

Stanford Law Review

A Presumption of Innocence, Not of Even Odds

Author(s): Richard D. Friedman

Reviewed work(s):

Source: *Stanford Law Review*, Vol. 52, No. 4 (Apr., 2000), pp. 873-887

Published by: [Stanford Law Review](#)

Stable URL: <http://www.jstor.org/stable/1229432>

Accessed: 15/12/2011 05:00

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at

<http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Stanford Law Review is collaborating with JSTOR to digitize, preserve and extend access to *Stanford Law Review*.

<http://www.jstor.org>

COMMENTARY

A Presumption of Innocence, Not of Even Odds

Richard D. Friedman*

Now I know how the Munchkins felt. Here I have been, toiling in the fields of Evidenceland for some years, laboring along with others to show how use of Bayesian probability theory can assist in the analysis and understanding of evidentiary problems.¹ In doing so, we have had to wage continuous battle against the Bayesioskeptics—the wicked witches who deny much value, even heuristic value, for probability theory in evidentiary analysis.² Occasionally, I have longed for law-and-economics scholars to help work this field, which should be fertile ground for them.³

So imagine my delight when the virtual personification of law and economics himself, Judge Richard Posner, came down from a star to help till the evidentiary soil with his powerful economic tools.⁴ And imagine my further delight when his house landed square on the noggins of the Bayesioskeptics. I do not suppose this will be a fatal blow. Like Napoleon’s Old Guard, they refuse to surrender—but neither will they die. Still, I appreciate having such an ally.

It may seem that I should be very gracious and roll out the welcome carpet to Evidenceland for Judge Posner. But I confess that my graciousness is

* Ralph W. Aigler Professor of Law, University of Michigan Law School. Many thanks to Judge Richard Posner, for entering into a dialogue that left him unpersuaded and spurred the writing of this essay, to Carlton Caves, Peter Donnelly, and David Kaye for offering helpful suggestions, and to Paul Diller for quick and valuable research help.

1. See, e.g., Richard D. Friedman, *Answering the Bayesioskeptical Challenge*, 1 INT’L J. EVID. & PROOF 276 (1997); D.H. Kaye, *Clarifying the Burden of Persuasion: What Bayesian Decision Rules Do and Do Not Do*, 3 INT’L J. EVID. & PROOF 1 (1999); Richard O. Lempert, *Modeling Relevance*, 75 MICH. L. REV. 1021 (1977).

2. I am calling the Bayesioskeptics, some of whom are good friends of mine, wicked witches in the nicest possible way, and I hope they will not take offense.

3. See Richard D. Friedman, *Economic Analysis of Evidentiary Law: An Underused Tool, an Underplowed Field*, 19 CARDOZO L. REV. 1531, 1531 (1998) (noting the potential value in applying economic legal analysis to the law of evidence).

4. See Richard A. Posner, *An Economic Approach to the Law of Evidence*, 51 STAN. L. REV. 1477 (1999) (presenting the first comprehensive economic analysis of the law of evidence).

tempered by the fact that Judge Posner commits a serious error in the use of Bayesian analysis. He asserts that an unbiased fact-finder should begin consideration of a disputed case with "prior odds of 1 to 1 that the plaintiff or prosecutor has a meritorious case."⁵ Indeed, Judge Posner considers this starting point the essence of the definition of an unbiased fact-finder. I believe that this view is wrong in principle as a matter of probability theory. It is indeterminate and also fundamentally at odds with the presumption of innocence. Finally, it leads to bizarre results that expose Bayesian analysis to gratuitous ridicule from the wicked Bayesioskeptics.

I. THEORETICAL UNDERPINNINGS

Under the view of probability most useful for evidentiary analysis, a probability assessment represents an observer's subjective level of confidence in the truth of a given proposition, based upon the information that the observer has at any given time. Probabilities may be stated on a scale from 0 (representing certainty that the proposition is false) to 1 (representing certainty that the proposition is true). For some purposes, it is easier to represent probabilities in terms of odds. The odds that a proposition is true equal the probability that the proposition is true divided by the probability that the proposition is not true. Accordingly, odds of 0 represent certainty that the proposition is false; infinitely high odds represent certainty that the proposition is true; and the oxymoronic-sounding "even odds" of 1:1—or 50/50—represent an assessment that the proposition is precisely as likely to be true as to be false.

The use of subjective probability theory in evidentiary analysis does not presuppose that fact-finders actually do, or even should, assess probabilities numerically or consciously at all. It only supposes that rational people act consistently with implicit probability assessments. An adult may run into the street to retrieve a bouncing ball if it is absurdly unlikely that an approaching car would hit her when she tried to do so, but she presumably would let the ball bounce away if that possibility seemed entirely plausible; on the other hand, because the stakes are so much different, she probably would run into the street to retrieve her bouncing baby even if the chances of her being run over during the rescue seemed quite high. None of this decision making requires calculation or numerification of odds.

When an observer receives new evidence relevant to the truth of the proposition at issue, she adjusts her probability assessment to take that evi-

5. *Id.* at 1514 ("Ideally we want the trier of fact to work from prior odds of 1 to 1 that the plaintiff or prosecutor has a meritorious case. A substantial departure from this position, in either direction, marks the trier of fact as biased."). See also *id.* at 1508 (hypothetically assuming prior odds of 1 to 1 "on the theory that the jury begins hearing the evidence . . . without any notion of who has the better case").

dence into account. An important principle indicating how this should be done rationally is Bayes' Theorem, which is so significant that the theory of subjective probability is often referred to, perhaps misleadingly, as Bayesian theory.⁶ A simple statement of Bayes' Theorem uses three terms. One is the *prior odds* of a proposition—that is, the odds as assessed before receipt of the new evidence. The second is the *posterior odds* of the proposition—that is, the odds that the proposition is true as assessed after receipt of the new evidence. And the third is the *likelihood ratio*. Simply defined, the likelihood ratio of a given body of evidence with respect to a given proposition is the ratio of the probability that the evidence would arise given that the proposition is true to the probability that the evidence would arise given that the proposition is false.⁷

Bayes' Theorem posits that the posterior odds of the proposition equal the prior odds times the likelihood ratio. In simple notation,

$$O_{\text{pos}} = O_{\text{pr}} \times L. \quad (1)$$

Thus, all other things being equal, the posterior odds will be higher: (a) the higher the prior odds; (b) the higher the probability that the evidence would arise given the truth of the proposition; and (c) the lower the probability that the evidence would arise given that the proposition is false. A likelihood ratio greater than 1 means that the proposition appears more probable in light of the new evidence; a likelihood ratio less than 1 means that the new evidence makes the proposition appear less probable; and a likelihood ratio of precisely 1 means that the new evidence leaves the probability unchanged.

Now, one problem in application of the theory is determination of the prior odds of a proposition in the face of virtual ignorance of information bearing on the truth of the proposition. I speak of virtual rather than of total ignorance, because if the observer understands the proposition at all then she must have at least some minimal level of information about the world and about the terms in which the proposition is articulated. Thus, she can make a very tentative assessment of how probable she believes the proposition to

6. One can be Bayesian without being subjectivist. See Friedman, *supra* note 1, at 277 n.3 ("Bayes' theorem . . . does not apply only to subjectively assigned probabilities, and there is much more to subjective probability theory than the theorem. But the theorem is crucial, and the label is handy."). The terminology and varieties of Bayesianism are noted more fully in D.H. Kaye, *Introduction: What is Bayesianism?*, in PROBABILITY AND INFERENCE IN THE LAW OF EVIDENCE: THE USES AND LIMITS OF BAYESIANISM 1 (Peter Tillers & Eric D. Green eds., 1988), published in slightly different form as D.H. Kaye, *What is Bayesianism? A Guide for the Perplexed*, 28 JURIMETRICS J. 161 (1988).

7. A probabilist, as opposed to an analyst of evidence, may prefer to speak in terms of propositions rather than of evidence; one could refer to the proposition that evidence arises.

be.⁸ That assessment is, of course, subject to being altered radically by the receipt of new evidence.⁹

It seems plain that the prior odds that an observer attaches to a proposition—even in the face of virtual ignorance—may be very low, very high, or somewhere in the middle, depending on the nature of the proposition and the circumstances in which it is posed. A fact-finder would certainly be rational in assigning low prior odds to the proposition, “Yesterday it literally rained cats and dogs,” and high prior odds to the proposition, “All mountains that exist today will exist tomorrow.”

Even in the face of virtual ignorance, there is no reason for an observer to adopt a general rule of setting prior odds of 1:1 for a proposition. Such a rule would be arbitrary and wrongheaded, putting out of mind all useful information with which the observer addresses a new problem. It would assume that, simply because a proposition has been articulated, the proposition is exactly as likely to be true as false. There is no justification for such a conclusion.¹⁰

I do not mean to deny that there are *some* propositions for which an observer may quite reasonably assess prior odds of 1:1, or even that there are some for which prior odds much different from 1:1 would *not* be reasonable. Suppose the observer is shown a disk that is perfectly symmetrical, except that one side is blue and the other side is red, and asked to assess the odds that the disk, if flipped from a vertical position, will land red side up. She may very well assess these odds as even, for she may have no basis at all for believing that one side is more likely to land up than the other. But, as the

8. For example, suppose the observer is asked, “How probable is it that it will be sunny tomorrow in Kiribati?” Given the context, even an observer ignorant of geography might well suppose that Kiribati is a place somewhere on earth, and if she understands the language and has experience of the world she has an understanding of what “sunny” means and some information bearing on how probable it is to be sunny at a given time in an arbitrarily chosen spot somewhere in the world. Cf. David Kaye, *The Laws of Probability and the Law of the Land*, 47 U. CHI. L. REV. 34, 44 (1979) (acknowledging that “[o]ne can construct a few intractable cases in which the choice of a subjective probability is either impossible or must be totally arbitrary,” one being that “[a] person deprived of all knowledge of physics, chemistry, biology, and astronomy would be puzzled if asked the probability that there are living beings on a planet orbiting the star Sirius.”).

9. If, for example, she learns what the weather was like this time last year in Kiribati, or what the weather is like today, she may radically alter her assessment.

10. A counterargument might accept the points made above that odds other than 1:1 are rational once a specific proposition is posed, and that the posing of a specific proposition conveys some information, but assert that before the observer knows the specifics of the proposition the only rational odds would be 1:1. For example, suppose the observer is asked: “I’m going to pose a yes-or-no question. What are the odds that the answer will be yes?” But even here, I do not believe the only reasonable answer would be 1:1; the observer’s ignorance is not total, for her experience might suggest that she is more (or less) likely to be asked a yes question than a no question. Perhaps more significantly, I think the argument—which applies in any event only to yes-or-no questions—gives up the game. The completely general question is uninteresting; as soon as it is made more specific, the observer has reason to alter the odds.

rather contrived nature of this hypothetical suggests, this is a relatively narrow type of situation.¹¹

Whatever appeal prior even odds may retain disappears altogether when we consider partitions of the possibilities into more than two groupings. Suppose that instead of a disk it is a perfectly fair cubic die that is to be tossed. Would a rational observer assess the odds that Face 1 will land up as 1:1? Of course not; that would mean that it is equally probable that the die will land on Face 1 as that it will not land on Face 1—and that, of course, is inconsistent with the premise that the die, with six faces, is fair. As the range of possibilities is partitioned into more and more smaller and smaller pieces—eventually becoming a continuous distribution—the point becomes stronger still. Suppose we select one point on our disk and then roll the disk, east to west. Eventually it will drop. The probability that the selected point will then be *the easternmost point on the disk* is infinitesimally small.

One might respond that the assumption of prior even odds really is only a specialized application for the two-possibility case—blue side or red side up, for example—of a more general “principle of indifference,” that all possibilities should be assumed to have equal prior odds. But for at least two reasons this response does not work. First, it makes the prior odds of a given hypothetical proposition depend on how the competing hypotheses are defined. Suppose there are balls in an urn, some red, some powder blue, and the rest royal blue, but we do not know how many there are of each. What are the prior odds that the first ball selected at random will be red? If we define the competing possibilities as red versus blue, then this indifference principle would prescribe that the prior odds that the ball is red are 1:1—but if the possibilities are defined as red, powder blue, and royal blue, then it would prescribe prior odds of 1:2.¹² Second, in the poly-possibility case, just as in the two-possibility case, it is not usually sound to assign equal prior odds to each possibility. Suppose you throw a pin up in the air. Where will its tip land? Points directly under the spot where you released the pin are substantially more probable than distant points.

It is not surprising, therefore, that the problem of assessing prior odds in the face of virtual ignorance has engaged (and perplexed) mathematical philosophers, statisticians, and even legal scholars for years, especially in the

11. See Max Black, *Probability*, in 6 ENCYCLOPEDIA OF PHILOSOPHY 464, 474 (Paul Edwards ed., 1967) (“It is hard, of course, to imagine a case in which the reasoner is *wholly* ignorant of evidence favoring either P or not- P and is therefore required to assign to each alternative the probability $\frac{1}{2}$. Is the present reader of these lines of the male sex? The writer fancies that women read philosophical articles less often than men do; if so, the relevant reasons are asymmetrical and the principle fails to apply.”).

12. This illustration is adapted from Barry Loewer, *Probability Theory and Epistemology*, in 7 ROUTLEDGE ENCYCLOPEDIA OF PHILOSOPHY 705, 710 (Edward Craig ed., 1998) (concluding that for this reason this principle “is very sensitive to exactly how the hypotheses are described”).

context of continuous distributions or other multi-part partitions.¹³ In asserting that the rational adjudicative fact-finder will use even prior odds, Judge Posner pays no attention to this discourse. But is there something about the adjudicative context—symbolized by the icon of the level scales of justice¹⁴—that warrants insisting that the fact-finder should use even prior odds?

II. INDETERMINACY

As we shift to the legal context, note first the indeterminacy of a rule of prior even odds: Is the rule to be applied to the case as a whole, as Judge Posner suggests, or to each element of the claim or charge? Take a case in which the defendant is charged with having committed murder, defined as “the unlawful killing of a human being, with malice aforethought.” Is a juror (or other adjudicative fact-finder)¹⁵ supposed to begin with prior odds of 1:1 that the whole charge—“unlawful killing of a human being, with malice aforethought” is true? That would necessarily mean that the juror would begin, notwithstanding the defendant’s plea of innocence, by believing that *some* aspects of the charge are more likely true than not,¹⁶ which seems clearly unacceptable. And now suppose there is a lesser included offense, such as manslaughter, defined as “unlawful killing of a human being.” The only way that both priors could be 1:1 is if the jury is supposed to take as established to a certainty the elements included in the definition of murder but not those in manslaughter; if the jury begins with any uncertainty at all about aggravating elements, then the prior odds of murder must be less than the prior odds of manslaughter.

13. See, e.g., JOSE M. BERNARDO & ADRIAN F.M. SMITH, BAYESIAN THEORY 277, 298, 357-67 (1994); CARLTON M. CAVES, *Predicting Future Duration from Present Age: A Critical Assessment*, in CONTEMPORARY PHYSICS (forthcoming 2000; available at <<http://xxx.lanl.gov/abs/astro-ph/0001414>>); Edwin T. Jaynes, *Prior Probabilities*, in E.T. JAYNES, PAPERS ON PROBABILITY, STATISTICS AND STATISTICAL PHYSICS 116 (R. D. Rosenkrantz ed., 1983) (seeking to determine an objective method for determining prior probabilities in the face of ignorance); HAROLD JEFFREYS, THEORY OF PROBABILITY 33-34 (3d ed. 1961); JOHN MAYNARD KEYNES, A TREATISE ON PROBABILITY 372-83 (1921); JOHN L. POLLOCK, NOMIC PROBABILITY AND THE FOUNDATIONS OF INDUCTION 5-8 (1990); BERNARD ROBERTSON & G.A. VIGNAUX, INTERPRETING EVIDENCE: EVALUATING FORENSIC SCIENCE IN THE COURTROOM 25 & n.33 (1995); D.H. Kaye, *The Probability of an Ultimate Issue: The Strange Case of Paternity Testing*, 75 IOWA L. REV. 75, 94 n.94 (1989) (explaining that 0.5 cannot be a mathematically neutral assignment in paternity cases).

14. See, e.g., Sir Joshua Reynolds’ painting “Justice,” reproduced as the frontispiece to ROBERT M. COVER, OWEN FISS & JUDITH RESNIK, PROCEDURE (1988). See generally Dennis E. Curtis & Judith Resnik, *Images of Justice*, 96 YALE L.J. 1727 (1987).

15. For simplicity, I will speak from here on of jurors, but the argument applies equally to other adjudicative fact-finders.

16. If the juror began by assessing prior odds of 1:1 on each element of the charge, she would necessarily assess prior odds of less than 1:1 on the whole charge, unless the elements were perfectly congruent with one another, in which case it would not make much sense to speak of separate elements.

The solution might appear to be that the juror should treat each element, not the charge as a whole, as having prior even odds. But that does not work, because the number of elements into which a charge or claim is divided is arbitrary. How many elements are there in the definition of murder given above? Two—unlawful killing of a human being and malice aforethought? Five—unlawful, killing, human being, malice, and aforethought? Some number between two and five? Some number greater than five (if one or more of the five has a multi-part definition)? Pick your favorite.

I do not believe an advocate of prior even odds could develop a satisfactory solution to this problem. But for the most part I will put it aside by treating each hypothetical case as if it involved the proof of a single, indivisible proposition.

III. THE PRESUMPTION OF INNOCENCE

If a casual observer walks into the middle of a criminal trial in the United States, it will not usually take her long to pick out the accused. And if, without learning more about the case, she assessed the odds that the accused is guilty as being substantially greater than 1:1, she would be acting perfectly rationally: most criminal defendants who are brought to trial in America are in fact guilty.¹⁷ So do prior odds of 1:1 give a break to the accused? Not at all. What is rational for a casual observer is not permissible for a juror.

The accused has a constitutional right to have the fact-finder apply a presumption of innocence to his case.¹⁸ An assumption of prior odds of 1:1 might be thought at first glance to be compatible with this presumption because the standard of persuasion—beyond a reasonable doubt—clearly requires that the jury acquit unless it concludes, in light of all the evidence, that the odds of guilt are much higher.¹⁹ Thus, if neither side presented any evi-

17. See, e.g., Samuel R. Gross, *Loss of Innocence: Eyewitness Identification and Proof of Guilt*, 16 J. LEGAL STUD. 395 (1987) (contending that the apparently low rate of erroneous convictions based on mistaken identifications results largely from police and prosecutorial processes that—relying on other, independent information, which may not be admissible evidence—usually weed out misidentifications before trial).

18. See, e.g., *Estelle v. Williams*, 425 U.S. 501, 503 (1976) (holding that the presumption of innocence “is a basic component of a fair trial under our system of criminal justice”); *Coffin v. United States*, 156 U.S. 432, 453 (1895) (“The principle that there is a presumption of innocence in favor of the accused is the undoubted law, axiomatic and elementary, and its enforcement lies at the foundation of the administration of our criminal law.”); cf. *Kentucky v. Whorton*, 441 U.S. 786, 789 (1979) (holding over dissent that failure to give instruction on presumption of innocence is not in itself constitutional error if the totality of the circumstances does not indicate a danger that the jury will convict on something other than the evidence).

19. See *Schlup v. Delo*, 513 U.S. 298, 325 (1995) (quoting with apparent approval the statement by Thomas Starkie, a leading treatise writer of the early nineteenth century, that “it is better that ninety-nine . . . offenders shall escape than that one innocent man be condemned”). Of course, the classic statement is that of 4 WILLIAM BLACKSTONE, COMMENTARIES 352 (photo reprint 1966) (1769) (“[I]t is better that ten guilty persons escape, than that one innocent suffer.”). See generally

dence at all, but the case was nevertheless submitted to the jury, which adhered to instructions on the standard of persuasion and to Judge Posner's rule of an even-odds prior, the jury would have to acquit.

But the presumption of innocence means more than that the prosecution cannot win without presenting some evidence. A basic aspect of the presumption is that the jury may not treat the fact that the defendant is on trial as itself an indication of guilt.²⁰ This rule is violated by the assumption of prior even odds.

To understand this, assume the venue has changed temporarily from the courtroom to a nearby street. Observer, seeing Passerby on the street and knowing nothing about him but what she notices as he walks by in an ordinary fashion, probably would not articulate to herself the issue of whether Passerby has recently committed a violent assault. But if she knew that he *had* done so, she might adjust her conduct accordingly, perhaps giving him wide berth on the sidewalk, or at least giving him a dirty look. Assuming she does not engage in some unusual conduct of this sort, she is probably acting on an implicit perception that the odds that Passerby has committed such an assault are very low. And she would certainly be reasonable in this assessment: most people do not violently assault their neighbors on or about a given date, and Observer has no information concerning Passerby that would increase the odds.

Moving back into the courtroom now, Observer becomes Juror and Passerby becomes Defendant. If Juror sets even, rather than very low, prior odds on the proposition that Defendant violently assaulted his neighbor on or about last June 25, what has caused the change? It can only be that Juror has

Alexander Volokh, *N Guilty Men*, 146 U. PA. L. REV. 173 (1997) (reciting numerous such statements). But getting from statements like this to the higher odds required for conviction is not so obvious. See Michael L. DeKay, *The Difference Between Blackstone-like Error Ratios and Probabilistic Standards of Proof*, 21 L. & SOC. INQUIRY 95 (1996).

20. See, e.g., *Taylor v. Kentucky*, 436 U.S. 478, 485 (1978) ("[O]ne accused of a crime is entitled to have his guilt or innocence determined solely on the basis of the evidence introduced at trial, and not on grounds of official suspicion, indictment, continued custody, or other circumstances not adduced as proof at trial. See, e.g., *Estelle v. Williams*, 425 U.S. 501 (1976). And it long has been recognized that an instruction on the presumption is one way of impressing upon the jury the importance of that right."). A well-known manual of federal jury instructions puts the matter this way:

I instruct you that you must presume the defendant to be innocent of the crime charged. Thus the defendant, although accused of a crime in the indictment, begins the trial with a "clean slate"—with no evidence against him. The indictment, as you already know, is not evidence of any kind. The law permits nothing but legal evidence presented before the jury in court to be considered in support of any charge against the defendant.

1A KEVIN F. O'MALLEY, JAY E. GRENING & HON. WILLIAM C. LEE, *FEDERAL JURY PRACTICE AND INSTRUCTIONS: CRIMINAL* § 12.10 (5th ed., 2000) (some variants in brackets omitted); *accord*, 1 JOSEPHINE R. POTUTO, STEPHEN A. SALTZBURG & HARVEY S. PERLMAN, *FEDERAL CRIMINAL JURY INSTRUCTIONS* § 1.03 (2d ed., 1993) ("An (indictment) (information) is only the government's claim that a crime was committed by the defendant; it is not evidence that a crime was committed or that the defendant committed it.")

taken into account the process by which the issue of Defendant's guilt on that charge has come to be presented in court. She believes that this process suggests guilt because Defendant is more likely to have been brought to trial if he is guilty than if he is not guilty. This reasoning may be plausible enough, but it is just what a conscientious juror is prohibited from using.

What, then, does the presumption of innocence mean? Or put another way, how should we conceptualize the prior odds that a conscientious juror should use? This is a more difficult problem than might appear.²¹ It may seem tempting to prescribe that the juror should use prior odds of 0. But that does not work. Prior odds of 0 mean that the juror is certain that the proposition at issue is false, and that no further evidence can change her mind.²² An adjudicative system cannot operate that way; if a juror manifested such obduracy during voir dire, she would be struck from the jury for cause.

The answer, I believe, is that Juror should assess prior odds just as if she were implicitly assessing the odds outside the adjudicative system—that is, if she were back to being Observer and Defendant back to being Passerby. Such odds would be greater than 0 but very small.

By speaking of *implicit* setting of odds, I am referring to a lurking complexity. If we want Juror to think about the prior odds of a proposition, we must pose the proposition. And doing so, even outside the adjudicative context, can be quite suggestive. Suppose, for example, Observer's friend asked her, "How probable do you think it is that this passerby has violently assaulted his neighbor recently?" Observer might well infer that the question would not be asked unless the proposition is at least plausible, and so assess the odds as higher than she did implicitly. So it is not quite accurate to tell Juror to assess the odds just as if the question were posed to her out of the adjudicative context.

We could improve the comparison conceptually by imagining that the question of interest were buried in a long questionnaire, thus eliminating any substantial suggestibility: "How probable is it that the building across the street will collapse tomorrow? the day after? the day after that? . . . five

21. See, e.g., Friedman, *supra* note 1, at 285 (criticizing tests under which "the presumption means that, before any evidence is presented, the fact-finder should treat the defendant as no more likely than anybody else, or, in a variant, as equally likely as anyone else, to be guilty").

22. This is a point made in *State v. Skipper*, 637 A.2d 1101, 1108 (Conn. 1994).

If we assume that the presumption of innocence standard would require the prior probability of guilt to be zero, the probability of paternity in a criminal case would always be zero because Bayes' Theorem requires the paternity index to be multiplied by a positive prior probability in order to have any utility.

Id. The *Skipper* court drew the incorrect inference that, because Bayesian analysis requires the assignment of a positive prior probability, "Bayes' Theorem can only work if the presumption of innocence disappears from consideration." *Id.* at 1108 (quoting Randolph N. Jonakait, *When Blood Is Their Argument: Probabilities in Criminal Cases, Genetic Markers, and, Once Again, Bayes' Theorem* 1983 U. ILL. L. REV. 369, 408 (1983)).

years from now? that this passerby violently assaulted his neighbor on or about January 1st? January 2nd? . . . yesterday? that the President will fly to Antarctica tomorrow? to Fiji? . . . ?"²³

I am not, of course, suggesting that a thought experiment of this sort actually be posed to the jury. Perhaps this problem of trying to pose the issue without suggestiveness is both too intractable and too insignificant to demand a complete solution. I believe an instruction of the following type should be sufficient:

Members of the jury, you may not treat the bringing of the accusation, or the prosecutor's opening statement, as evidence that the accused committed the crime charged. Accordingly, now, before the trial has begun, when you have no information on the matter apart from your general knowledge of the world, if you were to ask yourself how probable it is that the accused violently assaulted his neighbor last June 25, you would, to be reasonable, have to assess that probability as being very low. This is the point from which you must begin consideration of the evidence in this case.²⁴

One might well ask how realistic it is to expect that jurors will begin consideration with very low prior odds of guilt. I suspect that most jurors do start with relatively high odds, based largely on the anticipation that the prosecutor will present powerful evidence of guilt, for if she could not she would not press the case. In the end, such a juror might find her anticipation fulfilled. Then again, she might not. If she does not, she would assess the odds of guilt lower at the close of evidence than she did at the start. There may not be much harm in this approach, *so long as* at the close of evidence the juror gives no residual weight to the fact that those responsible—the prosecutor and, perhaps, a grand jury—decided to bring the charge. Engaging in such a course of reasoning is not a simple matter, but a juror asked to defend her conduct in light of the presumption of innocence might well say:

Yes, I knew I was supposed to start with very low prior odds and then adjust them in light of the evidence, and all I was doing was discounting for the evidence in advance, anticipating how strong I thought the evidence would be and what odds I would set at the end if I started where I was supposed to. When all the evidence was in, I had nothing more to anticipate, so I ended up right where I should have. I didn't give any credence to the idea that the prosecutor might know something she never told us.

Note that even under this line of reasoning there is no special role for prior odds of 1:1. Indeed, the juror may even assess prior odds considerably higher than 1:1 in anticipation of prosecution evidence. But in making that anticipation, this juror is predicting where she would likely end the case if

23. For similar suggestions, see Friedman, *supra* note 1, at 286 n.22, and my comment in Ronald J. Allen, David J. Balding, Peter Donnelly, Richard Friedman, David H. Kaye, Lewis Henry LaRue, Roger C. Park, Bernard Robertson & Alexander Stein, *Probability and Proof in State v. Skipper: An Internet Exchange*, 35 JURIMETRICS J. 277, 296 (1995) [hereinafter *Skipper Exchange*].

24. This instruction is drawn, with substantial alterations, from one I presented in *Skipper Exchange*, *supra* note 23, at 296.

she began it by using the very low prior odds that the presumption of innocence demands.

IV. BIZARRE RESULTS

If a juror really assumed prior even odds in adjudication, bizarre results would often follow.

A. *Satisfying the Criminal Standard of Persuasion*

It is important that the presumption of innocence and the standard of persuasion not be confused. The presumption of innocence tells the jurors where they should begin consideration of the case. The standard of persuasion tells the jurors where they must end if they are to convict. The presumption, I have said, demands very low prior odds. The standard of persuasion demands very high posterior odds for a conviction. Powerful evidence is necessary to carry a juror from the first point to the second. By prescribing the wrong starting point, the assumption of prior even odds would make satisfaction of the standard of persuasion far too easy.²⁵

Suppose, for example, that a witness, with metronomic regularity, answers 999 of 1000 yes-or-no questions correctly the first time they are asked.²⁶ She then gives an affirmative answer to the question, "Did Jack kidnap you and take you on a spaceship to the moon?" Now remember the simple formulation of Bayes' Theorem, that the posterior odds equal the prior odds times the likelihood ratio. The likelihood ratio of any answer she gives with respect to the proposition that she asserts is 999:1, no matter what the question. Given the hypothesis of prior odds of 1:1, the posterior odds simply equal the likelihood ratio—yielding odds of 999:1 that Jack did indeed kidnap her and take her on a spaceship to the moon. But those odds are clearly too high by several orders of magnitude;²⁷ a fact-finder *should* conclude that this is probably one of those occasional cases in which the witness has gotten it wrong, rather than that she has been to the moon. Plainly, the

25. To see the difference, note that some jurisdictions have had affirmative defenses on which the defendant has the burden of production—meaning that the presumption is *against* the defendant on that point—but if the defendant satisfies that burden, the prosecution has the burden of disproving the defense beyond a reasonable doubt. *See, e.g.*, United States v. Lewellyn, 723 F.2d 615, 617 (8th Cir. 1983) ("A [criminal] defendant is presumed sane, but the introduction of evidence of insanity dispels the presumption and subjects the prosecution to the burden of proving sanity beyond a reasonable doubt."); Byrd v. State, 297 So. 2d 22, 23 (Fla. 1974) ("The rule is well settled in Florida that all men are presumed sane, but the presumption vanishes when there is testimony of insanity sufficient to present a reasonable doubt as to the sanity of a defendant.").

26. Once she answers a given question, she will respond consistently to other questions.

27. Similarly, if the answer were negative, posterior odds of 999:1 would be far too low.

culprit in the reasoning that leads to the 999:1 odds is the assumption of prior odds of 1:1.²⁸

Additional light may be shed on the point by a somewhat different type of problem. An assault was admittedly committed at night in a large city. Though the police have been diligent in their investigation, there is no evidence pointing to the identity of the assailant except that he left footprints in the mud at the site of the crime. These footprints are of a very large shoe size, matching the shoes of only one person in 1000 in the general population, and they also match the shoes of the accused. If the accused is the assailant and left the shoe prints, they would almost certainly be of this size, but if the accused is not the assailant, then the probability that footprints left by the assailant would match these would be only about 1 in 1000. Thus, the likelihood ratio of this evidence would be essentially 1000:1. Under the assumption of prior even odds, the posterior odds would also be 1000 to 1—which seems to be sufficient for conviction.²⁹ But plainly this is absurd. In a large city there may be hundreds or even thousands of others with that shoe size, and we have no information suggesting that the accused is much more likely than the others to be the source of the print.

An advocate of even prior odds might respond to these hypotheticals by using the doctrine of missing evidence. The argument would be that if the shoe evidence were all that the prosecution had to offer, the jury should acquit, because if the defendant were guilty more evidence of guilt would be expected. It is true that the absence of evidence is often in itself significant, when the missing evidence is of the type that one would expect a given party to produce if the facts favored that party.³⁰ But not all cases are of that sort.

28. It does not do to say that the juror should reassess the odds downwards as soon as she learns the implausible nature of the charges. In addition to arguments already offered, see note 10 *supra* and accompanying text (responding to the idea that an observer might begin as a rule with prior even odds and then adjust them on learning the precise nature of the question), it is important to recognize that if information as to the nature of the charge causes the juror to lower the odds, that presupposes that information indicating that the charges were different would cause the juror to *raise* the odds. In other words, before hearing evidence, the juror on this hypothetical would have to begin some cases, before the presentation of evidence, with odds *greater than* 1:1.

29. See note 19 *supra* and accompanying text (99:1 ratio endorsed by Supreme Court in *Schlup*). The equation of posterior odds and likelihood ratio in contexts such as this is a form of what courts and scholars have called the “prosecutor’s fallacy.” See, e.g., United States v. Chischilly, 30 F.3d 1144, 1157 (9th Cir. 1994), cert. denied, 513 U.S. 1132 (1995); David J. Balding & Peter Donnelly, *The Prosecutor’s Fallacy and DNA Evidence*, CRIM. L. REV. 711, 711 (1994).

30. See, e.g., United States v. Pitts, 918 F.2d 197, 199 (D.C. Cir. 1990).

There are some persons . . . who potentially have so much to offer that one would expect them to take the stand. If such a person does not appear and one of the parties has some special ability to produce him, the law permits the jury to draw an inference—namely, that the missing witness would have given testimony damaging to that party.

Id.; Richard D. Friedman, *Dealing With Evidentiary Deficiency*, 18 CARDOZO L. REV. 1961, 1963-64 (1997) (explaining missing evidence inference); Dale A. Nance, *Missing Evidence*, 13 CARDOZO L. REV. 831, 848-49 (1991).

Sometimes, no matter how diligent the prosecution is, there may just be very little detectable evidence bearing on a proposition, such as the identity of the assailant in the shoe hypothetical.³¹

Of course, the jurors *would* expect the prosecutor to produce more evidence than the shoe print evidence in this hypothetical, and presumably if she did not do so, the jurors would acquit. But the reason is not hard to discern: a conscientious juror would start with prior odds not of 1:1 but of 1: X , where X is a large number perhaps on the order of the entire population of people who might have committed the crime. To justify conviction, the juror would have to reach high posterior odds, say 99:1. That would mean that for the shoe print evidence to justify conviction by itself the likelihood ratio would have to be at least $99X$. In other words, the shoe print size would have to be so extremely rare that the juror could conclude with great confidence that only one person in the relevant population had it. In ordinary circumstances, shoe print evidence would not satisfy this condition. DNA or fingerprint evidence may. If there is no single piece of evidence with overwhelming probative value, the prosecution will have to satisfy its burden of proof with an accumulation of evidence.³²

B. *Satisfying the Civil Standard of Persuasion*

An assumption of even odds would lead to similarly bizarre results in civil cases, in which the usual standard of persuasion is “preponderance of the evidence.” That standard is usually interpreted to mean that the plaintiff must only satisfy the jury that the facts more likely than not favor him—that is, that the probability in his favor is just a shade over 0.5.³³ But if the start-

31. Thus, there are many unsolved crimes, in which prosecutions are never brought because a reasonable jury could not be satisfied beyond a reasonable doubt of the guilt of any given defendant. In some circumstances, a successful civil claim may be possible even in the face of thin evidence of identity. See David Kaye, *The Limits of the Preponderance of the Evidence Standard: Justifiably Naked Statistical Evidence and Multiple Causation*, AM. B. FOUND. RES. J. 487 (1982).

32. Take a simple case. Defendant is arrested on a crowded street, a few minutes after a robbery and near the place where it occurred, when Victim and Eyewitness both tell police, “There’s the man!” The fact that Defendant is one of the relatively few people in the world who was in the vicinity of the robbery shortly after it occurred raises the prior odds from infinitesimally small to small but more substantial. The fact that Victim and Eyewitness both identified Defendant raises the odds much further, perhaps beyond reasonable doubt. If Defendant was the perpetrator, there would be a substantial probability (nowhere near certainty) that Victim and Eyewitness would both identify him; if Defendant was *not* the perpetrator, the probability that they would both pick him—among all other people in the crowd—as the perpetrator would usually appear very small.

33. See *Ethyl Corp. v. Environmental Protection Agency*, 541 F.2d 1, 28 n.58 (D.C. Cir. 1976) (en banc) (“[T]he standard of ordinary civil litigation, a preponderance of the evidence, demands only 51% certainty.”), cert. denied, 426 U.S. 941 (1976); *United States v. Shonubi*, 895 F. Supp. 460, 471 (E.D.N.Y. 1995) (Weinstein, J.) (“Judges—who deal with burdens of proof on a daily basis—are inclined to think of those burdens in probabilistic terms. A survey of judges in the Eastern District of New York found general agreement that ‘a preponderance of the evidence’ translates into 50+ percent probability.”).

ing point is even odds, or a probability of 0.5, the plaintiff is practically home even before presenting evidence, and a bare puff of wind will be enough to carry him there. Even if, as I suspect, the preponderance standard really means something more substantial, extraordinarily weak evidence would suffice.

To understand this, alter the nighttime assault case. Now the victim has brought a civil suit against the defendant, who is white. Despite energetic attempts to investigate, the only evidence of the assailant's identity is the victim's credible testimony, based on a view of his forearm, that he was white, together with proof that 60% of the people in the region, spread evenly throughout it, are white. Given that the victim had an opportunity to see the forearm, this is evidence that we would expect if the defendant was the assailant, but it would be only about 60% probable if someone else was. A jury would therefore be justified in giving this evidence a likelihood ratio of 100/60 or 5/3. Assuming prior odds of 1:1, that would mean that the posterior odds would also be 5:3, or a posterior probability over 60%, presumably enough to justify a verdict for the plaintiff. Absurd? Of course. But if instead we assume low prior odds the absurdity disappears.

C. *Multiple Hypotheses*

Recall the problems of the die and the rolling disk. They emphasize that not all problems are binary, with only two choices. And so it is in litigation. Suppose that patient *P* claims that she purchased and used a single bottle of pills and that it caused her serious injury. She then sues manufacturers *A* and *B*, claiming in the alternative that one or the other made the pills in question.³⁴ If the jury is supposed to begin with odds of 1:1 that *A* is liable, and 1:1 that *B* is liable, that would seem to exhaust the possibilities; where is there room for someone else to be liable, or for no one else to be liable at all? And if *P* now sues manufacturer *C* as well, I am utterly at a loss to understand how the jury is supposed to fit three 50% prior probabilities into a 100% total.

CONCLUSION

It is easy enough to understand the allure of the even-odds prior, and Judge Posner is not the first to be drawn to it.³⁵ The icon of equal scales

34. Modern pleading rules allow inconsistent pleading. *See, e.g.*, FED. R. CIV. P. 8(e)(2).

35. It is a common assumption made by experts presenting scientific evidence in paternity cases. *See, e.g.*, State v. Skipper, 637 A.2d 1101, 1105 (Conn. 1994); Kaye, *supra* note 13, at 83-89; ROBERTSON & VIGNAUX, *supra* note 13, at 25. In those cases, though, the prior assumes all the other evidence in the case, which may include testimony of the "he says, she says" variety. In such a case, a juror who, apart from the scientific evidence, reaches a probability of exactly .5 may be acting perfectly rationally. But there is nothing magical about that point; it is just one point on the

suggests symmetry and balance; presumably the scales will tilt one way or the other at the end of the case, but fairness may seem to demand that at the beginning they be on the same level. There is a simplicity, too, about even odds; given that a fact-finder's prior odds cannot be 0 or infinite, even odds are the only easily articulable starting point. But the world is not always as simple as our metaphors suggest; perhaps, that is why metaphors have value. Difficult though the task may be, I do think it is possible to articulate in Bayesian terms how the jury should determine where to start.

I hope Judge Posner enjoys a long stay in Evidenceland. Productive tilling of the Evidence soil with Bayesian tools, however, demands a premise that the fact-finder begins at an acceptable starting point. Even odds are not such a point.

continuum from 0 to 1. And in any event, it is no business of the expert to assume such a prior probability.