

**Preprints of the
Max Planck Institute for
Research on Collective Goods
Bonn 2015/14**



Modeling a Satisficing Judge

Christoph Engel
Werner Güth



MAX PLANCK SOCIETY



Modeling a Satisficing Judge

Christoph Engel / Werner Güth

4 October, 2015

Modeling a Satisficing Judge

Christoph Engel & Werner Güth

Sunday 4th October, 2015

Judges and juries frequently must decide, knowing that they do not know everything that would be relevant for deciding the case. The law uses two related institutions for enabling courts to nonetheless decide the case: the standard of proof, and the burden of proof. In this paper, we contrast a standard rational choice approach with a satisficing approach. Standard theory would want judges to rationally deal with the limitations of the evidence. We posit that this is not only descriptively implausible, but also normatively undesirable. We propose a theoretical framework for a judge who only considers scenarios that "she does not dare to neglect", and aims at decisions that are "good enough", given the undissolvable limitations of the evidence. We extend this approach to parties who strategically exploit the limited factual basis, and to judges who have to allocate limited resources for fact finding to more than one case.

JEL: C72, D03, D81, D82, K41

Contents

1	Research Question	3
2	The Myopic Judge	5
2.1	The Judge's Goal	5
2.2	The Divine Judge	6
2.3	The Human Judge	7
2.3.1	Scenario Construction	9
2.3.2	Removing Implausible Scenarios	10
2.3.3	Acceptable Outcomes	11
2.3.4	Decision Making	12
2.3.5	Beyond a Reasonable Doubt	13
2.4	Implications for the Normative Debate	14
2.5	Discussion	14
3	The Prevoyant Judge	15
3.1	The Judge's Goal	15

3.2	The Divine Judge	15
3.3	The Human Judge	16
3.3.1	Scenario Construction	17
3.3.2	Removing Implausible Scenarios	17
3.3.3	Acceptable Outcomes	17
3.3.4	Decision Making	18
3.4	Implications for the Normative Debate	18
4	The Professional Judge	19
4.1	The Judge's Goal	19
4.2	The Divine Judge	19
4.3	The Human Judge	20
4.4	Implications for the Normative Debate	21
5	Conclusion	21

1

¹We are grateful for helpful comments by Konstantinos Chatziathansiou and Philip Weinschenk, and seminar participants at the Max Planck Institute for Research on Collective Goods, Bonn (Germany)

1 Research Question

Why does a case go to court? Because there is a dispute.¹ Why is there a dispute? Usually because some facts are disputed. Sometimes, at trial all doubt is removed. But often the evidence only partly removes the uncertainty. However the court is not allowed to send the parties home without taking a decision. To make a decision possible, the law obliges the court to assign the risk of error resulting from the undissolvable uncertainty to one of the parties. Legal orders, and procedural settings for that matter, differ in the normative decision which of the parties has to bear that risk. Irrespective of legal order, the ultimate loser is plaintiff (in civil procedure) or prosecution (in criminal procedure): the burden of proof is on her. But it depends on the standard of proof at which point this default becomes relevant. If the standard is *preponderance of the evidence*, the law wants the court to distribute the risk equally between the parties. If it is a little more likely that plaintiff has a case, she should win. If it is a little more likely that plaintiff's case is not well founded, plaintiff should lose. Contrast this with the *beyond a reasonable doubt* standard. In the US it applies in criminal matters. On the European continent, it holds in all disputes. The legal orders care about different things (Engel, 2009). In a dispute among private parties, US law argues: ex ante the legal order does not have a reason to privilege one party over the other (Clermont and Sherwin, 2002). By contrast, Continental law is concerned about the fact that any court ruling is an exercise of sovereign powers. The legal order wants to reduce the risk that sovereignty is exercised unduly to a degree that seems acceptable, given circumstances (Taruffo, 2003).

In this paper, we aim at reconstructing these competing normative choices. In a first step, we formally show what either choice implies in a fictive world where the only limitation is epistemic: for some exogenous reason, the court only has incomplete access to the facts of the case about which it has to decide. But given this limitation, the court decides optimally. This is a strange world, though: the law acknowledges that the court is very likely to only have limited access to the facts. But the law assumes that limited access to information is the only constraint; jury and judge have unlimited cognitive resources. We begin by formally reconstructing this choice (see section 2.2 below). The purpose of this paper is to contrast this utterly unrealistic ("*divine*") court with an empirically more plausible and therefore normatively more appealing model. What do the competing normative choices (epitomized by the difference between US and Continental law) imply for a court that is, itself, populated by human beings? Note the status of the argument. We do not want to show that court practice does not live up to an overly demanding norm. Rather we want to reconstruct the normative difference between US and Continental standards of proof in the light of a more plausible descriptive model of judicial decision-making: what do the different standards imply for a "*human*" judge (see section 2.3 below)?

Of course, "realism" is not a well-defined concept. Potentially, reality diverges from a model in an infinite number of ways. In this paper, we decide to focus on the one dimension that arguably is most important for understanding the difference between legal orders. We

¹This increasingly even holds in criminal procedure. If the facts are undisputed, trial is routinely replaced by a deal between prosecution and defense.

posit: the position of US law implicitly assumes that judges are "*divine*". By contrast, the position of Continental law fits better for judges who, themselves, try as best they can, but are fully aware that their own cognitive abilities have limitations. Note that our approach is still consequentialist: satisficing judges are as forward looking as are "rational" judges. All we aim at offering is a "psychologically feasible" model of judicial decision making. i.e. a theory of decision making that is "absorbable by human beings". Arguably this also implies: while the superrational approach is beyond the cognitive abilities of real judges, the legislator might make an effort to teach judges and jury members to comply with our approach, or it might give them decision aids that help them do so.

In section 2 we make two simplifying assumptions: (1) the court does not expect the parties to anticipate how it deals with undissolvable uncertainty. Since we want to make a contribution to a normative debate, this assumption may seem tenable. When giving the court a rule, the legal order may not want to assume that the law's subjects strategically exploit weaknesses in court procedure. Yet it is important to understand whether the normative assessment of the difference between US and Continental law changes once one drops this assumption. This we do in section 3.

(2) All of the foregoing is a model for a court that decides a single case. This is in line with most of the normative debate, that focuses on the risk of falsely deciding individual cases. Yet real courts have a docket. They are only able to handle so many cases. One determinant of trial length is in the hands of the court. It can decide how much evidence to gather, and how much time to spend on each piece of evidence. Now additional evidence, or greater scrutiny in assessing the probative value of a certain piece of evidence for that matter, potentially reduce the risk of erroneous judgment. The court therefore effectively faces an allocative decision. It must decide which fraction of a limited budget of time and energy to devote to a single case. In section 4, we add this additional level of complexity to the analysis. We again start from the assumption of a *divine* judge, and contrast her with the decision made by a *human* judge, who relies on satisficing. This section also addresses another objection one might raise against the analysis presented in section 2: in that section we assume that cost is not an issue. Yet even if the court only had to decide a single case, it would neither be willing nor able to spend unlimited resources on finding the facts.

In this paper, we focus on a single source of uncertainty: the court knows that she does not know all the facts. Unless the parties precisely and correctly anticipate how the court will deal with this uncertainty, the uncertainty for the court translates into uncertainty for the parties (about the outcome of the trial). Frequently, the parties face a second source of prediction error: unless the case is exceptionally well defined, the court has some leeway in interpreting the applicable law; as a rule, law is ambiguous. We bracket this second source of uncertainty. This seems legitimate given we are not genuinely interested in prediction error for the parties, but in judicial error. On the European Continent, we could additionally rely on the rule derived from Roman law: *iura novit curia*, the court is supposed to know the law. Even if an individual judge is uncertain about the correct interpretation of the law, she is bound to come up with an interpretation.

2 The Myopic Judge

2.1 The Judge's Goal

As stated, we define a judge as being myopic if she ignores the possibility that the parties will anticipate her decision rule, and try to exploit it to their advantage. Such a judge (exclusively) faces an epistemic problem. In generic form, the problem can be characterized the following way:

Table 1: The Epistemic Problem

	acquit	convict
innocent	correct negative	false positive
guilty	false negative	correct positive

We want to contribute to a normative debate. This is why we assume that the judge is not biased, so that the utility from acquittal of an innocent defendant $u(d = 0|g = 0)$ is the same as $u(d = 1|g = 1)$, where d stands for the decision, that can either be acquittal (0) or conviction (1), and where g stands for guilt, which can either be given (1) or absent (0). In a private law dispute, plaintiff's claim can be well-founded ($g = 1$) or ill-founded ($g = 0$), and the decision d can be in favor of plaintiff (1) or defendant (0). For simplicity, we normalize the utility from correct decisions to 0. We thus assume $u(d = 1|g = 1) = u(d = 0|g = 0) = 0$. This implies: the judge is not concerned about the cardinal difference in utility between correct and false judgement. All she cares about is error. An unbiased judge dislikes passing false judgement. We must therefore have

$$\begin{aligned} u(d = 0|g = 0) &= u(d = 1|g = 1) > u(d = 0|g = 1) \\ u(d = 0|g = 0) &= u(d = 1|g = 1) > u(d = 1|g = 0) \end{aligned} \quad (1)$$

We consider two regimes, which differ by the standard of proof. If the standard is *preponderance of the evidence*, as typically in US civil law disputes, we have

$$u(d = 0|g = 0) = u(d = 1|g = 1) > u(d = 0|g = 1) = u(d = 1|g = 0) \quad (2)$$

Ideally, the judge just wants to minimize error. Contrast this with the standard *beyond a reasonable doubt*. It is characteristic for criminal law trials, and leads to

$$u(d = 0|g = 0) = u(d = 1|g = 1) > u(d = 0|g = 1) > u(d = 1|g = 0) \quad (3)$$

Convicting an innocent is (much more) troublesome than acquitting a guilty defendant.² Recall the status of the argument. We are not making empirical claims about the utility functions

²On the European Continent, the legal order has an additional degree of freedom. It may empower or even oblige the judge to become herself engaged in factfinding, the more so the more the legal order is troubled by false positives. Since in this paper we are interested in the competing normative positions of Continental and US law, we bracket this additional dimension.

of actual judges. Our argument is normative. We want to reconstruct the normative choice between standards of proof assuming judges holding utility functions that are psychologically sound. Hence the utility functions presented in this section are the utility functions the legal order *wants* judges to hold.

Since legal orders are divided over the standard of proof in private law disputes, let us illustrate the issue with an example from this area of law. A person has been at a rock concert. The day after she feels pain in her ear and visits the doctor. The doctor diagnoses her with a cracked tympanum. The company that had organized the concert would be liable if the music had transgressed the legal noise limit. Yet there have been no noise recordings during the concert, so that the noise level is disputed. Should the court hold in favor of plaintiff or defendant?³

2.2 The Divine Judge

In this section we normalize the judge's decision cost to 0. Another way of distinguishing this section from section 4 is the additional assumption that the court's choice is only motivated by the case it currently deals with. We consider the moment after all the attainable facts have been presented to the judge, so that the judge does also not have reason to consider the parties' transaction cost. Yet we model a situation that is to be found in most trials: at least some facts remain disputed. In our running example, the noise level at the concert is contested. How would a divine judge decide in the face of such insurmountable uncertainty?⁴

First consider the *preponderance of the evidence* standard of (2). Such a judge is indifferent between false positives and false negatives. If the true state of the world z was known and the court decides for plaintiff, with probability q_z this decision is correct. With counterprobability $1 - q_z$, the court commits an error. Likewise with probability q_z the court commits an error if it decides against plaintiff. Under the preponderance standard, the court decides for plaintiff if $u(d = 1) = q_z > u(d = 0) = 1 - q_z$. This is the case if $q_z > \frac{1}{2}$.

Now in the courtroom, z is normally not known with certainty. The court has to consider more than one state of the world. At best the court knows the set of possible states of the world Z , and the posterior probability p_z attached to each of these states, given the evidence. The "divine" court must then work with the *expected* utility, given it either decides in favor of plaintiff or of defendant. Specifically the court has to solve

$$E(d = 1) = \sum_{z=1}^Z p_z q_z > E(d = 0) = \sum_{z=1}^Z p_z (1 - q_z) \quad (4)$$

Solving the inequality, one finds that the court decides for plaintiff provided $\sum_{z=1}^Z p_z q_z > \frac{1}{2}$.

³The German Supreme Court held for plaintiff, stipulating a duty to make such recordings, *BGH* March 13, 2001, *Neue Juristische Wochenschrift* 2001, 2019.

⁴We are not interested in the way a judge uses the available evidence; one option would be Bayesian updating.

Let us illustrate the approach in an example. There are three possible states of the world. In state z_1 , it is quite likely that plaintiff has a case ($q_{z1} = 7/8$), but given the evidence the probability that this is the true state of the world is only $p_{z1} = 1/10$. In state z_2 , it is less certain that the claim is well founded, but the probability is still above chance $q_{z2} = 3/5 > 1/2$. This state of the world is more likely, $p_{z2} = 3/10$. The most likely state of the world is z_3 . Given the evidence, the probability that this is the true state of the world is $p_{z3} = 3/5$. If this is the true state of world, it is more likely than not that the claim is ill founded, $q_{z3} = 2/5$. If the court applies equation (4), her utility from deciding for plaintiff would be $203/400$, which is slightly more than her utility from deciding against plaintiff $197/400$. Plaintiff wins.

Yet the example demonstrates why this decision rule is unlikely to be normatively acceptable. Defendant would lose although the court leaves open what happened! Legal orders exceptionally make this possible, for instance if it cannot be decided whether a person has stolen a good, or trades the good knowing that someone else has stolen it. But normally procedural justice requires that the court specifies which alleged set of facts (in our language: which state of the world z) is sufficiently likely to be the factual basis of its decision. There are two formal ways to address this concern. The stricter requirement would be $d = 1$ if $\exists p_z q_z > \frac{1}{2}$. Yet this decision rule would favor defendant, the more so the larger Z . This might be considered to run counter the spirit of the preponderance standard. A less strict criterion would be $\exists p_z q_z | (q_z > \frac{1}{2}) > \forall p_z q_z | (q_z < \frac{1}{2})$: the largest product of $p_z q_z$ supporting plaintiff's case is larger than any product supporting defendant's case.

For a "divine" judge, deciding our real world example would be easy. There are only two potential states of the world. In state z_1 , the noise level surpassed the legal threshold. In state z_2 , this was not the case. Since there is the legal threshold, seemingly defining q_{zi} is also straightforward: if the threshold is surpassed, the probability that the claim is well-founded is (close to) 1. The court could concentrate on its (subjective) assessment of p_{zi} . This is how common lawyers would present the case.

With the *beyond a reasonable doubt* standard (3), all that changes is the decision rule. There is lively debate in the legal literature whether one should at all quantify the standard, and if so, how (Engel, 2009). For the purposes of this paper, we use the simplest approach. We again start with the counterfactual assumption that the state of the world z is known. The court deems convicting an innocent to be a much graver error. The court reflects this by positive weight $k > 1$ on acquitting defendant, so that the decision rule becomes $u(d = 1) = q_z > u(d = 0) = k \cdot (1 - q_z)$. Solving the inequality, conviction now requires $q_z > \frac{k}{1+k}$. In the realistic case where the court only knows the probabilities p_z with which several possible states of the world z obtain, the weighting factor k qualifies the right hand side (4).

2.3 The Human Judge

There are two very different reasons why the alternative approach modeled in this subsection might be preferable. The first reason results from a possible gap between positive and

normative theory. In principle, normative theory may of course formulate a goal, even if circumstances make it unlikely that the goal be attained. In the next step, one may then want to define "second best" outcomes, i.e. the best outcomes attainable, given some limitation. This is, for instance, the standard approach in (economic) contract theory where the necessity of incomplete contracts and the prevalence of private information allows at best optimization subject to constraints. By the same token, one might consider cognitive limitations as constraints. One might even want to empirically estimate how far the decision making abilities of a typical judge are off the normative mark. The adage "judex non calculat" might make the analyst suspect that the gap between normative theory and judicial practice is not small.

Yet there is a second reason for an alternative approach to modeling judicial choice that carries more weight. The theory developed above implicitly assumes that the state space Z is well defined, and that the judge precisely discerns and evaluates all its elements z . Continental lawyers would object: the seeming simplicity of the noise case only holds in the context of representation, not in the context of discovery. Initially plaintiff comes to court with a story about a concert and her illness. It is only after considerable interaction that the decision problem can be narrowed down to the issue of the legal threshold. Therefore originally the court has to consider a substantially larger set of potential states of the world Z : what else might have caused the illness? How long did plaintiff attend the concert? Which band did use which loudspeakers? In which ways did the organizer discipline the bands? Was plaintiff at risk, and should have protected herself? All of these factors could potentially be combined in multiple ways.

Of course, a hardnosed expected utility theorist can even handle a state space that is not well-defined. She will not only replace objective by subjective probabilities, but will also replace the objective state space Z by the subjectively expected one $E(Z)$. But even for a theorist, this is a difficult object. In the most simple case, she assumes the state space to be unidimensional, discrete, and all elements equidistant. Then the only subjective (decisionistic) element she needs is setting \hat{z}_{\min} and \hat{z}_{\max} . Yet not so rarely, these assumptions are not plausible. Then the theorist must further make assumptions about the distribution function $E(f(z))$. Which is its functional form: uniform, normal, exponential, gamma? What about the moments of the distribution: its central tendency, its variance, its skewedness, its kurtosis?

Most importantly: in hardly any legal case the uncertainty is plausibly unidimensional. In our example, there is potentially uncertainty about (1) possible causes of the illness, (2) the organization of the concert, (3) the specific circumstances of plaintiff's presence at the concert, (4) plaintiff's idiosyncratic vulnerability. Of course, technically multidimensional uncertainty can be handled. But it is not only unrealistic to assume that any "real" judge would do so, in whatever approximate way. The more complex the uncertainty, the more pronounced also the risk that the resulting estimate is far off reality. In such complex estimations, small changes in a single parameter may have big (and not so rarely also surprising) effects on outcomes.

In the following, we present an alternative modeling strategy. It is more "human" in a positive and a normative sense. On the positive side, it need not make utterly unrealistic assumptions about the decision making process, and thereby has greater descriptive validity.

On the normative side, it reacts to the fact that most legal problems are ill-defined, and presents a normative model for, nonetheless, taking decisions that are "good enough".

2.3.1 Scenario Construction

If the state space is continuous, a decision maker determined to exhaust the state space would have to consider infinitely many situations. The only way to do so effectively would require translating the characteristics of the state space into a function, and maximizing this function with respect to some decision criterion. If the state space is ill-defined, this exercise may easily lead to gross errors. A small misspecification of the function may translate into choices that turn out to be very inappropriate once the uncertainty dissolves. Modeling judicial decision-making as optimization under constraints is also empirically far off the mark. A robust psychological literature demonstrates that real judges decide by comparatively assessing competing "stories" that can be constructed, based on the available evidence.

Most of this evidence has been generated testing mock juries. Jurors are engaged in sense-making (Pennington and Hastie, 1991, 519) and constructing coherent stories from the evidence. Decision-making seems to be often interpretative (Pennington and Hastie, 1988, 524). It relies on reasoning about the evidence, rather than an algebra like process (Pennington and Hastie, 1988, 531). Jurors attempt at creating a narrative story from the pieces of evidence they have heard (Pennington and Hastie, 1986, 521), (Pennington and Hastie, 1988, 521), (Pennington and Hastie, 1993a, 194), (Pennington and Hastie, 1993b, 136), which can be considered as a mental model (Pennington and Hastie, 1988, 521), (also see Johnson-Laird, 1983).

Jurors decide by matching stories to the representation of the verdict categories given to them in the judge's instructions on the law (Pennington and Hastie, 1986, 243). Whether a story is accepted, or whether it is selected, depends on its goodness of fit (Pennington and Hastie, 1993a, 201). In this assessment, jurors look out for coverage, coherence and uniqueness (Pennington and Hastie, 1992, 190). Coverage, coherence and uniqueness also determine the level of confidence (Pennington and Hastie, 1988, 521, 528) (Pennington and Hastie, 1992, 190). Modelling the starting point of judicial decision making as scenario construction may also be motivated with a piece of doctrine: plaintiff has to present her "theory of the case".

The underlying process can be modeled as first constructing and then evaluating a limited set of scenarios $Z = \{z_1, z_2, \dots, z_n\}$. The theory of bounded rationality posits that a decision maker will only consider scenarios that "she does not dare to neglect" (more on the theory, and experimental tests from Gueth, 2013; Crosetto et al., 2015). This claim not only fits court procedure very well. Actually in court the process of scenario construction is even institutionalized. Legal orders differ in the precise design of this process. In the common law countries, the basic principle is adversarial. The court (i.e. the judge and the jury) only evaluate what the parties have presented to them. The parties are not just presenting pieces of evidence. They also "plead" their case, i.e. they tell a story. Plaintiff or prosecution always do. Defendant has a strategic choice to make. She either may tell a counter-story, or try

to cast doubt on the coherence of plaintiff's story. Note the power of the intervention. If the case is, as usual, essentially ill-defined, theoretically the number of possible scenarios is infinite. Yet procedure reduces the number of scenarios to a maximum of 2.⁵

In the civil law countries, court procedure keeps inquisitorial elements. The task of the judge is not confined to assessing the claims made by the parties. To a larger or smaller degree, she is actively seeking "the truth" (see e.g. §§ 138 f. German Code of Civil Procedure). In these legal orders, the judge must potentially herself create additional scenarios. Problem construction is part of judicial activity. In these legal orders, the constraints that make this process practical stem from the issue as defined by the parties, and substantive law.

Even if the judge is free to demand additional evidence, she cannot change the issue. It is fixed by the claim (in civil procedure) or the charge (in criminal procedure). Consequently scenarios are irrelevant that could not support the claim or charge, given the substantive legal rules governing the case. The judge will only consider alternative scenarios that might justify the intervention sought by plaintiff. In legal practice, this is an iterative process. The facts presented by the parties direct the search for potentially applicable legal rules. The conditions formulated in these rules direct the search for additional facts.

In principle, in the civil law countries the court is done if a story that matches the conditions of a relevant rule has been established, or if no such story can be found. Yet often judges also assess the plausibility of a story that would support the case by alternative stories that seem at least as plausible, but do not support the claim ⁶. This technique may also be used in a legal order following the adversarial principle. If the technique is employed, the set of relevant scenarios Z grows again. Yet it still remains bounded. Such counter-stories are only decision relevant to the extent that they are strong enough to cast sufficient doubt on the plausibility of a story that supports the claim.

2.3.2 Removing Implausible Scenarios

The set of scenarios only tells a decision maker which possible situations she has to consider. In the next logical step, she must qualify the relevant scenarios by their plausibility. In court, this element of boundedly rational decision making is also institutionalized. The standard of proof creates a relevance criterion. If the standard of proof is *beyond a reasonable doubt* and the scenario would in principle support the claim or charge, the judge must immediately discard it if "reasonable doubt" remains. It is at this (logical) point of the decision making process that the parties introduce evidence. If the legal order makes this possible, this is also when the judge, on its own initiative, gathers additional evidence, or calls for it.

The decision making process is more involved, though, if defendant has told a counter-story, or if an inquisitorial judge has created counter-scenarios on her own initiative. If the story and the counter-story were well defined and mutually exclusive, it would of course suffice to

⁵Unless defense decides to present multiple inconsistent counter-stories.

⁶There is, of course, a risk involved due to judges who after accepting one story told as reasonable, are reluctant to questioning it again (such reluctance would, for instance, be suggested by the theory of cognitive-dissonance, see (Festinger, 1957))

show that a counter-scenario is more than minimally plausible. That would logically imply: plaintiff's story does not pass the test. Yet if the judge knows the set of scenarios to be incomplete, the fact that one counter-story has a certain degree of plausibility is not more than a source of doubt. The judge is torn between alternative interpretations, and must comparatively assess the weight of wrongly discarding either of them.

In the spirit of finding the scenarios a decision maker "does not dare to neglect", satisficing theory posits that the judge need not define a subjective probability \hat{p}_z of some scenario being correct. Rather the judge (subjectively) assesses whether the likelihood l_z of the scenario being true is above some lower bound \underline{l} . This element gives the decision making process a lexicographic flavor: scenarios that do not pass the threshold \underline{l} are discarded altogether. As in maximum likelihood estimation, the likelihoods of several scenarios need not add up to 1. This allows for subadditivity and superadditivity: in the first case several scenarios are sufficiently probable, and in the second case all scenarios are not sufficiently probable. Note that competing scenarios may even be mutually exclusive, but still sufficiently likely to be considered. This may, for instance, be the case if both plaintiff and defendant offer plausible, but contradictory, interpretations of the facts. This does not lead to contradictory final outcomes. This part of the exercise is not definitive. All the judge aims to achieve at this point is reducing the set of scenarios. Scenarios are removed that would have been qualitatively relevant, but that are too unlikely to be supported by the evidence.

The distinction of probabilities and likelihoods also resonates with the legal debate on scientific evidence in court. The *Daubert*-rule makes the judge a gatekeeper for the admissibility of scientific evidence (*Daubert v. Merrell Dow Pharmaceuticals*, 509 U.S. 579 (1973)). The *Reference Manual on Scientific Evidence* (3rd), xiv stresses that the judge "must decide after considering the best available science", knowing that it tends to be imperfect.

2.3.3 Acceptable Outcomes

We keep the assumption that the judge is unbiased. If $g = 1$, the judge would want to decide for plaintiff $d = 1$. Likewise, if the claim is unfounded $g = 0$, the judge would want to decide against plaintiff $d = 0$. By implication, the judge suffers if she commits an error. Again the judge is content with satisficing. She does not aim at picking precisely the one scenario with the highest overall satisfaction. All she is concerned with is choosing a scenario that leads to a gain/loss balance that is "good enough". In finding the set of decisions D that meet this criterion, the judge is again guided by procedural rules. Usually, the law does not want the judge to be guided by potential gains and losses for the parties.⁷ All the judge shall be concerned with is error. This considerably simplifies the analysis. The judge need not (and should not for that matter) estimate how much the parties stand to gain or lose. Consequently the normative weight of basing choice on any (sufficiently plausible) scenario is the same (and can therefore be normalized to 1). That implies a very simple procedure for defining the set of acceptable outcomes D .

⁷This can be different if the court has power to provisionally govern the real life conflict until the case has been definitively decided, as in an injunction.

First consider a civil law dispute under US law, i.e. under the *preponderance of the evidence* standard. At this point, the distinction between probabilities and likelihoods is critical. It becomes most important if neither scenario on its own has a likelihood $l_i > .5$, but if there is more than one scenario that meets the threshold $l_i > \underline{l}$. The most plausible interpretation of the standard of proof would ask the judge to take the decision mandated by the most likely scenario, $d_i | l_i > l_j, \forall j$. This introduces a second lexicographic element into the decision function.

Now when taken literally, this standard of proof would force a sharp line, built on an exact cardinal definition of likelihoods. This would run counter the spirit of the satisficing approach. One way of accommodating this objection modifies the decision rule. The judge is only required and entitled to decide in favor of d_i if $\Delta(l_i - l_j) > \underline{\Delta}, \forall j$. The likelihoods must be sufficiently distinct, with sufficiency following from some judicial norm. This approach is similar in spirit to the role most practitioners assign to probability thresholds in statistics. They do not take p-values at face value, and rather check whether the p-value is below some conventional level (usually 5 %), and only note whether the p-value is even below a more stringent conventional level (usually 1 %) , and speak of "significance at conventional levels" and "highly significant", respectively. Coarsening the evaluation avoids assigning meaning to excessively precise numerical estimates.

Note the critical role of \underline{l} . The stricter the legal order (or the individual judge in applying the law, for that matter) is in setting the standard \underline{l} , the more often the decision must be taken by burden of proof: the evidence is insufficient to make any definite claim about the events. By default, the claim (in civil disputes) or the charge (in criminal disputes) is dismissed. On the opposite end, the more the standard \underline{l} is lenient, the more often *both* a scenario l_i that supports conviction *and* a scenario l_j that supports acquittal will be established to the requisite standard. In that event, the decision must again be taken by burden of proof. The claim or charge is dismissed, not because it could not be established, but because the evidence in favor of conviction is ambiguous. The burden of proof has a further noteworthy effect. When non-legal satisficing decision makers face undissolvable uncertainty, they may react by lowering their aspiration level (the idea goes back to Sauermann and Selten, 1962). The legal order excludes this option, and rather wants the judge to decide by burden of proof whenever the stable threshold defined by the standard of proof is not passed. The case is decided (by dismissing it), but not on its merits.

2.3.4 Decision Making

The advantage of the satisficing approach is that the subset of decisions $\dot{D} \subset D$ increases where the judge is indifferent. By a basic rule of procedural law, indifference is resolved to the detriment of plaintiff. Our formalization thus makes the actual normative decision visible. The smaller \dot{D} , the more the inevitable indeterminacy resulting from the fact that the judge is a satisficer is resolved to the detriment of defendant, and vice versa. Hence the decision

rule is given by

$$d = 1 \quad \text{if} \quad (i) \quad l_i - l_j > \underline{\Delta}, \forall j \quad \text{and} \quad (ii) \quad g_i = 1$$

$$d = 0 \quad \text{otherwise} \quad (5)$$

In words: plaintiff wins if there is a scenario that backs her claim ($g_i = 1$) and is sufficiently more likely than any alternative (counter) scenario. Note that this decision rule diverges from the decision rule (4) used by the "divine" judge. The judge no longer minimizes the risk of a wrong decision, but is content with a decision that is "good enough", in the sense of sufficiently reducing the risk of a false positive, or a false negative, decision.

This decision rule implies that a case may be dismissed for one of three reasons: (1) the decision set is empty, $\dot{D} = \emptyset$: no party has brought forward a scenario z that is sufficiently backed by facts (nor has the judge herself, if procedural rules empower her to do so). (2) the evidence is inconclusive, $D = \{d_1 = 1, d_2 = 0\}$. (3) defendant is proven to be innocent, $\dot{D} = \{d_1 = 0\}$. Note that (1) and (2) are decisions by burden of proof. The court does not judge on the merits of the case. Legal orders differ as to whether, in terms of the final ruling, there is a difference between (1) "not proven" and (3) "not guilty". In the comparative law literature, the latter is called the *Scottish Rule* (Duff, 1999) (more from Bray, 2005) (Leipold, 1999).

2.3.5 Beyond a Reasonable Doubt

It is straightforward to extend this approach to legal matters with a different standard of proof, and with the *beyond a reasonable doubt* standard in particular. The critical element of this standard is not the difference between the most likely scenario z_i and the next likely scenario z_j . The law is thus not content with $\underline{\Delta}_{\text{brd}} > \underline{\Delta}_{\text{poe}}$, where brd stands for beyond a reasonable doubt, and poe stands for preponderance of the evidence. Rather the law constrains the set of acceptable outcomes D : $l_i > \kappa | g_i = 1, \frac{1}{2} < \kappa < 1$.

Remains to determine the decision rule if more than one (relevant) scenario is sufficiently likely. There are two ways to go about this, and both exist in criminal law practice. If acquittal is all defendant cares about, the only case to consider is $l_i > \kappa, g_i = 1$ and $l_j > \kappa, g_j = 0$: there is a highly likely story supporting guilt and a highly likely counter-story speaking against guilt $g = 0$. As with the preponderance standard, this conflict must be resolved in favor of defendant.

If the legal order distinguishes "acquittal for cause" from "acquittal for lack of evidence", it must not only define $\underline{l}_{\text{convict}}$, but also $\underline{l}_{\text{acquit}}$. In criminal procedure, the normative rationale of the *beyond a reasonable doubt* standard is the presumption of innocence. This constitutionally protected value suggests $\underline{l}_{\text{acquit}} \ll \underline{l}_{\text{convict}}$: acquittal for cause does not require that the counter-story be extremely likely. It suffices if it meets a considerably more modest standard.

2.4 Implications for the Normative Debate

For a "divine" judge, decision making in the face of undissolvable uncertainty is a mechanical exercise. If an outside observer is given access to the court's subjective probabilities, and to its subjective definition of the state space, she will come to the same conclusion. By contrast, decision making by a "human" judge, in the sense modeled in this paper, is decisionistic in a much deeper sense. At all steps (scenario construction, removing implausible scenarios, defining acceptable outcomes, choosing among acceptable outcomes) the court is engaged in a way that is not, and cannot be predetermined by procedure. The parties, and the legal order, are at the mercy of these human decision makers who happen to hold the office of a judge, or who have been elected as jury members. In a much deeper sense than suggested by the utterly unrealistic model of a "divine" judge, judges exercise power. They are engaged as subjective persons. This gives more weight to concerns that motivate the position of Continental European law. The purpose of the standard of proof is not confined to defining a decision criterion, and thereby distributing the risk of error among the parties. The standard of proof also, and critically, informs the judge (as a fallible human being) about the normative expectations the legal order has about her idiosyncratic way of deciding the case. Specifically, the *beyond a reasonable doubt* standard tells the judge: "We know you must decide despite imperfect knowledge about the case. The legal order will accept and enforce your decisions. But it cares about the impression this leaves on society. Do not shy away from deciding. But if you have serious doubts about a claim, do not force a decision on the merits. Rather decide by burden of proof, and thereby immunize the legal order against the reproach of material error".

2.5 Discussion

Our model deliberately stays close to the standard rational choice approach. We drop the utterly unrealistic assumption that a judge or jury member "rationally deals with its own irrationality" (resulting from understanding that the evidence is objectively limited). But we do not deviate from rational choice further than necessary. This motivates the term bounded rationality for our concept of satisficing judges (for background see Gueth, 2013; Crosetto et al., 2015). Specifically our approach remains consequentialist: when choosing how to handle a case, the judge engages in forward looking reasoning, and considers the likely effects of her action on her ex ante defined utility. It may well be that judges out there in court do not deliberate at all, but content themselves with recognizing cues they have experienced to discriminate sufficiently well (see Dhimi and Ayton, 2001; Gigerenzer and Engel, 2006). It is not unlikely either that real judges are also guided by foreseeable reactions of the parties, or the likely decision of superior courts. We do not deny these effects. But they do not matter for our normative endeavor. We do not want to describe or predict courts. We want to help the legal order understand the implications of choosing a standard of proof.

3 The Prevoyant Judge

3.1 The Jugde's Goal

After German reunification, a pessimistic sentence has often been heard: we wanted justice, and we have gotten the rule of law. Cynical suspects not so rarely tell the police: if you believe I am guilty, prove it. Even a "divine" judge may be the target of strategic action by criminals or defendants, for that matter. The more stringent the standard of proof, the easier strategic concealment. A perpetrator gets away with her illegal action. Anecdotal evidence has it that not so few judges get cynical themselves, after having seen too many cynical defendants. Defendants are lying anyway. Let's see how we can frame them. While such reactions are a worthy topic for empirical research, they are clearly normatively undesired. Judges are expected to abide by procedural rules, even if this means that they have to acquit many defendants whom they suspect to be guilty.

Rule of law thus severely limits the room for strategic judicial action. Yet such action is not completely outlawed. The allocation of the burden of proof is a practical application. An illustration is the *res ipsa loquitur* rule, which plays a major role in medical malpractice trials. The judge is entitled to instruct the jury that harm suffered by plaintiff is caused by defendant's negligence provided (a) an event of this kind does not ordinarily occur without negligence, (b) competing explanations may be reasonably ruled out and (c) the indicated negligence is within the scope of defendant's duty towards plaintiff. If the judge relies on this power, the burden of proof shifts to defendant (*Restatement of Torts, Second, § 328D*). Knowing that, otherwise, she loses, this reallocation of the burden of proof forces defendant to reveal her private information.

We keep the definition of the judge's goal. The judge is unbiased (1), and either applies the *preponderance of the evidence* standard (2) or the *beyond a reasonable doubt* standard (3), as stipulated by the law. But we now assume the judge to take into account that either party might aim at biasing the court ruling in her favor.

3.2 The Divine Judge

We first consider a prevoyant and divine judge, i.e. a judge that optimally reacts to the prevailing uncertainty about the true state of the world z . We assume that at least one party, plaintiff or defendant, knows the state of the world with certainty. We introduce the strategic component through the fact that the informed party sends a signal s which we assume to be positively, but imperfectly correlated with z . Since the party possesses private information, the divine judge exploits her statutory discretion to write a direct mechanism (more from Gueth, 2015). Formally, if the judge expects defendant to possess critical private information, the judge commits to a measure m such that defendant's utility v is given by the incentive compatibility constraint

$$v(g = 1, s = 0, m = 0) < v(g = 1, s = 1, m = 1) \tag{6}$$

It is in the best interest of defendant to send a true signal $s = 1$ if she actually is guilty, and to send signal $s = 0$ only if she is truly innocent. The judge capitalizes on her procedural discretion to elicit defendant's private information. In criminal procedure, this approach might not survive constitutional scrutiny. Arguably it violates the maxim *nemo tenetur seipsum accusare*, which US law protects in the Fifth Amendment to the Federal Constitution: no (criminal law) defendant is obliged to contribute to her own conviction. Yet depending on the technology the judge uses, the constitutional verdict may be avoided. And in private law disputes, constitutions tend to give the courts more freedom in the first place. By analogous intervention, the court may aim at inducing prosecution or plaintiff to reveal their private information. Note that the judge need not worry about participation; she exploits the sovereign powers bestowed upon her and forces the party(s) to accept the mechanism. Budget balance is also not an issue since procedural law is concerned with error, not with welfare.

In the simplest case, the set of possible mechanisms M coincides with the set of possible final outcomes D . The judge credibly threatens the party that holds relevant private information with losing in court unless she reveals this information. Shifting the burden of proof is a measure of this kind. If $M = D$ holds, the action space of the myopic and of the prevoant judges is the same. It is defined by the binary choice between acquittal ($d = 0$) and conviction ($d = 1$). Many practical applications are of a different kind. The judge threatens a party reluctant to reveal her private information with exercising her procedural discretion in an unfavorable way. The judge might, for instance, define the order in which witnesses are heard (and thereby the jury is influenced), or order the jury to ignore a piece of evidence when this is a borderline case. Then the threat power stems from the fact that a verdict becomes more or less likely: $\text{Prob}\{d = 1\} = f(m)$. In criminal procedure judges routinely take it into account at the sentencing stage if defendant has pleaded guilty, and has helped the court investigate the case. If they do, formally they exploit the fact that the game is more complicated. The court actually takes two decisions: d_g , the binary decision about conviction or acquittal, or guilt. Provided $d_g = 1$, a second decision $d_s \in [\text{sentence}_{\min}, \text{sentence}_{\max}]$ is taken. The judge exploits $d_s = f(m)$ to force defendant into revealing her private information.

3.3 The Human Judge

A direct mechanism only delivers if the designer gets the problem right, including the determination of addressees to maximize their personal benefit given whatever constraints they face. This challenge is familiar from the debate over penalty defaults (i.e. direct mechanisms) vs. majoritarian defaults (Ayres and Gertner, 1999). In criminal and civil procedure, it is all the more relevant. On the one hand, professional advice may bring behavior in court closer to outcome maximization. Yet on the other hand, to a non negligible degree the judicial system relies on defendants exhibiting honesty and remorse. The judicial system is also concerned about the impression court procedure makes on outsiders who might be tempted to commit similar infractions. If judges made it plain that they try to coerce defendants, these valuable social resources might deplish. Last, but not least, judges are no trained mechanism designers. A normative model that expects them to do so may therefore ask for too much. For

all these reasons, also when taking into account that defendants might act strategically, a normative model is more appropriate that assumes a "human" judge, in the sense defined in section 2. Such a model has the same building blocks as a normative model for a myopic, human judge.

3.3.1 Scenario Construction

If a judge assumes defendant not to act strategically, the only source of uncertainty is stochastic. It results from incomplete knowledge about the true state of the world z . If the judge drops this (sometimes) heroic assumption, her problem becomes more complicated. A scenario now has (at least) two elements: the true state of the world z , and the strategic action of defendant s . By the design of court procedure, this signal typically is a vector, consisting of defendant's explicit statements, the means of evidence she offers, and the impressions she tries to convey through these means. Yet for the purposes of our model, it suffices to compress this vector into a single signal s . It may be interpreted as the overall probability of guilt as derived from exclusively considering all the action by defendant or her representatives. Note that, again, procedural law considerably simplifies the problem. Information is decision relevant only if it has been formally introduced into procedure. Thereby s is quite clearly defined. In normal court procedure, along with defendant's signal s_d only one more signal has to be taken into account by the jury and the judge: the signal s_p sent out by prosecution or plaintiff. It is constructed analogously. Hence a complete scenario consists of three elements $\{z_i, s_d, s_p\}$. Note that signals are not indexed by states of the world i . The two signals are perfectly observed. The problem is what they tell the court about which of the possible states of the world Z is the true one.

3.3.2 Removing Implausible Scenarios

The same way as a myopic judge, a satisficing prevoyant judge does not consider all theoretically conceivable scenarios. She does also not try to estimate an exact subjective probability \hat{p}_z of each scenario being true. Rather she confines herself to checking whether the likelihood l_z of the scenario being true surpasses threshold \underline{l} . Actually the more complete definition of the scenario by a true state of the world z as predicted by two signals s_d, s_p makes it clear by which means the court arrives at this assessment. Yet the prevoyant judge departs from the myopic judge in that she does not take these two signals at face value. She considers the possibility that they might not only be false, but deliberately misleading.

3.3.3 Acceptable Outcomes

The critical difference between a myopic and a prevoyant satisficing judge lies in the construction of the set of acceptable outcomes. The myopic judge only considers the ultimate goal of minimizing error. The prevoyant judge balances this out with the proximate goal of inducing the parties to reveal their private information. The myopic judge treats all outcomes as acceptable that meet the criterion $\Delta(l_i - l_j) > \underline{\Delta}, \forall j$, where the likelihoods are constructed from

taking the evidence at face value. By contrast a prevoyant judge works with $l_i(s_d(m), s_p(m))$: the likelihoods are a function of the signals sent out by the parties which, in turn, are functions of the mechanism designed by the court. Moreover provided the set of mechanisms $M \supset D$, acceptable outcomes must be defined in more than one dimension. Hence the judge no longer works with the acceptable set for a myopic judge $A_m \in \{d = 0, d = 1\}$, but shifts towards the acceptable set for a prevoyant judge $A_p \in \{d, m\}$.

Of course, acceptability is a normative issue. If the procedure is criminal and the constitution is strict with *nemo tenetur seipsum accusare*, there is not much room for m ; the judge is essentially forced to live with defendant's strategic concealment (but may, at least, be able to coerce prosecution not to hide information). Even if the legal order is not opposed to eliciting private information of the parties as a matter of principle, the court faces a normative challenge. Mechanism design has been developed assuming that addressees optimize, interpreting the mechanism as a constraint. Empirically, many individuals do not think ahead, or only very imperfectly so (see only Arad and Rubinstein, 2012). Since courts tend to meet defendants only once, it is difficult for them to infer the level of sophistication of the concrete parties in front of them. Therefore the very fact that they employ a mechanism may itself bias judgment. Specifically courts may wrongly infer that $g = 1$ because defendant does not reveal information. Typically, all courts have is a rather diffuse sense of population composition, and maybe some second order signals about the sophistication of the concrete parties in front of them. In defining acceptable outcomes, they must balance this court-induced risk of false judgement with the court-induced chance of eliciting the parties' private information.

3.3.4 Decision Making

A myopic satisficing judge convicts defendant if (i) at least one acceptable scenario meets the decision criterion $l_i - l_j > \underline{\Delta}$ and (ii) according to this scenario defendant is guilty ($g = 1$). If a prevoyant satisficing judge is constrained by $M = D$, condition (a) still holds (with l_i and l_j being functions of the respective signals s_d and s_p), but condition (b) is replaced by (6). The judge may even convict if the scenario in question does not indicate guilt, but conviction is necessary to make defendant reveal her private information. Of course if defendant is rational in the standard sense, and if the mechanism is properly designed, this is an action off the equilibrium path. In anticipation, defendant reveals her private information. The judge need no longer decide in execution of the mechanism, but may decide based on reliable information about the true state of the world. If, by contrast, $M \supset D$ holds, the actual decision encompasses more than one element, possibly to be implemented before (in the case of procedural threats) or after (in the case of a threat with exercising sentencing discretion) the decision D about guilt.

3.4 Implications for the Normative Debate

The US approach to defining the standard of proof is motivated by reasoning that assumes rationality. Mechanism design also assumes rationality. In that sense, the concept of a prevoyant judge is in the spirit of the US approach. Yet it has been the purpose of this

section to explain in which ways a satisficing judge may think and act strategically. As we have seen, this is not an oxymoron. Moreover a satisficing judge has reason to consider the possibility that the parties in front of her are also satisficers. This does not exclude an approach in the spirit of mechanism design, but calls for additional caution. In the typical court room situation, the judge's problem will not be well-defined. The judge senses that the probability of her being strategically misled by one of the parties is non negligible. She considers taking strategic measures herself. But both in designing the intervention, and in interpreting the parties' reactions, she will need judgment. Such a cautious approach of course increases the risk that one of the parties frames the court. Yet if the court refuses to decide on the merits since it has not been convinced *beyond a reasonable doubt*, the court makes it transparent that normatively relevant uncertainty could not be removed. This concern speaks in favor of the more conservative Continental approach, even taking into account that court cases tend to be situations of strategic interaction between the parties and the court.

4 The Professional Judge

4.1 The Judge's Goal

A major difference between common and Continental law lies in the role of the jury. While Continental law also uses lay judges, on the Continent court procedure is not built on a strict division of labor between the jury and the judge. Even if laymen sit on the bench, on the Continent they are typically not called for a single case, but become part of the judiciary for an extended period of time. This institutional difference explains why an additional normative concern is patent on the Continent. The court does not only strive for minimizing error in the case at hand. It must also decide about the allocation of judicial resources between cases. The more time and energy a judge devotes to one case, the less she has at her disposition for other cases that compete for her attention. Taking into account that judicial resources are limited, courts face a tradeoff between minimizing error in the case at hand, and doing so in other cases on their docket. We refer to a judge who tackles this additional challenge as "professional".

Ultimately, common law jurisdictions cannot avoid this tradeoff either. It is only less visible. But if procedure takes endless time, the presiding judge cannot deal with other cases, the same way as her counterparts on the Continent. And jury members pay an opportunity cost in terms of lost time for work and leisure which the legal order may not want to become excessive.

4.2 The Divine Judge

For a "divine" judge (or jury), this additional concern introduces an investment choice. She must decide which fraction of maximum investigation effort \bar{C} to invest in the case at hand. For our purposes, it suffices to consider a second case 2 that competes with the current case 1 for resources. (1) can in generic notation be expressed as $u(d, g)$: the judge's utility

from deciding the case is a function of her decision d , and of the true state of the world, i.e. defendants guilt g . If judicial resources are limited, the judge's utility function has an additional element c , the cost, consisting of judicial effort devoted to the case. We assume $\frac{\partial u}{\partial c} > 0$: additional investment reduces the risk of false judgment, but $\frac{\partial^2 u}{\partial c^2} < 0$: investment has diminishing returns. On these assumptions, the court obviously chooses c_1 and c_2 such that $\frac{\partial u_1}{\partial c_1} = \frac{\partial u_2}{\partial c_2}$, $c_1 + c_2 = \bar{C}$: The marginal effect of reducing error in either case is the same, and the court does not hold back judicial resources.

4.3 The Human Judge

Iudex non calculat. This skeptical assessment may be overstated for simple arithmetic operations. But calculus is certainly beyond most professional lawyers, and out of the question for a jury. Yet again, ultimately such technical constraints are not essential. If the legal order cared about the optimal allocation of judicial resources to cases, it might endow courts with a decision aid. Since the mathematical operation is fairly simple, a spreadsheet would do the trick, which could be provided by the court administration. Again the critical issue is precision. For calculating the marginal effect of investing one more unit of court effort on the expected risk of convicting an innocent or acquitting a guilty defendant, one must specify the functional form of $u(d, g, c)$, and must fix parameters. The additional complexity resulting from the investment decision exacerbates the concerns that motivated a satisficing approach to judicial decision making in the first place. If courts tried to calculate the exact optimal effort for each case, they would quite likely be grossly misled by misspecifying the investment problem. A satisficing approach imposes itself *a fortiori*.

In principle, all steps of the satisficing process are effortful. Yet in stylized manner, one may distinguish deliberation cost c_d and fact finding cost c_f . In a typical court situation, c_f is a much bigger problem than c_d . More technically: it is very unusual that $c_{d1} + c_{d2} \leq C_d$ binds when $c_{f1} + c_{f2} \leq C_f$ is not binding. We therefore focus on c_f , and for simplicity just denote it with c . If, in an unusual case, c_d is critical, the analysis could be extended analogously.

If the adversarial principle applies (strictly), scenario construction is exclusively in the hands of the parties. Then the court does not have to invest any resources in this step of the procedure. If procedure has inquisitorial elements, the court is involved in scenario construction. Scenario construction is an exercise in sense making. It therefore requires deliberation cost c_d , not fact finding cost c_f . The final step of making a decision is mechanical: the court checks whether there is a scenario that supports conviction, and whether the likelihood of this scenario is sufficiently distinct from the likelihood of all competing scenarios. This activity again only engages deliberation cost c_d .

By contrast, the two remaining steps can be affected by the amount of fact finding cost c_f the court is willing to invest. The lower the relevance threshold \underline{l} , the more scenarios the court must consider. Discriminating among more scenarios requires a higher fact finding cost c_f . Consequently, the threshold taking the negative externality on competing cases into account \underline{l}_c is strictly more demanding than the threshold chosen by a court that only considers a single case: $\underline{l}_c > \underline{l}$. The same argument can be made for the definition of acceptable outcomes.

The bigger the normatively required gap $\underline{\Delta}$ between the accepted and the discarded scenarios, the less effort the court needs to discriminate between those scenarios that have surpassed the first hurdle \underline{l}_c . We thus also posit $\underline{\Delta}_c > \underline{\Delta}$.

Some jurisdictions, and some courts for that matter, are flooded with cases. In other jurisdictions, individuals rarely go to court. At some point in time, a certain type of cases piles up. Say the country reacts to a catastrophe somewhere else in the world by accepting refugees. Then the courts must discriminate between those who truly deserve asylum, and others who just want to come because they expect a better life. Courts must deal with a heavy caseload until this stream of refugees has been handled. At another point in time, the courts may face a fact intensive type of cases more frequently. If, for instance, the standard of care is defined more stringently, the number of medical malpractice suits may jump up. For such reasons, the opportunity cost c of dealing with the case at hand more intensely is not fix, but reacts to circumstances. This is a further justification for adopting a satisficing approach. Courts are not expected to get the (estimated) effect on (current and future) competing cases exactly right. All they try to avoid is a gross misallocation of judicial effort.

4.4 Implications for the Normative Debate

If a jurisdiction applies the *preponderance of the evidence* standard, a satisficing approach affects the criterion for decision making: $\sum_{z=1}^Z p_z q_z > \frac{1}{2}$ is replaced by $l_i - l_j > \underline{\Delta}, \forall j$. The only possibility to introduce sensitivity to fact finding effort c_f is through the definition of the two thresholds \underline{l}_c and $\underline{\Delta}_c$. Technically, this is no different if the jurisdiction applies the *beyond a reasonable doubt* standard; D_{poe} is replaced by D_{brd} . Yet the need to adopt a flexible threshold is easier to accept for such a legal order, given the standard had to be "reasonable" in the first place. All one has to do is interpret this as "reasonable, given circumstances". More importantly even: handling the conflict between multiple cases introduces an additional element of indeterminacy and hence power. This makes it even more advisable to constrain courts to "reasonable" choices, and to admonish them to take on personal responsibility.

5 Conclusion

Judges frequently have to decide even though they know full well that they have not completely understood the case. If such undissolvable uncertainty remains, judges must choose between the risk of a false positive ruling (prosecution or plaintiff win although defendant is innocent) and a false negative ruling (defendant is acquitted although she is guilty). The standard of proof assigns a weight to either error. Legal orders basically agree that in criminal procedure false positives are very troublesome. This motivates the *beyond a reasonable doubt* standard. By contrast, legal orders disagree about the standard of proof in private law disputes. In such disputes, US law assigns the same weight to either risk, and consequently adheres to a *preponderance of the evidence* standard. Continental law, however, applies the *beyond a reasonable doubt* standard in private law disputes as well. In this paper, we aim at reconstructing this normative choice in the light of alternative behavioral assumptions about

judicial decision making. We argue that the US approach implicitly assumes a judge who holds empirically implausible cognitive abilities. We contrast such a "divine" judge with a more plausible "human" judge. Specifically we model decision making in the face of undissolvable uncertainty by a judge who acts as a satisficer. Note that this is neither an empirical paper nor an empirical argument. Rather, based on extant empirical evidence about judicial decision making, we argue that a satisficing approach yields a more appropriate understanding of the *normative* issue.

In the second section, we extend this approach to a "prevoyant" judge who takes into account that the parties might aim at purposefully biasing her. Even taking this possibility into account, the satisficing approach remains appropriate, as does the normative reaction to adopt the *beyond a reasonable doubt* standard in civil disputes as well. In the final, third, step we consider a "professional" judge, whom we define as a judge taking into account that judicial resources are limited, and have to be divided between cases. From this angle, the cautious Continental approach to private law disputes becomes even more appealing. The term "reasonable" is flexible enough to capture the concrete circumstances. And decision making is bounded away from any excessively precise assessment, while focusing on the inevitable indeterminacy of judicial action, and admonishing the individual judge to exercise this discretion responsibly.

This paper is meant as a contribution to the normative legal debate, and to the understanding of one of the not so many real differences between US and Continental law. Yet in the interest of making this contribution, we had to translate rules of procedural law and observations from court practice into a model of judging by a satisficer. Now judging is a basic aspect of human behavior and, more narrowly, of forward deliberation in order to select among choice options by anticipating their likely consequences. And like the professional judge one hardly ever can be perfectly sure to have taken into account all circumstances. Actually when judging the risk of financial investments it is common practice to assess "the value at risk". This term means that one has rather accurate risk assessments for 98%-probability events, but leaves completely unspecified what happens in the remaining 2% cases. This is similar in attitude to common practice in empirical research: one treats a hypothesis as not disconfirmed when p-values do not exceed 10%, 5% or 1%. Acknowledging that, ultimately, one can never be sure, one is nonetheless willing to pass judgment. Such analogies make us wonder whether our model of judicial decision making might actually have a wider field of application. But we only flag this possibility out, and leave the serious analysis of such potentially similar tasks to future work.

References

- Ayala Arad and Ariel Rubinstein. The 1120 money request game. a level-k reasoning study. *American Economic Review*, 102(7):3561–3573, 2012. [18]
- Ian Ayres and Robert Gertner. Majoritarian vs. minoritarian defaults. *Stanford Law Review*, 51:1591–1613, 1999. [16]

- Samuel Bray. Not proven. introducing a third verdict. *University of Chicago Law Review*, 72: 1299–1329, 2005. [13]
- Kevin M. Clermont and Emily Sherwin. A comparative view of standards of proof. *American Journal of Comparative Law*, 50:243–276, 2002. [3]
- Paolo Crosetto, Werner Gueth, Luigi Mittone, and Matteo Ploner. Equity seeking and punishment in a public good experiment, 2015. [9, 14]
- Mandeep K. Dhami and Peter Ayton. Bailing and jailing the fast and frugal way. *Journal of Behavioral Decision Making*, 14:141–168, 2001. [14]
- Peter Duff. The scottish criminal jury. a very peculiar institution. *Law and Contemporary Problems*, 62:173–201, 1999. [13]
- Christoph Engel. Preponderance of the evidence versus intime conviction. a behavioral perspective on a conflict between american and continental european law. *Vermont Law Review*, 33:435–467, 2009. [3, 7]
- Leon Festinger. *A Theory of Cognitive Dissonance*. Row Peterson, Evanston, Ill., 1957. [10]
- Gerd Gigerenzer and Christoph Engel, editors. *Heuristics and the Law*. MIT Press, Boston, 2006. [14]
- Werner Gueth. Satisficing players. *Research in World Economy*, 4(1):1–13, 2013. [9, 14]
- Werner Gueth. Mechanism design and the law. In Francesco Parisi, editor, *Oxford Handbook of Law and Economics*, page forthcoming. Oxford University Press, Oxford, 2015. [15]
- Philip N. Johnson-Laird. *Mental Models. Towards a Cognitive Science of Language, Inference and Consciousness*. Cambridge University Press, Cambridge, 1983. [9]
- Andrew D. Leipold. The problem of the innocent, acquitted defendant. *Northwestern University Law Review*, 94:1297–1346, 1999. [13]
- Nancy Pennington and Reed Hastie. Evidence evaluation in complex decision making. *Journal of Personality and Social Psychology*, 51:242–258, 1986. [9]
- Nancy Pennington and Reed Hastie. Explanation-based decision making. effect of memory structure on judgement. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 14:521–533, 1988. [9]
- Nancy Pennington and Reed Hastie. A cognitive theory of juror decision making. the story model. *Cardozo Law Review*, 13:519–557, 1991. [9]
- Nancy Pennington and Reed Hastie. Explaining the evidence. tests of the story model for juror decision making. *Journal of Personality and Social Psychology*, 62:189–206, 1992. [9]
- Nancy Pennington and Reed Hastie. The story model for juror decision making. In Reed Hastie, editor, *Inside the Juror. The Psychology of Juror Decision Making*, pages 192–221. Cambridge University Press, Cambridge, 1993a. [9]
- Nancy Pennington and Reid Hastie. Reasoning in explanation-based decision-making. *Cognition*, 49:123–163, 1993b. [9]

Heinz Sauer mann and Reinhard Selten. Anspruchsanpassungstheorie der unternehmung. *Journal of Institutional and Theoretical Economics*, 118:577–597, 1962. [12]

Michele Taruffo. Rethinking the standards of proof. *American Journal of Comparative Law*, 51:659–677, 2003. [3]