B|e) > 1]$, or in terms of the likelihood ratio $[\text{LR}(e) > P(B) / P(H)]$. And we can readily translate between them using the odds ratio version of Bayes' Theorem:

$$[P(H) / P(B)] \times \text{LR}(e) = P(H|e) / P(B|e).$$

Even here, there is arguably some virtue in presentation of the requirement as a likelihood ratio test: It makes clear the need to ascertain both the likelihood ratio, $\text{LR}(e)$, and the ratio of the prior probabilities, $P(B) / P(H)$. The explicit use of Bayes' Theorem in adjudication has proved controversial, among other reasons, because of the need to determine the prior probabilities, which are variously suggested to be outside the factfinder's knowledge, difficult to appreciate, an inappropriate subject for the presentation of evidence, or distracting with regard to the proper focus of decision making.[37] These views are strange, however, if indeed the factfinder is supposed to assess liability using the preponderance rule (or other rules of that genre) because the rule requires knowledge of the posterior probabilities which, in principle, can only be determined using information on the prior probabilities and on the likelihood ratio.[38] Accordingly, what is sometimes proffered as simplification may more often constitute obfuscation.

Differences of substance emerge as we move from conventional formulations of legal decision rules to optimal ones. In each instance, we saw how optimal criteria depend on consequences that are omitted from conventional rules—which focus on a version of truth (Bayesian posterior probabilities) rather than consequences (social welfare). In the setting analyzed in subsection 3.2, involving future conduct, the difference between formulation as a likelihood ratio test $(\text{LR}(e) > [P(B) / P(H)] \times [\text{Loss} / \text{Gain}])$ and a restated version using Bayesian posterior probabilities (which would be $P(H|e) / P(B|e) > \text{Loss} / \text{Gain}$),[39] is not great, consistent with the previous discussion. However, we have seen that in the setting analyzed in subsection 3.3, involving incentives for ex ante behavior, the prior probabilities needed to determine the Bayesian posterior probabilities do not appear anywhere in the optimal decision rule, so there is a strong sense in which it cannot usefully be stated otherwise. (Moreover, this point renders ironic the controversy in legal scholarship about the use of Bayesian prior probabilities in that they are necessary under the decision rules that opponents of their use seem to endorse but irrelevant under an optimal criterion of which they are unaware.) Finally, moving to the optimal rule for multistage adjudication that is derived in section 4, where the concern is again taken to be with ex ante behavior, it would be unhelpful to restate the more complex

---

[37]See, for example, the references cited in Kaplow (2012, pp. 796–99).

[38]In some settings, there may exist direct information on the posterior probabilities, in which case starting (and ending) the analysis with them would be appropriate, but even then the information implicitly is derived from priors and the likelihood ratio.

[39]This formula can readily be verified by starting with the likelihood ratio version, moving the priors to the left side, and substituting for the result on the left side using the odds ratio version of Bayes' Theorem from subsection 2.2.

optimal rule for that case as a test based on the Bayesian posterior probabilities.[40]

Likelihood ratio tests prove to be a useful unifying lens through which to examine and compare legal decision rules. In many respects, Bayesian posterior probabilities do not. The most central differences emerge in subsection 3.3 and section 4, where the concern is with the incentives provided for ex ante behavior: the deterrence of harmful (*H*) acts and the chilling of benign or beneficial (*B*) acts. The reason is that what matters in assessing marginal adjustments of legal decision rules is how they change these incentives, and the relative impacts on the two types of acts depend (in an important but not exclusive manner) on how likely they are to generate the evidence *e* that is observed in the scenario at hand. The likelihood ratio, LR(*e*), captures that notion in relative terms. Bayesian posterior probabilities do not: Through Bayes' Theorem, we can see that they do depend on the likelihood ratio, but they also depend on the prior probabilities, P(*H*) and P(*B*), which are immaterial for purposes of assessing incentive effects. Moreover, even when the prior probabilities do matter, such as in regulating future conduct, the likelihood ratios matter as well, and presenting decision rules as likelihood ratio tests is at least as transparent and illuminating.

Why is it that most legal analysts are so much more used to thinking in terms of Bayesian posterior probabilities than likelihood ratios? One reason is that, as noted, posterior probabilities correspond to truth in some sense, and many take adjudication to be aimed at finding the truth. Although greater accuracy does tend to be advantageous in many settings, more modern analysis takes accuracy as a means to an end, and the pertinent end is the promotion of social welfare.[41]

Another explanation for current views is that decision analysis has typically been developed in settings in which the concern is with determining appropriate future conduct—such as in choosing a course of medical treatment or in deciding whether to make some investment. By contrast, much legal policy analysis concerns the provision of ex ante incentives, typically with a focus on deterrence but sometimes as well with attention to the problem of chilling desirable conduct. This difference has not been sufficiently appreciated. As one indication—which perhaps combines these two reasons—most would endorse the proposition that, in making legal decisions, we should strive to "get it right" in every case, where correct outcomes are taken to be those that reflect our best guess at the truth (implicitly, the preponderance rule), or something along these lines. That there is a substantial, qualitative disjunction between this perspective and optimal policy is not usually recognized, much less taken as a beacon that should guide analysis..

---

[40]If one examines the optimal rule, it does contain the prior probabilities in the latter components of the numerator and the denominator on the right side, but they are absent in the former components, which are analogous to those in subsection 3.3 for the optimal burden of proof.

[41]For extended discussion, including of some respects in which truth may be relevant to welfare (e.g., in influencing system legitimacy), see Kaplow (1994) and Kaplow and Shavell (2002, ch. 5). And for criticism of alternative ex post objectives proffered in the legal literature, see Kaplow (2012, pp. 799–805).

References

Cullison, Alan D. 1969. "Probability Analysis of Judicial Fact-Finding: A Preliminary Outline of the Subjective Approach," 1 *University of Toledo Law Review* 535–98.

DeKay, Michael L. 1996. "The Difference Between Blackstone-Like Error Ratios and Probabilistic Standards of Proof," 21 *Law and Social Inquiry* 95–132.

Finkelstein, Michael O., and William B. Fairley. 1970. "A Bayesian Approach to Identification Evidence," 83 *Harvard Law Review* 489–517.

Friedman, Richard D. 1986. "A Close Look at Probative Value," 66 *Boston University Law Review* 733–59.

Johnston, Jason S. 1987. "Bayesian Fact-Finding and Efficiency: Toward an Economic Theory of Liability under Uncertainty," 61 *Southern California Law Review* 137–81.

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

Kaplow, Louis. 1994. "The Value of Accuracy in Adjudication," 23 *Journal of Legal Studies* 307–401.

Kaplow, Louis. 2011a. "On the Optimal Burden of Proof," 119 *Journal of Political Economy* 1104–40.

Kaplow, Louis. 2011b. "Optimal Proof Burdens, Deterrence, and the Chilling of Desirable Behavior," 101 *American Economic Review: Papers and Proceedings* 277–80.

Kaplow, Louis. 2012. "Burden of Proof," 121 *Yale Law Journal* 738–859.

Kaplow, Louis. 2013a. "Multistage Adjudication," 126 *Harvard Law Review* (forthcoming).

Kaplow, Louis. 2013b. "Optimal Multistage Adjudication" (unpublished).

Kaplow, Louis. 2013c. "Instrument Choice and the Incentive to Sue" (unpublished).

Kaplow, Louis, and Steven Shavell. 2002. *Fairness versus Welfare*. Cambridge, MA: Harvard University Press.

Karlin, Samuel, and Herman Rubin. 1956. "The Theory of Decision Procedures for Distributions with Monotone Likelihood Ratio," 27 *Annals of Mathematical Statistics* 272–99.

Koehler, Jonathan J. 1996. "On Conveying the Probative Value of DNA Evidence: Frequencies, Likelihood Ratios, and Error Rates," 67 *University of Colorado Law Review*

859–86.

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

Luce, R. Duncan. 1963. "Detection and Recognition," in R. Duncan Luce, Robert R. Bush, and Eugene Galanter, *Handbook of Mathematical Psychology*, vol. 1: 103–89. New York: John Wiley and Sons.

Macmillan, Neil A., and C. Douglas Creelman. 2005. *Detection Theory: A User's Guide*, 2nd ed., New York: Psychology Press.

Martin, Anne W., and David A. Schum. 1987. "Quantifying Burdens of Proof: A Likelihood Ratio Approach," 27 *Jurimetrics Journal* 383–402.

Milgrom, Paul R. 1981. "Good News and Bad News: Representation Theorems and Applications," 12 *Bell Journal of Economics* 380–91.

Neyman, Jerzy, and Egon Pearson. 1933. "On the Problem of the Most Efficient Tests of Statistical Hypotheses," 231 *Philosophical Transactions of the Royal Society of London, Series A, Containing Papers of a Mathematical or Physical Character* 289–337.

Tribe, Laurence H. 1971. "Trial by Mathematics: Precision and Ritual in the Legal Process," 84 *Harvard Law Review* 1329–93.