Neglect the Base Rate:
It's the Law!

Christoph Engel

Neglect the Base Rate: It's the Law!

Christoph Engel

December 2012

Neglect the Base Rate: It's the Law!

Christoph Engel

Abstract

If accurate prediction is the goal, and if information about the unconditional probability of the predicted event is available, a strong case can be made for using this information, i.e. for a Bayesian approach to inference. Not so rarely, the law calls for accurate prediction, e.g. if a bailing decision hinges on an estimate of recidivism risk. Yet for other questions of law, and for the law of evidence in particular, accuracy is not the exclusive goal. Substantive law determines who should bear the risk that doubt cannot be removed. These rules decide whether several individuals, or several acts for that matter, shall be treated as members of a class. Applying Bayes' rule also implicitly treats the person or the action in question as a member of a class. If in conflict, the normative decision of substantive law overrides Bayes' rule, and forces judges and administrators to neglect the base rate.

Keywords: base rate, Bayes' rule, acting in concert, product share liability, alternative causes

1. Introduction

Base rate neglect is a classic in the judgment and decision making literature.[1] It is typically considered as a fallacy[2] since it violates Bayes' theorem. The decision maker is exclusively (or overly) attentive to new information, and fails to integrate it with what she knows (or should have learned) about the abstract probability of the event in question.

This can lead to actions hardly anyone would find appealing. Take the gynecologist who has sent a woman to a routine mammography. The test turns out positive. With trepidation, the woman asks the doctor: what is the probability that I actually have breast cancer? The doctor may know for a fact that the probability of a woman who has breast cancer to test positive is around 90%. If this is all he tells his poor patient, she will endure terrible anxiety on very misleading grounds. For the overall risk of a woman to attract the disease is in the order of 1%. Moreover mammographies are overly sensitive. In about 9% of all cases, women test positive although they actually are not ill. If the doctor knows Bayes' rule, he will calculate that, despite the test being positive, the probability of this woman being ill is only 9.17%.[3] Less than one out of ten women

with a positive test actually has cancer. Who would not want this woman to know, and who would not want her doctor to understand? Actually, before they are trained, the large majority of gynecologists heavily overestimate the risk. Almost half of them say it is 90%.[4]

Not so rarely, the law faces comparable problems. Sometimes, the application of a legal rule rests on prognostic judgment: is a sexual offender likely to recidivate, and may therefore be held in custody?[5] Does a psychiatry patient pose a threat of violence to third parties and may

---

[1] PAUL SLOVIC & SARAH LICHTENSTEIN, *Comparison of Bayesian and Regression Approaches to the Study of Information Processing in Judgment*, 6 Organizational Behavior and Human Performance (1971);DANIEL KAHNEMAN & AMOS TVERSKY, *On the Psychology of Prediction*, 80 Psychological Review (1973);DON LYON & PAUL SLOVIC, *Dominance of Accuracy Information and Neglect of Base Rates in Probability Estimation*, 40 Acta Psychologica (1976);MAYA BAR-HILLEL, *The Base-Rate Fallacy in Probability Judgments*, 44 Acta Psychologica (1980);BARUCH FISCHHOFF & RUTH BEYTH-MAROM, *Hypothesis Evaluation from a Bayesian Perspective*, 90 Psychological Review (1983).

[2] But see JONATHAN KOEHLER, *The Base Rate Fallacy Reconsidered. Descriptive, Normative, and Methodological Challenges*, 19 Behavioral and Brain Sciences (1996).

[3] Bayes rule is, in the notation I will later use for the legal application, where *g* stands for guilty, and *e* stands for a piece of evidence,

$$p(g|e) = \frac{p(e|g) * p(g)}{p(e)}$$

In the concrete case, $p(g)$ is the anterior probability of the woman having cancer. $p(e)$ is not given, but must be reconstructed as $p(e|g) * p(g) + p(e|\neg g) * p(\neg g)$. $p(\neg g) = .9$. It follows from the fact that the two events (woman has cancer; woman does not have cancer) are mutually exclusive.

[4] GERD GIGERENZER, et al., *Helping Doctors and Patients Make Sense of Health Statistics*, 8 Psychological Science in the Public Interest (2007).

[5] WILLIAM M. GROVE, et al., *Clinical versus Mechanical Prediction. A Meta-analysis*, 12 Psychological Assessment (2000);R.KARL HANSON & KELLY E. MORTON-BOURGON, *The Accuracy of Recidivism Risk Assessments for Sexual Offenders. A Meta-analysis of 118 Prediction Studies*, 21 Psychological Assessment (2009);RICHARD WOLLERT, *Low Base Rates Limit Expert Certainty when Current Actuarials are Used to Identify Sexually Violent Predators. An Application of Bayes's Theorem*, 12 Psychology, Public Policy, and Law (2006);RICHARD WOLLERT, *Poor Diagnostic Reliability, the Null-Bayes Logic Model, and Their Implications for Sexually Violent Predator Evaluations*, 13 Psychology, Public Policy,

therefore be detained?[6] How likely is a proposed regulatory intervention to mitigate a risk to the environment and may therefore interfere with freedom or property?[7] In such cases, it is hard to deny that the normative goal is prognostic accuracy, and that accuracy is (potentially dramatically) improved if the base rate of the respective event is taken into account.

Another area of application is the probative value of evidence presented in court. If a drug dog has alerted, the court should take into account how often that dog got it right in the past.[8] DNA evidence is also inherently Bayesian. It rests on generic knowledge about the very low probability of a random match between the identifying DNA particles and tissue that actually is not from the individual, i.e. on a base rate.[9] If a court wants to rely on this evidence, it should integrate it with an estimate of the much higher risk of a lab mistake,[10] and of the likelihood that the evidence has been doctored,[11] using Bayes' rule.[12]

A third area of law where a Bayesian approach is easy to justify is the quantification of damage, provided causation and, if required, intent or negligence have been established. If this damage consists of lost income or profit, the court must engage in a prognostic assessment. Provided generic knowledge was sufficiently reliable, courts have allowed plaintiffs to rely on a base rate for that purpose.[13]

If they interpret the probability of the claim being well founded as established by the existing evidence as a base rate, lawyers have a controlled technique for assessing the relevance of a new piece of evidence.[14] If investigators are trained to take base rates into account, this reduces the risk of confirmatory search.[15] Lawyers negotiating, say over a plea bargain, mitigate the

---

and Law (2007); ANDREAS MOKROS, et al., *Assessment of Risk for Violent Recidivism Through Multivariate Bayesian Classification*, 16 Psychology, Public Policy, and Law (2010).

[6]   NICHOLAS SCURICH & RICHARD S. JOHN, *Prescriptive Approaches to Communicating the Risk of Violence in Actuarial Risk Assessment*, 18 Psychology, Public Policy, and the Law (2012).

[7]   DAVID E. ADELMAN, *Scientific Activism and Restraint. The Interplay of Statistics, Judgment, and Procedure in Environmental Law*, 79 Notre Dame Law Review (2003).

[8]   MARK E. SMITH, *Going to the Dogs. Evaluating the Proper Standard for Narcotic Detector Dog Searches of Private Residences*, 46 Houston Law Review (2009).

[9]   For an easy to understand introduction into the basics see
      http://learn.genetics.utah.edu/content/labs/gel/forensics.

[10] JONATHAN J. KOEHLER, *Why DNA Likelihood Ratios Should Account for Error (Even when a National Research Council Report Says They Should Not)*, 37 Jurimetrics (1996); BOAZ SANGERO & MORDECHAI HALPERT, *Why a Conviction Should Not Be Based on a Single Piece of Evidence: A Proposal for Reform*, 48 Jurimetrics (2007).

[11] This is disturbingly easy to do, DAN FRUMKIN, et al., *Authentication of Forensic DNA Samples*, 4 Forensic Science International: Genetics (2010).

[12] DALE A. NANCE, *Naturalized Epistemology and the Critique of Evidence Theory*, 87 Virginia Law Review (2001) ; also see JON TODD POWELL, *How to Tell the Truth With Statistics. A New Statistical Approach to Analyzing the Bendectin Epidemiological Data in the Aftermath of Daubert v. Merrell Dow Pharmaceuticals*, 31 Houston Law Review (1994).

[13] JONATHAN J. KOEHLER, *When Do Courts Think Base Rates Statistics Are Relevant?*, 42 Jurimetrics, 398 f. (2001) reports one successful case: *Contemporary Mission, Inc. v. Famous Music Corp.*, 557 F.2d 918, 926 f. (2d. Cir. 1977), and one failed case: *Wilson v. B.F.Goodrich Co.*, 627 P.2d 1280 (Or.Ct.App. 1981), *aff'd* 642 P.2d 644 (Or. 1982).

[14] LAWRENCE H. TRIBE, *Trial by Mathematics. Precision and Ritual in the Legal Process*, Harvard Law Review, 1351 (1971).; also see ADELMAN, *supra* note 7, at 504 and passim.

[15] MICHAEL J. SAKS & MICHAEL RISINGER, *Baserates, the Presumption of Guilt, Admissibility Rulings, and Erroneous Convictions*, 1051-1063 Michigan State Law Review (2003).

risk of overoptimism if they are attentive to the base rate of winning, were the case to go to trial.[16]

This seemingly makes the advice compelling: Train our Jurors![17] and, *a fortiori,* our professional judges. All the more so since tests with mock jurors show that they heavily violate Bayes' theorem.[18] If one just reads the equation,[19] training may seem mission impossible. Conditional probabilities are difficult to understand, and most lawyers do not excel in math. Yet there are straightforward decision aids, like the representation of the problem as a decision tree, that massively improve performance.[20]

Now for decades it has been debated whether lawyers in general, and courts dealing with evidentiary issues in particular, should become Bayesians.[21] Critics have said: it is immoral to condemn a defendant although one attaches positive probability to her being innocent.[22] This leads to a dehumanization of justice.[23] The norms of probability theory, and of Bayes' theorem in particular, are at variance with judge and jury intuition, i.e. with the way how actual triers find their decisions.[24] Overtly statistical argument violates the standard of proof: beyond a reasonable doubt requires actual belief.[25] Verdict accuracy, which is served by a Bayesian approach, is not the only normative goal.[26] Procedural justice carries independent weight.

---

[16]   Id.

[17]   JONATHAN KOEHLER, *Train our Jurors*, *in* Heuristics and the Law 315, (Gerd Gigerenzer & Christoph Engel eds., 2006).

[18]   DALE A. NANCE & SCOTT B. MORRIS, *Juror Understanding of DNA Evidence. An Empirical Assessment of Presentation Formats for Trace Evidence with a Relatively Small Random-Match Probability*, 34 Journal of Legal Studies (2005) ; also see MICHAEL J. SAKS & ROBERT F. KIDD, *Human Information Processing and Adjudication. Trial by Heuristics*, 15 Law and Society Review (1980).

[19]   See note 3.

[20]   GERD GIGERENZER & ULRICH HOFFRAGE, *How to Improve Bayesian Reasoning Without Instruction. Frequency Formats*, 102 Psychological Review (1995).

[21]   Good overviews of the debate are provided by NANCE, *supra* note 12; JONATHAN J. KOEHLER & DANIEL N. SHAVIRO, *Veridical Verdicts. Increasing Verdict Accuracy through the Use of Overtly Probabilistic Evidence and Methods*, 75 Cornell Law Review (1990); RONALD J. ALLEN & BRIAN LEITER, *Naturalized Epistemology and the Law of Evidence*, 87 Virginia Law Review (2001); ROGER C. PARK & MICHAEL J. SAKS, *Evidence Scholarship Reconsidered. Results of the Interdisciplinary Turn*, 46 Boston College Law Review, 988-995 (2006). A comprehensive treatise is being prepared by Mark Schweizer, who forcefully advocates a Bayesian approach to the law of evidence.

[22]   TRIBE, *supra* note 14, at 1372; DAVID T. WASSERMAN, *The Morality of Statistical Proof and the Risk of Mistaken Liability*, 13 Cardozo Law Review (1991).

[23]   TRIBE, *supra* note 14, at 1375.; but see DANIEL N. SHAVIRO, *Statistical-Probability Evidence and the Appearance of Justice*, 103 Harvard Law Review (1989).

[24]   ALLEN & LEITER; LEA BRILMAYER & LEWIS A. KORNHAUSER, *Quantitative Methods and Legal Decisions*, 46 University of Chicago Law Review (1978) ; but see NANCE, *supra* note 12; KOEHLER & SHAVIRO, *supra* note 21; DAVID FAUST, *Holistic Thinking is not the Whole Story. Alternative or Adjunct Approaches for Increasing the Accuracy of Legal Evaluations*, 10 Assessment (2003).

[25]   For an extensive survey of that debate see ANDREA ROTH, *Safety in Numbers? Deciding When DNA Alone is Enough to Convict*, 85 New York University Law Review (2010). Actually, I share that concern, but do not want to pursue it in this note, see CHRISTOPH ENGEL, *Preponderance of the Evidence Versus Intime Conviction. A Behavioral Perspective on a Conflict Between American and Continental European Law*, 33 Vermont Law Review (2009).

[26]   For a discussion of alternative goals see RONALD J. ALLEN, *A Reconceptualization of Civil Trials*, 66 Boston University Law Review (1986); RONALD J. ALLEN, *On the Significance of Batting Averages and Strikeout Totals. A Clarification of the Naked Statistical Evidence Debate, the Meaning of Evidence, and the Requirement of Proof beyond a Reasonable Doubt*, 65 Tulane Law Review (1990).

Overtly probabilistic methods may clash with the justice sentiments of the parties, and of the public at large.[27]

Other concerns originate in the limitations inherent in judicial practice. Exceptions like DNA evidence notwithstanding, normally courts will at best have probability estimates individually generated by the triers. While subjective probability theory shows how to align them with Bayes rule,[28] not all feel comfortable with the translation.[29] More mundanely, critics are afraid of the "costs of precision".[30] Attention may be distracted from facts that may not easily be quantified, but that are normatively no less relevant.[31] Statistics may be abused to hide, e.g. racial, bias.[32] Impressive posterior probabilities may result from an inappropriate definition of the reference class.[33] Plaintiffs and prosecutors may be tempted to fish for quantitative evidence.[34] Courts may be overwhelmed by the computational complexity of a Bayesian approach if there are multiple, in particular multiple dependent pieces of evidence.[35]

In this note I bracket this debate. I leave aside the incentive arguments against using base rates in court: individuals from a group with a low prior probability of committing the antisocial act in question could exploit that they are shielded from liability;[36] groups who are initially discriminated against would run an ever increasing risk of liability, so that discrimination would become self-fulfilling.[37] I also do not want to contribute to the still small body of empirical knowledge about the use of statistics in general, and of Bayes' rule in particular in court. Courts have been divided over accepting base rate information.[38] In line with the cited evidence on doctors, when provided with appropriate decision aids, the decisions of mock jurors

---

[27]   KOEHLER & SHAVIRO, *supra* note 21, at 250-252.

[28]   LEONARD J. SAVAGE, The Foundations of Statistics (Wiley. 1954).

[29]   ALLEN & LEITER, *supra* note 21, at 1508;ALEX STEIN, *Judicial Fact-finding and the Bayesian Method. The Case for Deeper Scepticism About their Combination*, 1 International Journal of Evidence and Proof (1996) ; but see the more nuanced position of CHARLES YABLON, *Meaning of Probability Judgments. An Essay on the Use and Misuse of Behavioral Economics,* University of Illinois Law Review (2004).

[30]   TRIBE, *supra* note 14, at 1358.

[31]   Id., at 1366.

[32]   MICHAEL I. MEYERSON & WILLIAM MEYERSON, *Significant Statistics. The Unwitting Policy Making of Mathematically Ignorant Judges*, 37 Pepperdine Law Review (2010).

[33]   KOEHLER & SHAVIRO, *supra* note 21, at 259;DAVID H. KAYE, *DNA Evidence. Probability, Population Genetics, and the Courts*, 7 Harvard Journal of Law and Technology (1993).

[34]   This is the core of the debate over data base trawls for DNA matches, see ROTH, *supra* note 25, for a skeptical view: admissible at most if the probability standard is elevated to 99.9%; but see DAVID H. KAYE, *Rounding Up the Usual Suspects. A Legal and Logical Analysis of DNA Trawling Cases*, 87 North Carolina Law Review (2008);DAVID H. KAYE, *Trawling DNA Databases for Partial Matches. What is the FBI Afraid of?*, 19 Cornell Journal of Law and Public Policy (2009);NICHOLAS SCURICH & RICHARD S. JOHN, *Trawling Genetic Databases. When a DNA Match is Just a Naked Statistic*, 8 Journal of Empirical Legal Studies (2011).

[35]   TRIBE, *supra* note 14, at 1366-1369;ALLEN & LEITER, *supra* note 21, at 1507;RICHARD D. FRIEDMAN, *Generalized Inferences, Individual Merits, and Jury Discretion*, 66 Boston University Law Review (1986) . Mark Schweizer, in his forthcoming treatise, proposes a very elegant, and an easily manageable solution to this problem, namely computerized Bayes networks as decision aids.

[36]   Leo Katz pointed me to this implication.

[37]   Ralf Poscher pointed me to this implication.

[38]   JONATHAN J. KOEHLER, *When Do Courts Think Base Rates Statistics Are Relevant?*, 42 Jurimetrics (2002).

on a DNA case came closer to the Bayesian norm.[39] If mock jurors are exclusively provided with base rate information, their estimate of the probability that the defendant committed the crime increases, but they are unwilling to convict on this basis.[40]

I want to make a normative point. If courts rely on base rate information, this can violate normative decisions previously taken by the legal order elsewhere. I thus call for normative coherence. I claim that these normative concerns trump the accuracy goal. As long as the law does not change, to the extent that these inconsistencies exist, base rates must be neglected.

I proceed in two steps. I use hypotheticals to demonstrate the conflict between a Bayesian approach and the existing law. I then ask whether these inconsistencies are fatal for any Bayesian approach to evidence in court, or whether they leave room for a differentiated solution.

2.　When Does Using the Base Rate Violate the Law?

Consider my first hypothetical:

> *Immigrant 1*: Three men meet an immigrant in the street. They call him names, and start beating him up. Things turn violent. The immigrant is severely wounded. A doctor certifies that a single blow caused the wound. During the brawl, there was no witness. It was dark, so that the immigrant could not discern who actually hit him.

Seemingly, there is not much room for Bayesian analysis here. All we know is the base rate. We can safely exclude that a stranger has caused the wound, or that the immigrant has hurt himself. The prior probability of either of the three men having caused the wound is 1/3. If the immigrant sues for damages, this is a private law suit. The standard of proof, under US law, is preponderance of the evidence. If it is interpreted quantitatively,[41] it is not met. The probability is below .5.

---

[39]　NANCE & MORRIS, *supra* note 18.

[40]　GARY L. WELLS, *Naked Statistical Evidence of Liability. Is Subjective Probability Enough?*, 62 Journal of Personality and Social Psychology (1992);KEITH E. NIEDERMEIER, et al., *Jurors' Use of Naked Statistical Evidence. Exploring Bases and Implications of the Wells Effect*, 76 Journal of Personality and Social Psychology (1999) ; also see KEVIN JON HELLER, *The Cognitive Psychology of Circumstancial Evidence*, 105 Michigan Law Review (2006).

[41]　Many scholarly observers call for quantification, DOROTHY K. KAGEHIRO & W. CLARK STANTON, *Legal vs. Quantified Definitions of Standard of Proof*, 9 Law and Human Behavior (1985);DOROTHY K. KAGEHIRO, *Defining the Standard of Proof in Jury Instructions*, 1 Psychological Science (1990);HARRY D. SAUNDERS, Quantifying Reasonable Doubt: A Proposed Solution to an Equal Protection Problem (2005);PETER TILLERS & JONATHAN GOTTFRIED, *Case comment—United States v. Copeland, 369 F. Supp. 2d 275 (E.D.N.Y. 2005): A Collateral Attack on the Legal Maxim That Proof Beyond A Reasonable Doubt Is Unquantifiable?*, 5 Law, Probability and Risk (2006);JACK B. WEINSTEIN & IAN DEWSBURY, *Comment on the Meaning of 'Proof Beyond a Reasonable Doubt'*, 5 Law, Probability and Risk (2006) ; but see JAMES FRANKLIN, *Case Comment - United States v. Copeland, 369 F.Supp.2d 275 (E.D.N.Y. 2005): Quantification of the 'Proof Beyond Reasonable Doubt' Standard*, 5 Law, Probability and Risk (2006).

Now consider the following variant:

> *Immigrant 2:* The police have been called to the crime scene, and have ordered a forensic investigation. An expert testifies that, if the blow comes from a left-handed aggressor, this type of injury is observed in ¾ of all cases, while it is only observed in ¼ of all cases if the aggressor is right-handed. One, and only one, of the aggressors is left-handed.

Now Bayesians are in business. The prior probability that the blow came from the one left-handed person is 1/3. In Bayesian parlance, the expert's statement is about the conditional probability of the signal (signs of the blow coming from a left-handed person) being present if the one left-handed aggressor has indeed hit the immigrant: $p(e|g) = \frac{3}{4}$. The remainder is an exercise in arithmetics. The prior of the defendant not having hit the immigrant is $p(\neg g) = 1 - p(g) = \frac{2}{3}$. If we take the expert's statement at face value, arguably the conditional probability of finding these signs on the victim's body even if the perpetrator was not left-handed (i.e. was right-handed) is given by $p(e|\neg g) = 1 - p(e|g) = \frac{1}{4}$.[42] With this information, we first can reconstruct the total probability of finding the signs of the immigrant being hit by a left-handed person: $p(e) = p(g) * p(e|g) + p(\neg g|e) * p(e|\neg g) = \frac{5}{12}$. Plugging this in, we have

$$p(g|e) = \frac{p(g) * p(e|g)}{p(s)} = \frac{\frac{1}{3} * \frac{3}{4}}{\frac{5}{12}} = \frac{3}{5}$$

If Bayes' rule is the right approach, if we apply the preponderance of the evidence standard, and if we translate it into $p(g|e) > \frac{1}{2}$,[43] in the *Immigrant 2* case, the left-handed person should be ordered to pay damages. If the immigrant did not have or use the additional evidence in the first place, and sued one of the other two aggressors, he should loose. Using Bayes' rule again, the posterior probability that one of them actually hit the immigrant is $\frac{1}{5} < \frac{1}{2}$.

Yet this is not the law. A court would instead rely on Sec. 876 Restatement 2d of Torts

> "For harm resulting to a third person from the tortious conduct of another, one is subject to liability if he
>
> (a) does a tortious act in concert with the other or pursuant to a common design with him, or
>
> (b) knows that the other's conduct constitutes a breach of duty and gives substantial assistance or encouragement to the other so to conduct himself, or

---

[42] In theory, the two likelihoods need not add up to 1. But in the concrete case, it seems plausible that the likelihood of finding these signs on the victim's body if the actual perpetrator was right-handed are 1 minus the likelihood of finding them if the perpetrator was left-handed.

[43] An alternative interpretation uses posterior odds, i.e. $\frac{p(g|s)}{p(\neg g|s)}$. Given the evidence, how much more is it likely that the defendant is guilty, rather than innocent? If both is equally likely, posterior odds are 1. Consequently, the preponderance standard requires that posterior odds are strictly larger than 1.

(c) gives substantial assistance to the other in accomplishing a tortious result and his own conduct, separately considered, constitutes a breach of duty to the third person."

The defendant is already liable under *Immigrant 1*. The immigrant could sue either of the three aggressors under *Immigrant 2*. If he had sued the "wrong" one, and this defendant brought the evidence about the likely left-handedness of the aggressor, this would not exonerate him.[44]

There is of course nothing conceptually wrong with the Bayesian approach. Yet the faithful Bayesian would solve the wrong problem. If someone "does a tortious act in concert with the other or pursuant to a common design with him", the law no longer cares about causation. Specifically, it replaces the standard requirement (defendant has caused harm to plaintiff) by a more lenient requirement (defendant has teamed up with others; in some way this team has caused harm to plaintiff). With this provision, substantive law reacts to the characteristic difficulty in finding evidence. Reconstructing the details of a brawl is frequently impossible. The law reacts by changing the definition of the tortious act. Teaming up with the intent to cause harm is enough. All who participate are jointly and separately liable. Plaintiff is free to pick the one member of the gang that was easiest to find, or most likely solvent. This one member may later sue other gang members for their fair share.

Is this but an oddity of the law of torts, the proverbial exception that proves the rule (that a proper evidentiary lawyer should be Bayesian, that is)? Consider the next hypothetical:

> *Asbestos 1:* A works for a construction company. A has attracted lung cancer. It can be shown that the cause is longterm exposure to asbestos. The construction company has been insolvent and dissolved. Over the years, A's company has used asbestos by three different suppliers, in about equal shares.

For the faithful Bayesian, there is no difference between *Immigrant 1* and *Asbestos 1*. Under the rules on product liability, the worker has a direct claim in torts against the manufacturer,[45] provided it can be shown that the product sold by this manufacturer has indeed caused the disease. Given the evidence, this probability is only 1/3, so that a Bayesian will dismiss the claim under the preponderance of the evidence standard. Now consider the following variant:

> *Asbestos 2:* A provides a list of construction sites where he has been working over the years. A's former employer provides a list of bills for these sites. In 3/4 of these sites, asbestos by supplier S has been used.

The faithful Bayesian will treat *Asbestos 2* the same way as *Immigrant 2*. By the additional evidence, the conditional probability of S having supplied the material, given A has been employing it, i.e. $p(e|g)$, is ¾. Consequently, and formally in exactly the same way as in *Immi-*

---

[44]   DAVID KAYE, *The Limits of the Preponderance of the Evidence Standard. Justifiably Naked StatisticalEvidence and Multiple Causation*, 7 American Bar Foundation Research Journal (1982)   is closest to the claim made in this note. Yet he proposes a different solution, namely (in appropriate cases) to deviate from the preponderance standard. I do not think the standard is the problem, but the base rate of the event, provided it is defined as being the member of a class of events.

[45]   Restatement 2d on Torts, Sec. 402a.

grant 2, the Bayesian calculates the posterior probability of S having supplied the asbestos in question to be 3/5 > 1/2. This meets the preponderance standard. S has to compensate A for medical expenses, and for pain and suffering.

Yet again this is not the law. There is one obvious difference between *Immigrant 2* and *Asbestos 2*: the three suppliers of asbestos have neither acted in concert nor pursued a common design. All they have in common is selling asbestos. Actually, provided there is no cartel, on this market they do not coordinate but, to the contrary, compete with each other for market share. The normative justification of easing the burden of proof inherent in Sec. 876 Restatement 2d of Torts is clearly not present.

In many legal orders, this would be the end of the story. US law is different. The case would come under the purview of the rules on market share liability.[46] These rules imply that A would no longer lose under *Asbestos 1*. He would have a claim against S, but it would be limited to 1/3 of all damages.

Now A might object: S may be a more attractive defendant than the remaining two suppliers. Why not use the Bayesian argument to get full compensation from A? Yet both the rules on brawls and on market share liability are normative decisions about who should bear which portion of the risk of uncertainty. If tortfeasors act in concert, the legal order has them bear the entire risk of taking a materially wrong decision because it is not possible to find out with certainty which member of a jointly acting group has actually caused harm. By contrast, even under the consumer friendly solution of US law, each producer is only held liable for that fraction of the damage that corresponds to its share of the market in question. Using Bayes rule, and the preponderance standard, one would overrule this normative standard. Producers would bear a risk above their market share whenever additional, but inconclusive evidence is produced. Specifically, provided this additional evidence pushes the posterior probability above the threshold of 50%, the individual producer would be liable for the entire damage although the normative justification for such an outcome resulting from the rules on tortfeasors acting in concert is not present.

The next problem is typically discussed in criminal, not in private law. To maintain the parallel with the previous two problems, I transpose it to private law. Consider the following hypothetical:

> *Clock 1:* Burglars have stolen the precious grandfather clock from A's home. A year later, the clock is offered for sale on an Internet platform. There are three possibilities how seller S may have come into possession of the clock. He may have been one of the burglars himself. He may have bought the clock from one of the burglars, knowing that it was stolen. He may have bought the clock on a flea market without knowing that it was stolen property. After trial, any of these three courses of action still seem equally likely. A sues S. He not only wants his clock

---

[46]   *Sindell vs. Abbott Laboratories,* Cal Sup Ct, March 20, 1980, 26 C3d 588.

back. During the year, the clock has been damaged. A also claims money to restore the clock.

Whether A gets his clock back is not the issue. Repossession does not require tort. In most legal orders stolen goods may also not be acquired in good faith. Yet damages require tort. Let us apply the Bayesian logic to the case. A is liable if he was one of the burglars. Yet this was only the case with probability 1/3 < 1/2, so that the preponderance standard is not met. Fencing is also tortious. If the second chain of events is what happened, S would again be liable. But this too only happened with probability 1/3 < 1/2. So eventually, a Bayesian must dismiss the claim. Compare this with the following variant:

> *Clock 2:* A proves to the requisite standard that S has sold hundreds of stolen items on this same Internet platform. S could not produce evidence for having bought legally any of these items. Given this circumstantial evidence, the court deems the probability that S is a professional fencer, and that he acquired the clock in exercise of this profession, to be ¾.

These are, of course, again the same numbers. Before A presented the additional evidence, the probability of S having committed the tortious act of knowingly selling a stolen good, i.e. $p(g)$, is 1/3. By the additional evidence, a Bayesian has to update her prior with $p(e|g) = \frac{3}{4}$, to get $p(g|e) = \frac{3}{5}$. The Bayesian would have the defendant pay damages in *Clock 2,* but not in *Clock 1*.

Most likely, this is not the law. If S had burglarized A's house, he would be liable. If S had knowingly sold goods that a burglar had stolen from A's house, he would be liable. Arguably, if it was established with certainty that S had committed either crime, he would still be liable. At least this is how the case would typically be treated under criminal law.[47] Arguably, if one applies the preponderance standard, given the total probability is 2/3 that S has committed either of these two crimes, he should be held liable. S would be liable in *Clock 1* as well. Now consider

> *Neighbour 1:* Over decades, A and his neighbor N have been enemies. When he returns from the annual holiday, A finds his cat dead, and a tree in his garden felled. He suspects N to have done both, in continuation of their feud. The pecuniary value of the cat and of the tree is $ 1000.- each. A sues for $ 1000.- in damages. A offers witness testimony for both acts. Yet both witnesses are engaged in the feud, on A's side. The Court therefore is distrustful and attaches a probability of 1/3 only to N having committed either act.[48]

---

[47] For detail see ALON HAREL & ARIEL PORAT, *Aggregating Probabilities Across Cases. Criminal Responsibility for Unspecified Offenses*, 94 Michigan Law Review (2009).

[48] Strictly speaking, in this example, given the estimates are generated after trial, they are not priors but posterior estimates of probability. Yet Bayesian updating may be iterative. A Bayesian may start from an uninformative prior, update it for a first time in the light of the initial evidence, and reupdate if new evidence becomes available, always using Bayes' rule.

By the now familiar line of reasoning, the Bayesian judge dismisses the claim. Compare this with

> *Neighbour 2:* A additionally finds a stranger who happened to observe how a person whom he describes in ways that resemble N to attack a cat with a stone on one of the days when A was on holiday. Since the stranger cannot identify N with certainty, the court attaches probability ¾ to his testimony showing that N has actually killed A's cat.

Again, by the familiar line of reasoning, and under the preponderance standard, the court would now order N to pay $ 1000.- as the pecuniary value of the cat. By the additional testimony, the posterior probability of N having committed this tort is 3/5.

Even if N actually had committed both acts, they would only be related by the identity of the tortfeasor and the victim. This would not be enough to decide against N on the argument that it is more probable than not that he has committed at least one of those acts. The law does not give the court power to consider both acts as members of one class.[49]

Seemingly, the *Asbestos* and the *Neighbor* cases are analogues. In neither of these cases the law is content with the case being a member of a class. Yet there is a difference. If the court uses Bayes' rule to decide either case, by defining a base rate, of necessity the court treats the *Asbestos* case as a member of a class. This is not so in the *Neighbor* case. For in the *Neighbor* case, the two events to which the court attaches a probability of 1/3 each are not mutually exclusive. There is no conceptual reason why N should not have committed both acts, or yet other tortious acts that A has not yet detected. There is also no reason to conclude that the prior probability of N having committed neither act is also 1/3. The probability of 1/3 is just an assessment based on less conclusive evidence. Note that, by contrast, in the *Clock* cases, the base rate *is* a statement about mutually exclusive interpretations of an outcome (S got into possession of the clock), and hence assumes that the interpretations are members of a class.

3.   Implications for a Bayesian Approach to Evidence in General

Thus far we have found an inconsistency. If courts treat any uncertainty about the facts the Bayesian way, they may be at variance with the normative decision of substantive, not procedural, law when to treat events as members of a class, and when not. As pointed out in the introduction, this is not an objection against any Bayesian argument in court. Yet it might be an argument against attaching a base rate to the disputed event if this base rate is not based on specific evidence, but merely on the membership of the event in a class. If the criterion can be upheld, it also addresses (one interpretation of) "naked statistical evidence".[50] Such evidence may not be used in court if it implies treating disputed events as members of a class.

---

⁴⁹   See again HAREL & PORAT, *supra* note 47.
⁵⁰   It is the one also used in WELLS, *supra* note 40;NIEDERMEIER, et al., *supra* note 40.

This criterion allows to decide some of the classic hypotheticals that have kept this literature engaged. Consider

> *Buses:* A passenger has been hit by a bus at night. The bus absconded. It was so dark that the passenger could not discern the color of the bus. It is established that 80% of all buses in town are operated by company A and are painted blue. The remaining 20% are operated by company B and are painted green.[51]

A faithful probability theorist (who need not even be Bayesian) will order company A to pay damages since the probability of this being the correct decision is 80%, which is much more than the preponderance standard requires.

Many have taken issue with this result. It would drive the larger company out of the market since, in anticipation, that company would have to pay for almost all accidents, and in particular for those committed by its competitor.[52] The fact that plaintiff offers no additional evidence may itself have probative value;[53] this argument would, however, not be relevant in the current version of the problem since plaintiff could not be expected to produce additional evidence. Others have suggested to change the standard of proof and to require a higher standard than 51% if the evidence is scanty, the stricter the scantier the evidence.[54] In the perspective proposed in this note, relying on the proportion of buses in town is an obvious violation of the prohibition to treat the two companies as members of a class although they have clearly not acted in concert. The conditions for market share liability are not fulfilled either (and would anyhow not lead to full liability).

Consider next

> *Gatecrasher:* 1000 attended a rodeo although the promoter only sold 499 tickets. The promoter sues all for damages.[55]

This is a strange case in the first place. How come the promoter is able to identify visitors, but unable to identify those who have paid? Why do those who actually have paid not simply produce their tickets? But the criterion proposed here will also take care of this hypothetical. Attending the same show does not turn visitors into a group acting in concert.

> *Prison Guard:* 24 of 25 prisoners collaborate in the murder of a prison guard.[56]

---

[51]   This is the original version of the problem proposed by TRIBE, *supra* note 14, at 1340.. RICHARD A. POSNER, *An Economic Approach to the Law of Evidence*, 51 Stanford Law Review, 1508 (1999) proposes a more radical version: 51% of the busses are operated by one company, so that the preponderance standard is exactly matched. Decision theorists have added additional testimony by a witness who says the cab is green, and have used the example to test base rate neglect, BAR-HILLEL, *supra* note 1, at 211 f;DANIEL KAHNEMAN & AMOS TVERSKY, *On Prediction and Judgement*, 12 Oregon Research Institute Bulletin (1972).

[52]   POSNER, *supra* note 51, at 1510.

[53]   TRIBE, *supra* note 14, at 1349;POSNER, *supra* note 51, at 1509;RICHARD O. LEMPERT, *The New Evidence Scholarship. Analyzing the Process of Proof*, 66 Boston University Law Review, at 457 (1986).

[54]   RICHARD D. FRIEDMAN, *Generalized Inferences, Individual Merits, and Jury Discretion*, 66 Boston University Law Review (1986).

[55]   L.JONATHAN COHEN, The Probable and the Provable (Clarendon Press Oxford. 1977);L.JONATHAN COHEN, *Subjective Probability and the Paradox of the Gatecrasher*, Arizona State Law Journal (1981).

Here, a Bayesian approach would be misplaced for the opposite reason. The 24 prisoners have teamed up and should therefore be held jointly and individually liable.

Finally consider the court case that has been at the start of the literature:

> *People vs. Collins:* In a criminal case, defendants were convicted after an instructor of mathematics testified to the product rule, "assuming the robbery was committed by a Caucasian woman with a blond ponytail who left the scene accompanied by a Negro with a beard and mustache, there was an overwhelming probability that the crime was committed by any couple answering such distinctive characteristics. […] There was but one chance in 12 million that any couple possessed the distinctive characteristics of the defendants".[57]

One prominent commentator has tried to justify this decision on Bayesian grounds.[58] This is a striking violation of treating individuals as members of a class to which they do not obtain. Actually the class in question is the entire population of the country!

All of this seems to imply: the base rate of the event should not be admitted as an argument in court. Yet consider the following hypothetical:

> *Sexual Assault:* Victim V was sexually assaulted. The perpetrator was wearing a balaclava, which is why the victim could not recognize the face. The victim scratched the perpetrator. Skin particles can be recovered from underneath her nails. The police performs a search in a DNA database and finds a match. The DNA is from a 40year old woman whose DNA was entered into the database as supporting evidence in a paternity suit.

In principle, DNA evidence is very reliable. Depending on circumstances, the probability of a random match is at least as small as 1/50,000,000 and can well be as small as 1/204,800,000,000,000.[59] The risk of a lab mistake, and of the evidence being doctored, can be much higher. But even if this is taken into consideration, the posterior probability of a false positive is no bigger than 1.9%.[60]

Yet obviously something is fishy here. Sexual assaults of women on women are very rare. One dated source, covering Alabama, North Carolina and South Carolina in 1991, estimates that only .8% of all incidences fall into that category.[61] If this is taken into consideration, the defendant is treated as a member of a class (of all persons committing sexual assault, that is). This class is almost as general as the class in *People v. Collins*. These perpetrators do by no

---

[56]   CHARLES R. NESSON, *Reasonable Doubt and Permissive Inferences. The Value of Complexity*, 92 Harvard Law Review, 1192 f. (1979).; also see the discussion in KOEHLER & SHAVIRO, *supra* note 21, at 252 f.

[57]   68 Cal.2d 319 (1968).

[58]   MICHAEL O. FINKELSTEIN & WILLIAM B. FAIRLEY, *A Bayesian Approach to Identification Evidence*, 83 Harvard Law Review (1970)., but see TRIBE, 1334-1338.

[59]   It is between $\left(\frac{1}{5}\right)^{11}$ and $\left(\frac{1}{20}\right)^{11}$ since, in US practice, 11 independent snippets of DNA are analyzed, and the probability of finding other individuals with the same snippet is between 1/5 and 1/20.

[60]   Assuming that all three risks are independent, that both concomitant risks are no bigger than 1/100, and that the random match probability is no bigger than 1/50,000,000, and using the product rule. Of course, this is *not* a complete application of Bayes' rule, in that no prior is specified.

[61]   http://www.mincava.umn.edu/documents/sexoff/sexoff.html.

means act in concert. Given the criterion proposed in this note, the court would not be allowed to use this information. The defendant will sure see this as a striking violation of justice. Were the low base rate taken into account, the posterior probability of the defendant being guilty (and in a civil law suit being liable) would go down to .2843[62] – way below even the preponderance standard, let alone the beyond a reasonable doubt standard if the defendant was prosecuted.

Seemingly this creates a dilemma. When I have advocated neglecting the base rate I have also argued normatively. Seemingly justice is pitted against justice. It would be unjust to ignore the base rate in the *Sexual Assault* case. And it would be unjust to treat the defendants as members of the class of all inhabitants of the US in *People vs. Collins*. Yet there is a difference. The law only treats multiple aggressors as members of a class if they act in concert. The law only treats producers as members of a class if they sell sufficiently similar products on the same market. The law only treats multiple acts by the same individual as items of a class if these acts are sufficiently related and similar. All these "onlys" are meant as decisions against holding individuals liable for no other reason than that they or their acts may be subsumed under a broader category, as long as the additional qualifying factors are not present. By defining the conditions under which the uncertainty may be held against the defendant the law draws a line. Beyond this line, it does not want liability. The fact that there is uncertainty shall protect the defendant. In the *Sexual Assault* case, no similar normative decision is discernible. The legal order does not want to make it easier for prosecution to get a person convicted that is charged with sexual assault. If the case is tried in torts, the legitimate interests of the plaintiff may be factored in. This is the main justification for the preponderance, rather than the more stringent beyond a reasonable doubt standard.[63] But there is no normative reason why the plaintiff should win more easily, despite the base rate being very low.

A further illustration is evidentiary rules on the admissibility of character evidence. If the only goal is accuracy of verdict, it is relevant information that the defendant has previously been convicted for other crimes. The information is the more relevant the closer the earlier crime is related to the current crime. If such information is admitted, the crime at trial is treated as an element of a class of acts, namely all acts for the commission of which the defendant has previously been convicted. As a rule, US law prevents prosecution from bringing such evidence (Rule 404 (b) (1) Federal Rules of Evidence). In principle, the same rule applies for character evidence in general (Rule 404 (a) (1)). Yet the defendant may bring character evidence to exculpate herself (Rule 404 (a) (2) (A)). We thus find the same asymmetry. Moreover we find

---

[62] This result obtains on the following assumptions: from the match with the DNA database, taking the risks of lab failure and the evidence being doctored into account, one calculates $p(e|g) = .9801$. One assumes $p(e|\neg g) = 1 - p(e|g) = .0199$. The prior probability of female-female sexual assault is $p(g) = .008$. Since events are mutually exclusive, $p(\neg g) = .992$. Using this information, one reconstructs $p(e) = p(g) * p(e|g) + p(\neg g) * p(e|\neg g) = .02758$. The final step is calculating
$p(g|e) = \frac{p(g)*p(e|g)}{p(e)} = \frac{.008*.9801}{.02758} = .2843$.

[63] KEVIN M. CLERMONT & EMILY SHERWIN, *A Comparative View of Standards of Proof*, 50 American Journal of Comparative Law (2002).

exceptions for sexual assault and child molestation (Rules 413-414). As a matter of legal poli-cy, I find these exceptions questionable. The presumption of innocence should not depend on the charge. But the exceptions make it patent that introducing base rate information is a nor-mative issue and that, constitutional limitations notwithstanding, it is for the legislator to make this decision.

This leads to the following result: (1) if the law asks for prognostic evaluation, the base rate of the event is legally relevant. (2) If additional evidence is made available, it may be integrated with the previously available evidence using Bayes rule. This implies that from the previous evidence a base rate of the event is derived. (3) The base rate of the event may be introduced as a defense against otherwise compelling evidence. But (4) the base rate of the event may not be held against the defendant if this implies treating the defendant as a member of a class of multiple individuals. And (5) the base rate of the event may not be held against the defendant if this implies attaching a probability to mutually exclusive alternative causal chains.

While it covers many cases, this list is likely not exhaustive, and compiling the list has not been the main purpose of this note. Rather I wanted to point to the underlying consistency problem. In court, the risk of making a mistake is not the exclusive normative concern. Who of the conflicting parties shall bear the risk of critical facts remaining uncertain can be a mat-ter of substantive, not procedural law. If substantive law assigns this risk to one party, this normative decision may not be overruled by applying Bayes rule.

A final case illustrates how my criterion could be extended to other evidentiary issues:

> *Shonubi:* The Nigerian national Shonubi was caught at JFK Airport with 427.4 grams of heroin in his stomach. It could be proven that this was his eighth round-trip, and that he had been smuggling heroin on all occasions. Yet for sentencing, it was relevant whether the total amount exceeded 1000 grams, or even 3000 grams. For his earlier trips, no quantity information was available. The trial judge relied on FBI information about the amounts smuggled by other Nigerians and caught at the same airport, using the same method, in a closely related time period, and e-stimated that Shonubi had at least smuggled another 572.6 grams, which qualified him for the considerably higher sentence.[64]

Critics claim that the decision suffers from the indeterminacy of the reference class: why compare Shonubi with other Nigerian drug traffickers caught at the same airport?[65], and from disguised racial bias: the decision implicitly holds his nationality against the defendant.[66] In my view, the issue must be settled by substantive law. US criminal law has decided to con-strain trial judges' discretion in sentencing, and to oblige them to apply the law like sentenc-

---

[64]   *United States vs. Shonubi,* 802 F. Supp. 859 (E.D.N.Y. 1992); 998 F.2d 84 (2d. Cir. 1993); 895 F. Supp. 460 (E.D.N.Y. 1995); 103 F.3d 1085 (2d. Cir. 1997).

[65]   MARK COLYVAN, et al., *Is it a Crime to Belong to a Reference Class?*, 9 Journal of Political Philosophy (2001).

[66]   PETER TILLERS, *If Wishes Were Horses. Discursive Comments on Attempts to Prevent Individuals from being Unfairly Burdened by their Reference Classes*, 4 Law, Probability and Risk (2005).

ing guidelines.[67] The Guidelines expressly regulate how to proceed if facts relevant for determining the sentence are disputed:

> "When any factor important to the sentencing determination is reasonably in dispute, the parties shall be given an adequate opportunity to present information to the court regarding that factor. In resolving any dispute concerning a factor important to the sentencing determination, the court may consider relevant information without regard to its admissibility under the rules of evidence applicable at trial, provided that the information has sufficient indicia of reliability to support its probable accuracy.[68]"

Being a German lawyer who does not believe too much into harsh sentences deterring future crime, I do not find this decision appealing. But it seems to say quite clearly that the (quasi) legislator only cares about accuracy. Since the accuracy of the estimate does not seem to be in doubt, trial judges seem empowered to use this base rate.

In the legal discourse, Bayes rule is typically equated with deriving a prior from the empirical base rate. In this spirit, I have used the two terms exchangeably. Yet in conclusion, I should stress that one can be a Bayesian without using statistical evidence, or any information about empirical distributions, for that matter. All a faithful Bayesian needs is a prior. Ultimately, this prior can be uninformative (one assumes that all possible events are equally likely), or it can be determined by an exogenously imposed normative criterion. Eventually, my argument could therefore also be presented as an argument within a Bayesian framework, but calling for caution when using empirical base rates to decide evidentiary problems. The critical point is this: whether the empirical base rate of the disputed event may be used in the law of proof is a normative issue. As a rule, it is decided by substantive, not evidentiary law.

---

[67]   Sentencing Reform Act 1984, but see *Booker*, 543 U.S. 220 (2005), and Sentencing Guidelines 2010, Chapter 1, p. 12.

[68]   Sentencing Guidelines 2010, § 6A.1.3 (a).