Information Acquisition and the Exclusion of Evidence in Trials

Benjamin Lester　　　　　　Nicola Persico
University of Western Ontario　　New York University

Ludo Visschers
Simon Fraser University

March 29, 2008

Abstract

A peculiar principle of legal evidence in common law systems is that probative evidence may be excluded in order to increase the accuracy of fact-finding. A formal model is provided that rationalizes this principle. The key assumption is that the fact-finders (jurors) have a cognitive cost of processing evidence, an assumption well grounded in the psychological literature. Within this framework, the judge excludes evidence in order to incentivize the jury to focus on other, more probative evidence. Our analysis sheds light on two distinctive characteristics of this type of exclusionary rules. First, that broad exclusionary powers are delegated to the judge. Second, that exclusion by undue prejudice is peculiar to common law systems. Both features arise in our model.

Information Acquisition and the Exclusion of Evidence in Trials

Abstract: A peculiar principle of legal evidence in common law systems is that probative evidence may be excluded in order to increase the accuracy of fact-finding. A formal model is provided that rationalizes this principle. The key assumption is that the fact-finders (jurors) have a cognitive cost of processing evidence, an assumption well grounded in the psychological literature. Within this framework, the judge excludes evidence in order to incentivize the jury to focus on other, more probative evidence. Our analysis sheds light on two distinctive characteristics of this type of exclusionary rules. First, that broad exclusionary powers are delegated to the judge. Second, that exclusion by undue prejudice is peculiar to common law systems. Both features arise in our model.

1 Introduction

A peculiar principle of legal evidence in common law systems is that probative evidence may be excluded in order to increase the accuracy of fact-finding. A formal model is provided that rationalizes this principle. The key assumption is that the fact-finders (jurors) have a cognitive cost of processing evidence, an assumption well grounded in the psychological literature. Within this framework, the judge excludes evidence in order to incentivize the jury to focus on other, more probative evidence. Our analysis sheds light on two distinctive characteristics of this type of exclusionary rules. First, that broad exclusionary powers are delegated to the judge. Second, that exclusion by undue prejudice is peculiar to common law systems. Both features arise in our model.

A significant amount of recent literature in economic theory deals with endogenous information acquisition in elections, committees, and juries. The premise of this literature is that agents (voters, jurors) must be incentivized to acquire information that is socially valuable but costly to acquire. This literature then looks at the optimality of the voting procedure in question in light of the information acquisition activity. In this paper we extend this research agenda within the context of jury trials. We consider a framework that is functionally identical to the information acquisition one—we posit that jurors incur a cognitive cost when evaluating or information. We use this framework to shed light on a hitherto neglected phase of the decision-making process—not the voting rule, but rather the prior stage in which evidence is brought forward. This stage of the decision-making process is governed by the *rules of evidence*, which form a complex body of law.

An important subject matter of evidence law is the *admissibility of evidence*, i.e., what evidence can be shown to the jury and therefore influence the decision. In common law trials, the judge pre-screens the evidence the jury will get to see. Clearly, what evidence is deemed admissible has a large impact on the outcome of a trial. The jury in a trial should, as a general rule, be presented with all relevant evidence. There are,

however, exceptions to this rule. An important exception is based on the principle that excluding probative evidence may *increase the accuracy of fact-finding.* In the US system, this principle is referred to as exclusion on grounds of unfair prejudice.[1] According to this principle, the judge is given wide discretion to exclude evidence that, while probative, is seen as "unfairly" biasing the fact-finder (the jury). The principle also underlies several other more specific exclusionary rules and powers.[2] This principle is remarkable for the latitude it affords the judge to influence the outcome of the trial.

The principle that excluding probative evidence may increase the accuracy of fact-finding is peculiar of common law systems,[3] and it has received a great deal of scrutiny over the centuries. Among the early economists who have addressed this exclusionary principle is Jeremy Bentham, who thought that excluding evidence impaired jury deliberation, and devoted a large part of his 1827 treatise to exposing what he perceived as the drawbacks of exclusion of evidence (see Bentham 1827). More recently, Gordon Tullock also took a dim view of this exclusionary principle.[4] Yet, given its long history,[5] exclusion by reason of unfair prejudice should be presumed to play some important functional role in common law systems.

The conventional justification is a paternalistic one: given certain kinds of evi-

---

[1] The fundamental principle underlying such exclusions is expressed in Rule 403 from the US Federal Rules of Evidence, which states: "Relevant evidence may be excluded if its probative value is substantially outweighed by the danger of unfair prejudice [...] "

[2] Such as the exclusion of character and prior acts, the power to bar expert witnesses from testifying who are "hired guns," as well as, perhaps, the rule against hearsay evidence.

[3] Damaska (1997, p. 15) remarks "Rules typical of common law can be found only among those that reject probative information, on the belief that its elimination will enhance the accuracy of fact-finding." He calls such exclusionary rules "intrinsic."

[4] "One would rather suspect that as the result of many of the laws of evidence (not all of them), the [fact finder] is automatically somewhat erroneous as it simply ignores certain parts of the valid evidence." Cited from Tullock (1996), p. 7.

[5] Since the dawn of the common law system, the judge has had the power to exclude evidence. Originally, that power was unremarkable because the judge was so dominant in the trial. From an earlier system in which the trial judge himself collected evidence and examined the witnesses, while attorneys played a limited role, the modern system evolved in which the judge plays a much less active role. This evolution took place at the end of the 18th century, and it is then that the modern evidentiary rules developed. (See Langbein (1996), p. 1201.) The Federal Rules of Evidence, which became law in 1975, mark the greatest retrenchment yet of judicial power over the trial. Still, the judge retains the power to exclude evidence by reason of unfair prejudice.

dence, juries make systematic mistakes in updating and so they need to be protected from themselves. This paternalistic paradigm, while intuitively appealing, may not necessarily be a productive way of conceptualizing (or justifying) exclusion. First, its scope is difficult to circumscribe: how would we know when productive updating stops and "undue bias" begins? Second, and related, once we subscribe to the notion of boundedly-rational or even irrational jurors, it becomes conceptually treacherous to define notions of "more probative evidence," and "more accurate decision;" what looks more informative to the paternalistic observer might not be so to the irrational fact finder, and vice versa. Third, it is not clear why the exclusionary rules in question should not also have arisen in continental legal systems.

In this paper we articulate a model of "undue prejudice" which is not based on mistaken updating on the part of the jury. The logic is very simple. In our model, jurors can be fully rational, but they have a cognitive cost of processing information. The cost captures the idea that some information is hard to understand. Jurors are assumed to (consciously or subconsciously) optimize their mental effort in processing information – they behave as "cognitive misers."[6] Thus, jurors may focus on information that is easy to understand, though not necessarily very probative, instead of evaluating information that is very probative but hard to understand. In this framework, excluding easy-to-understand information can provide the proper incentives for the jury to focus on more probative information, thus improving the quality of the decision.

We view this model as offering a fairly conventional formalization of "undue prejudice." In the model, jurors have a tendency to ignore evidence and *pre-judge* based on their prior beliefs – hence they are "pre-judiced."[7] Excluding evidence may help induce jurors to focus on better evidence and thus to rely less on their prior beliefs – it makes them less "pre-judiced." Within this framework, the notion of "undue" prejudice can

---

[6] In Section 6 we provide evidence from the psychology literature suggesting that people are conscious of the mental effort needed to analyze evidence and that they act so as to reduce it.

[7] The *American Heritage Dictionary of the English Language* describes prejudice as : "[A] judgment or opinion formed beforehand or without knowledge or examination of the facts."

be coherently articulated and distinguished from appropriate (in a statistical sense) reliance on the juror's prior beliefs. The juror shows "undue" prejudice if he does not process a piece of available information due to his cognitive costs. The resulting over-reliance on the prior is properly viewed as "undue" because a benevolent social planner, conscious of the large positive externalities of a correct decision, would greatly discount (in the limit, ignore) the cognitive costs incurred by the jury and command the jury to evaluate that piece of evidence. In this account, the judge who excludes evidence is not behaving paternalistically. Rather, he is employing the (limited) means at his disposal to induce the jury to exert more effort. In this respect, the problem is analogous to a principal-agent relationship in an economic setting.

We study the comparative statics properties of this model, and demonstrate some fairly counterintuitive properties. For example, making evidence more probative may lead it to be optimally excluded. Similarly, a judge may optimally choose to exclude more evidence when faced with a more competent jury. Finally, we show that the decision to exclude evidence is a contextual one: optimal exclusion of one piece of evidence requires taking into account its relationship with all other pieces of evidence. We interpret these complex and counterintuitive properties as evidence that general rules mandating exclusion are unlikely to be optimal. In our model, rather, optimal exclusion can be implemented straightforwardly by giving the judge broad exclusionary powers. In this sense, our analysis finds virtues in the broad latitude currently afforded the judge. Finally, the analysis suggests a reason why exclusion by reason of unfair prejudice is characteristic of common-law systems. This is because common law systems are adversarial; we shall develop this argument later in the paper, after we have introduced some key concepts.[8]

The aim of this paper is certainly not to contend that juries are perfectly rational in their updating.[9] Indeed, our theory can properly be viewed as a model of bounded

---

[8] See Glaeser and Shleifer (2002) for a historical perspective on the divergence between common law and continental trials.

[9] Indeed, in our analysis jurors may, or may not, update in a Bayesian fashion. There is much

rationality. Rather, by proposing a model based on optimizing agents, we challenge the notion that irrationality is *required* to justify exclusionary rules. A conceptualization based on the juror as a "cognitive miser" has several advantages. From an empirical viewpoint, there is strong support in the psychology literature for the hypothesis that people incur mental effort in evaluating evidence.[10] From a model-building viewpoint, it is less arbitrary to introduce costs of evaluating evidence than to build a brand-new model of updating. And, because of the minimal departure from the standard decision-theoretic models, we can still make conceptual sense of notions such as "better information," and "more accurate decision," and our results are directly comparable with those from standard decision theory.

1.1 Related Literature

This paper is related to the economic literature on information acquisition. In particular, we share in common a focus on identifying the proper incentives for an agent, or a group of agents, to acquire costly information. The existing literature analyzes optimal information acquisition when information is aggregated through voting (e.g., Gersbach 1995, Mukhopadhaya 2003, Persico 2004, Feddersen and Sandroni 2006, Martinelli 2006, 2007, Gerardi and Yariv 2008a, and Gershkov and Szentes 2008); information acquisition in bureaucratic settings (Stephenson 2007); and the optimal transmission of information from "experts" to less informed principals (Gerardi and Yariv 2008b). Borgers et al. (2007) study the valuation of multiple pieces of information in a setting with costly information acquisition. Overall, the techniques used in this literature are quite close to those utilized here, even though our focus is not exclusively on Bayesian updating. None of these papers, however, address the specific institution or mechanism discussed in this paper.

Our paper is also related to the large legal literature inquiring about the rationale

---

evidence that they are not always rational.

[10] See Section 6.

for "intrinsic" exclusionary rules (using the terminology of Damaska 1997). Sanchirico (2001), in dealing specifically with exclusion of character evidence, provides a useful classification of such rationales. Rationales supporting the exclusion of evidence can be classified into two categories, based on the interest that exclusion is assumed to promote.

The first category of rationales are based on an incentive argument: by excluding evidence, the legislator might seek to provide incentives for potential wrongdoers. Thus, an extreme rendition of the argument goes, incentives ought to be conditioned preferably on signals that the wrongdoer can affect by his actions; all other evidence (character evidence, in particular) should not be used to provide incentives. This clever argument, proposed by Sanchirico (2001),[11] seems best suited to explain rules that *mandate* exclusion. That is because the argument relies on the predictability of exclusion on the part of the potential wrongdoer. A salient feature of Rule 403 in the US Federal Rules of Evidence, in contrast, is the *latitude* given the judge to exclude evidence on a case by case basis. That latitude seems to run counter the incentive-giving argument, because it makes it difficult for the potential wrongdoer to foresee what evidence might be excluded. Our theory, as we will show, is consistent with this latitude. Therefore, we view Sanchirico (2001) as providing a rationale for *mandatory* exclusionary rules, while our theory can explain *discretionary* exclusion. Our contribution is thus complementary, not substitute, to Sanchirico (2001).

The second category of argument is based on the view that the legislator seeks to improve the quality (accuracy) of the outcome of the trial. This is the more conventional view, and it is the one that is taken in this paper. The challenge, of course, is to explain how excluding evidence can improve the quality of the decision. Some authors simply did not believe it could. Other authors find a role for exclusion of evidence. Some authors believe that evidence of past crimes, for example, might tempt the jury into punishing the past crime as opposed to the (alleged) present one. Other authors appeal

---

[11]Schrag and Scotchmer (1994) provide a related argument.

vaguely to a tendency to be overly, or *unduly,* affected by certain kind of evidence. These arguments are similar in that they focus on some "undue bias," which may arise either because the juror's goals might be swayed by the presentation of certain kind of evidence, or because jurors might update incorrectly. By comparison, we view our approach as a small and well-defined departure from the fully rational model with zero cost of processing information. One advantage of being able to stick close to the rational model is that the structure provided by the rationality hypothesis affords some comparative statics implications, which would be difficult to obtain (or be rather arbitrary) if one departs from the rational model.

2   The Benefits of Exclusion: An Example

In this section, we introduce a simple example to illustrate how a judge's ability to exclude evidence can be welfare-improving. Suppose that a juror is asked to decide whether the speed of a car that was involved in an accident exceeded 50 miles per hour. Let x denote the speed of the car, which in the juror's mind is equally likely to be any real number between 0 and 100 miles per hour. Formally, we think of x as a random variable that is uniformly distributed in the interval $[0, 100]$. If the juror rules correctly, he receives a payoff of 0, and if he rules incorrectly, he receives a payoff of -100. There are two pieces of evidence available, but each is costly to process. The first piece of evidence, which we denote E_1, has a cost of 5 and informs the juror whether or not x is in the interval $[0, 20]$. The second piece of evidence, E_2, has a cost of 35 and informs the juror whether or not x is in the interval $[10, 50]$.

E_1 is less valuable information than E_2 on average. Observing the realization of E_1 and learning whether the car was driving below 20 miles per hour provides relatively little information about whether the car was driving below 50 miles per hour. Indeed, after learning that the driver's speed did *not* lie in the interval $[0, 20]$, the probability that the driver's speed was less than 50 miles per hour, $\frac{3}{8}$, is still relatively large. On the

other hand, observing the realization of E_2 and learning whether the car was driving between 10 and 50 miles per hour provides a lot of information about whether the car was driving below 50 miles per hour; after learning that the driver's speed did *not* lie in the interval $[10, 50]$, the probability that the driver's speed was less than 50 miles per hour is reduced to $\frac{1}{6}$. Finally, note that observing the realization of both pieces of evidence will allow the juror to know with certainty if the speed exceeded 50.

The table below illustrates the payoffs to both the juror and society when the juror chooses to process just the first piece of evidence, just the second, both, and neither.[12] We assume that the cognitive costs of processing information incurred by the jurors are negligible to society.[13]

Table 1: Payoffs to Juror and Society

Juror			Society	
Process	Expected Payoff		Process	Expected Payoff
E_1	-35		E_1	-30
E_2	-45		E_2	-10
E_1, E_2	-40		E_1, E_2	0
Neither	-50		Neither	-50

Clearly, then, the optimal outcome for society is for the juror to process both pieces of evidence. However, the juror would choose to only process E_1, as the gains from a more accurate ruling associated with E_2 are outweighed by the costs of processing.

The same point is illustrated graphically in Figure 1. Evaluating E_1, E_2 is socially ideal, but consideration of the private costs will lead the juror to evaluate only E_1 instead. To prevent this outcome, which is socially undesirable, the judge can exclude E_1 (thereby also excluding the package E_1, E_2) and the juror will choose to process E_2. Though this outcome is second best, the payoff to society clearly dominates the payoff

---

[12] An explanation of how these values were derived is provided in the appendix.

[13] One justification for this assumption is that the direct benefits of a correct ruling on the lives of those involved in court cases, as well as the indirect benefits of maintaining a fair, trustworthy legal system, far outweigh the cognitive costs of a few selected jurors.

in the absence of exclusion.

[Insert Figure 1 Here]

3 Model

We assume, for simplicity, that the jury is composed of only one juror.[14] Let θ be a random variable denoting the true state of the world. The realization of this random variable is what needs to be determined, but it is unknown to both the judge and the juror. There is a set of random variables $\mathcal{S} \equiv \{E_1, E_2, ..., E_n\}$ that are correlated with θ. We will refer to these variables as *pieces of evidence*, and to any subset of it as an *information set*. The information system $\mathcal{S}$ represents all the evidence that could conceivably be presented to the juror.

The juror, but not the judge, has the ability to *evaluate* the evidence, which means that the juror can extract the information contained in the evidence. We think of the evaluation process as analogous to opening a box and observing the realization e_i of the piece of evidence E_i. Opening the box entails a cost, associated with the cognitive process of evaluating a piece of evidence and using the information to update beliefs. For example, evaluating the accounting evidence presented in a complex financial fraud case, and drawing implications concerning the guilt of the defendant, can be mentally quite taxing for the jury. After the evidence is evaluated, it may turn out that the evidence exonerates the defendant or that it incriminates him, or that the evidence is not relevant. The "box opening" metaphor captures this costly evaluation process. We also assume that, before going through the evaluation process, the jury and the judge have a sense of the probative value contained *on average* in the piece of evidence; formally, the probability distribution over realizations of E_i, conditional on θ, is known to all. That is, the judge and the jury can foresee the expected benefits of delving into the accounting evidence. This may, for example, lead the jury to rationally "tune off"

---

[14]This assumption is relaxed in Section 6.1.

11

the accounting evidence, and to rely instead on other evidence which may be cognitively easier to process (evidence of the defendant's wealth, for example).

Not all pieces of evidence need to be presented to the juror, nor is the juror obliged to evaluate every piece of available evidence. At the judge's discretion, the juror may be presented with any subset $\mathbf{S} \subseteq \mathcal{S}$ of all the possible evidence. The juror, in turn, may choose to restrict attention to any subset $\mathbf{s} \subseteq \mathbf{S}$ of the evidence that is presented to her; for example, the juror may choose not to evaluate any piece of evidence, in which case $\mathbf{s} = \varnothing$, or the juror may choose to evaluate only the first two pieces of evidence, in which case $\mathbf{s} = \{E_1, E_2\}$.

If the juror evaluates a subset $\mathbf{s}$ of the evidence presented to her, she receives an expected payoff

$$V\left(\mathbf{s}\right) - C\left(\mathbf{s}\right).$$

The function $C\left(\mathbf{s}\right)$ represents the cost the juror incurs from evaluating the information set $\mathbf{s}$. The function $V\left(\cdot\right)$ represents the expected benefit to the juror from adjudicating the case. We shall assume that V is monotonic in the sense that

$$V\left(\mathbf{s}\right) < V\left(\mathbf{s}'\right) \text{ if } \mathbf{s} \subset \mathbf{s}'.$$

This implies that every piece of evidence is valuable, in that it helps increase the accuracy of the decision.

3.1 Social Welfare and the Problem of the Judge

We stipulate that the expected value to society from adjudicating the case based on consideration of information set $\mathbf{s}$ is given by $V\left(\mathbf{s}\right)$. This amounts to assuming that C, the juror's disutility from processing information, is negligible to society relative to the benefit of reaching the correct decision. Because V is monotonic, the maximum value of V is achieved when all information is utilized by the juror. The juror, in contrast, does not necessarily want V to be maximal because her objective function also involves

C. Thus, the juror does *not* generally have the socially proper incentives to process all available information, and an agency problem arises.

The judge, whose utility coincides with society's, simply wants V to be maximal. The divergence of interests between the juror and the judge (or society) leaves room for socially beneficial intervention on the part of the judge. We assume that the act of evaluating evidence is not contractible, so the juror cannot be compensated based on the evidence she might choose to evaluate. Such, of course, is the case in real-world courts. The only instrument that the judge may use to intervene in the adjudication process is the exclusion of evidence. By restricting the set of evidence presented to the juror, the judge may induce the juror to evaluate more probative evidence.

In our model, the judge chooses the subset of evidence $\mathbf{S}$ to present to the juror so as to maximize the probative value of the evidence evaluated by the juror. The judge cannot, or at lepast does not, perform the task of evaluating evidence before deciding on exclusion.[15] Formally, the judge's problem is

$$\max_{\mathbf{S}} V\left(\mathbf{s}^*\right)$$
$$\text{s.t. } \mathbf{s}^* \in \arg\max_{\mathbf{s} \subseteq \mathbf{S}} V\left(\mathbf{s}\right) - C\left(\mathbf{s}\right).$$

3.2 Special Case: A Bayesian Juror.

The model developed above is not necessarily tied to the assumptions of Bayesian updating. Instead, it operates at a more abstract level by taking as primitive the function V, which may or may not derive from Bayesian updating. In the special case of Bayesian updating, the function $V\left(\mathbf{s}\right)$ would represent the expected gains from a correct ruling, conditional on the information contained in $\mathbf{s}$, less the expected losses from type one and type two errors. In this framework, evaluating evidence means observing the realization e_i of E_i[16] and updating the probability distribution over θ

---

[15] In our model, this role is reserved for the jury. This assumption embodies the common-law principle that fact-finding is for the jury, and the judge is supposed to act as a referee.

[16] We assume that n, the joint distribution of θ and $E_1, ..., E_n$, as well as the cognitive costs, are common knowledge to the judge and juror. This assumption allows us to avoid the possibility that a

according to Bayes' rule.[17]. After choosing a subset of evidence $\mathbf{s}$ to process, and observing the realized value of each piece of evidence in this subset, a juror must make a decision, d (conviction or acquittal, for example). The decision gives rise to a payoff which depends on the true state of the world via a loss function $L(d, \theta)$. For example, the juror may feel a loss of zero if the decision is correct (acquit the innocent, convict the guilty) and experience a negative payoff if the decision is incorrect. The expected loss of a juror who makes a decision upon observing $\mathbf{e} \equiv \{e_i : E_i \in \mathbf{s}\}$, a specific realization of $\mathbf{s}$, is

$$v(\mathbf{e}) = \max_d \mathbb{E}\left[L(d, \theta) | \mathbf{e}\right],$$

where the letter $\mathbb{E}$ represents the (conditional) expectation operator applied to θ. The function V in our model is the expected loss *ex ante*, before observing the realization $\mathbf{e}$. Thus, in this case the function V is given by

$$V(\mathbf{s}) = \mathbb{E}\left[v(\mathbf{e})\right].$$

Of course, $V(\mathbf{s}) < V(\mathbf{s}')$ if $\mathbf{s} \subset \mathbf{s}'$, because a Bayesian decision maker can make a better decision when he has more information.

4   The Absence of General Principles Guiding Exclusion

In this section we present several results pointing to the difficulty of eliciting general principles that can inform the exclusion of specific pieces of evidence as a general rule. One source of this difficulty is that optimal exclusion is necessarily *conditional on the*

---

juror would update his beliefs based on what evidence is *not* presented. In our model, excluding E_i is equivalent to admitting it but with a very high cost c_i.

[17]It is worth emphasizing that we do not restrict pieces of evidence to be conditionally independent. In the literature that attempts to model juror judgment and decision making, the label "Bayesian updating" is sometimes equated with Bayesian updating for the special case where all evidence is conditionally independent (see e.g. Hastie 1993). In this case, all information is captured in the conditional likelihood functions. In other words, this restrictive interpretation of Bayesian jurors rules out any more elaborate interdependence of evidence, e.g. complementarity and substitutability of evidence. We argue, to the contrary, that interdependencies of evidence are important, and illustrate it for the case of the costs and benefits of exclusion.

totality of the evidence at one's disposal. To see this, note that if $\mathcal{S} = \{E_1\}$ is a singleton, then E_1 should always be admitted, regardless of its informational content or cognitive costs. But if exclusion leads some other piece of evidence to be evaluated, as in the initial example, then it may be optimal to exclude E_1. Hence the key point: excluding a piece of evidence (E_1 in our case) may be beneficial or detrimental, depending on the characteristics of other available pieces of evidence. In theory, therefore, a general rule which attempted to *mandate optimal* exclusion will need to condition the exclusionary rules on the fine details of the *other* evidence available in the case.

The fact that optimal exclusion is conditional is not the only reason why general principles concerning exclusion are difficult to come by. In the remainder of this section we show that optimal exclusion can have some counterintuitive properties. We interpret these findings as suggestive that it is difficult, within our model, to give general prescriptions about what evidence ought to be excluded. We take these cautionary results as supportive of the practice of delegating to the judge a broad authority to exclude evidence.

4.1 The Informational Content of Evidence, Outcomes, and Exclusion

First, we will use an example to illustrate that improving the accuracy of evidence may lead to a worse decision. We heed strictly to the fully rational, Bayesian updating framework; the V functions in the examples are derived from Bayesian decision making, and the cost function is actually additive (a special case of submodularity). Given the purpose of the (counter)examples, the fact that they obtain in a very conventional environment should help convince the reader that they are a robust feature in this framework.

Consider the example described in Section 2, in which the juror must rule on whether the speed of a car, which we denote x, was greater or less than 50 miles per hour. Let us maintain all previous assumptions on the distribution of x, the juror's payoffs, and

the properties of E_2. First, consider evidence E_1^A, which has cost $c_1^A = 5$ and reveals whether x lies in the interval $[0, 10]$ or whether it lies in the interval $(10, 100]$. The payoffs to the juror and society are characterized in the Table 2 below.

Table 2: Payoffs to Juror and Society

Juror			Society	
Process	Expected Payoff		Process	Expected Payoff
E_1	-45		E_1	-45
E_2	-45		E_2	-10
E_1, E_2	-40		E_1, E_2	0
Neither	-50		Neither	-50

Under this information system, the juror chooses to process both E_1^A and E_2 when both pieces of evidence are available. In words, E_1^A is sufficiently uninformative that the juror seeks out additional information in the form of E_2. Note that the outcome is the first best.

Let us now replace evidence E_1^A with evidence E_1^B, which also has cost $c_1^B = 5$, but reveals whether x lies in the interval $[0, 10]$, $(10, 20]$, or $(20, 100]$. Notice two characteristics of E_1^B. First, for this decision problem, it is equivalent to evidence E_1 in the original example in Section 2. Therefore, the juror's optimal decision is the same as in Section 2: E_1^B is sufficiently informative that the juror does *not* find it optimal to process E_2, given its cost. As a result, the juror only processes E_1^B and the payoff to society is strictly lower than the payoff under the information system E_1^A, E_2. Secondly, note that E_1^B is *more informative* than E_1^A in the sense of Blackwell (1951): it is more valuable in any decision problem. Therefore, we conclude that more informative evidence may lead to worse outcomes in the absence of exclusion.

Result 1. *Absent exclusion, more informative evidence (in the sense of Blackwell) can lead to worse outcomes.*

A counterintuitive corollary (or re-interpretation) of this result is that finding jurors

who are better able to evaluate evidence is not necessarily desirable from a social viewpoint. Such jurors may rely on a smaller subset of evidence (E_1^B in the example above), while less able jurors, aware of their limitations, will continue to seek out additional information (both E_1^A and E_2) before reaching a decision.

Corollary 1. *A jury that has the ability to interpret evidence more accurately (in the sense of Blackwell) may make less accurate decisions.*

Note that these counterintuitive results cannot be eliminated by optimal exclusion. Returning to the previous example, a judge may want to allow a certain piece of evidence (i.e. E_1^A) and yet, caeteris paribus, the judge may want to exclude a *more informative* version of the same piece of evidence (i.e. E_1^B). Indeed, in the example the first best outcome was achieved under information system (E_1^A, E_2), since the juror's optimal choice implied the maximal payoff to society, zero. However, once E_1^A is replaced with more informative evidence, the judge optimally excludes E_1^B and the payoff is reduced to -10. This shows that *even with optimal exclusion*, better evidence may lead to worse outcomes.

Proposition 1. *(Quality of evidence and exclusion)* *Improving the probative value of a piece of evidence (in the sense of Blackwell) may lead that piece of evidence to be optimally excluded. Even with exclusion, more informative evidence can lead to worse outcomes.*

The corollary below again translates our finding to speak about the jury's level of ability. A judge facing a jury who is capable of a more accurate reading of the evidence (E_1^B instead of E_1^A) may be lead to exclude the first piece of evidence (E_1^B).

Corollary 2. *A "better" jury (one that has the ability to interpret evidence more accurately in the sense of Blackwell) may lead the judge to optimally exclude more evidence.*

This result, while counterintuitive, is also insightful in that it makes clear that, within our model, the judge excludes evidence not to *protect the jury* from evidence that it is unfit to process, but rather to *provide incentives* for the jury to seek out more informative evidence.[18]

In conclusion, since excluding evidence is as much about the evidence that is admitted as about the evidence that is not, even the most basic properties we would expect can fail to be true. As such, general principles guiding optimal exclusion are difficult to identify. Fortunately, a judge is available in our setup who can be trusted with the power to optimally exclude evidence on a case-by-case basis. It seems natural in this setup that there will be few rules mandating exclusion, and that broad exclusionary powers would be delegated to the judge. We therefore interpret our negative results as making the case for delegating to the judge a broad authority to exclude evidence.

5 Complementary and Substitutable Evidence

As illustrated above, the decision to exclude a piece of evidence relies heavily on how it relates to other pieces of evidence. In this section, we formalize the notions of complementary and substitutable pieces of evidence, and pursue exclusionary principles that might be based on these notions. Intuitively, two pieces of evidence are complementary if possessing one makes it more desirable for the jury to acquire the other. We will show that if *all* pieces of evidence in an information system are complementary, then excluding any subset of them cannot improve the quality of the decision. We then go on to suggest that in an adversarial system it is unlikely that the entire information system in a trial is complementary. Rather, it is more likely that the information sets put forth by the two parties – plaintiff and defendant – will contain substitutable pieces of evidence, even though the pieces of evidence presented by a single party may well be complementary among themselves. We then introduce a notion of complementarity

---

[18] We expand upon this example in the appendix to provide further insights into the intuition behind our results.

"within information sets," one which does not extend to the whole information system. Elements of an information sets that are complementary with each other we call *stories*. We then show another negative result, namely, that the judge does not necessarily want to admit evidence that complements a story.

5.1 Definitions

The mathematical notion of complementarity is related to the legal notion of "conditional relevance." This notion enters the decision of what pieces of evidence are to be considered relevant, and so may be admitted to trial. When the probative value of a piece of evidence is positively dependent on the presence of another piece of evidence, the judge needs to weigh the *joint* probative value of the two pieces of evidence.[19]

A formal definition of complementary information is based on the notion of supermodularity.

Definition 1. *A function f is said to be **supermodular** if, for any two information sets $\mathbf{s}_1, \mathbf{s}_2$,*

$$f\left(\mathbf{s}_1 \cup \mathbf{s}_2\right) + f\left(\mathbf{s}_1 \cap \mathbf{s}_2\right) \geq f\left(\mathbf{s}_1\right) + f\left(\mathbf{s}_2\right).$$

*We say that a function f is **submodular** if $-f$ is supermodular.*

Definition 2. *An information system is **complementary** if the associated value function V is supermodular. In that case the separate pieces of evidence $E_1, ... E_n$ of the information system are said to be **complementary** to each other. If V is submodular then the pieces of evidence are called **substitutes**.*

To illustrate the meaning of complementarity in our context, let $\mathbf{s}_1 = E_1, ... E_{n-1}$ and $\mathbf{s}_2 = E_n$ in the equation above. Rearranging terms yields

$$V\left(E_1, ... E_n\right) - V\left(E_1, ... E_{n-1}\right) \geq V\left(E_n\right) - V\left(\emptyset\right).$$

---

[19]Rule 104(b) provides that "(w)hen the relevancy of evidence depends upon the fulfillment of a condition of fact, the court shall admit it upon, or subject to, the introduction of evidence sufficient to support a finding of the fulfillment of the condition."

In words, information piece E_n is more valuable – that is, it leads to a greater increase in the value function – when it is paired with the set $E_1, ... E_{n-1}$ than when it is considered in isolation. When the value function is supermodular, each piece of information is most valuable when considered in the context of other information.

For an example of complementary pieces of evidence, suppose the question to be adjudicated is whether a US citizen defendant is or is not a member of the Yakuza, the Japanese mafia. It is known that many Yakuza members are missing a pinky finger, owing to their custom of severing it as a self-imposed penalty for unsatisfactory conduct with regards to the criminal organization. Now consider the following two pieces of evidence: ethnicity, and whether a pinky finger is missing. Each piece of evidence on its own has almost no probative value of membership in the criminal organization— the great majority of Japanese-Americans do not belong to the Yakuza, and the great majority of US citizens with missing fingers are presumably unlucky carpenters. Yet the two pieces of evidence together represent somewhat probative evidence. Thus, the two pieces of evidence are complements.

For an example of substitute pieces of evidence, suppose the question to be adjudicated is whether the defendant committed a particular crime that occurred in New York City. There are two pieces of evidence. One is computer records from a toll booth indicating that the defendant's car entered New York City. The other is a parking violation incurred on the streets of New York City. Either piece of information may be quite informative about the whereabouts of the defendant on the day in question. However, knowing one decreases the jury's value of knowing the other.

When the function V is derived from a Bayesian decision problem, whether or not an information system $E_1, ... E_n$ is complementary depends not only on their joint distribution conditional on θ, but also on the prior over θ and on the loss function, all of which enter the expression for V. For instance, two pieces of evidence may be

complementary for a certain prior over θ and substitute for another prior.[20]

Definition 3. *The cost function C is said to have nondecreasing returns to scale if C is submodular.*

The assumption of nondecreasing returns implies that the "marginal" cost of evaluating a piece of evidence decreases when other pieces of evidence are also considered. This is a property of returns to scale in the evaluation of costly evidence. A special case of submodularity is additivity, the case in which for every disjoint $\mathbf{s}_1$ and $\mathbf{s}_2$ we have $C(\mathbf{s}_1 \cup \mathbf{s}_2) = C(\mathbf{s}_1) + C(\mathbf{s}_2)$. In the additive case the marginal cost of evaluating evidence is independent of the amount of other evidence being evaluated.

5.2 Exclusion and Complementary Evidence

In the introduction, we showed how excluding a piece of evidence can be welfare-improving. By removing cheaper, less informative evidence, the judge manipulated the juror's choice set so that more informative evidence was processed, and a better decision (on average) was handed down. We now describe circumstances in which excluding evidence can *not* be beneficial.

Assumption 1. *If the jury is indifferent among processing several information sets, the jury will choose the one that is most informative (i.e., the one with the highest social welfare).*

This assumption is weak in that it *only* restricts the choices made when the jury is indifferent among several subsets of evidence. We should expect this occurrence to be very unlikely, in the sense that it does not occur in a generic set of primitives.[21]

Proposition 2. *If the information system is complementary, the cost function exhibits nondecreasing returns to scale, and Assumption 1 holds, then excluding information cannot improve the quality of the decision.*

---

[20] Persico (2005) provides an example of this phenomenon in the context of a jury model.

[21] An alternative to Assumption 1 that yields equivalent results is to assume independence of irrelevant alternatives.

Proof. We will prove the result by contradiction. Suppose that all pieces of evidence in information system $\mathcal{S}$ are complementary, and let $\mathbf{s}^* \subseteq \mathcal{S}$ denote the subset of information that the juror chooses to process when all pieces of evidence in $\mathcal{S}$ are allowed. Let $\hat{\mathbf{s}}$ denote the juror's choice when only pieces of evidence in the set $\mathbf{S}^A \subset \mathcal{S}$ are admitted. Due to our assumptions of complementarity and returns to scale, the function $f(\cdot) = V(\cdot) - C(\cdot)$ is supermodular. Then the following holds:

$$f(\hat{\mathbf{s}} \cup \mathbf{s}^*) - f(\mathbf{s}^*) \geq f(\hat{\mathbf{s}}) - f(\hat{\mathbf{s}} \cap \mathbf{s}^*) \geq 0, \tag{1}$$

where the second inequality follows from the fact that $\hat{s}$ is the jury's choice within the set $\mathbf{S}^A$, which contains $\hat{s} \cap s^*$. It follows from equation (1) that it must be $f(\hat{s} \cup s^*) \geq f(\mathbf{s}^*)$. Strict inequality cannot hold, by definition of $\mathbf{s}^*$, so that it must be

$$f(\hat{\mathbf{s}} \cup \mathbf{s}^*) = f(\mathbf{s}^*). \tag{2}$$

This shows that the jury must be indifferent between processing $\hat{\mathbf{s}} \cup \mathbf{s}^*$ and $\mathbf{s}^*$. Suppose, towards a contradiction, that $\hat{\mathbf{s}}$ were strictly more informative than $\mathbf{s}^*$. Then $\hat{\mathbf{s}} \cup \mathbf{s}^*$ is also strictly more informative than $\mathbf{s}^*$. By Assumption 1, then, the jury could not have chosen to process $\mathbf{s}^*$ when all pieces of evidence $\mathcal{S}$ are allowed. This establishes the required contradiction. ∎

The relevance of this result, of course, depends on the likelihood that the entire information system is complementary. In adversarial systems, these circumstances would seem unlikely. Though each side (plaintiff and defendant) may present a subset of evidence composed of complementary pieces of evidence – what one might call an argument or *story* – in general the pieces of evidence within one party's argument may very well be substitutes in relation to the opposition's argument. We formalize these ideas, and show that general results regarding the optimal use of exclusion are hard to come by in this context. Again, we interpret this dearth of general prescriptions as an affirmative argument for delegating the unfettered exercise of exclusionary powers to the judge.

5.3 Complementarity of Evidence in an Adversarial System

When an information system is complementary, there is no role for exclusion. However, in an adversarial system, it may be unlikely that all available pieces of evidence are complementary. In such a system, the plaintiff and defendant each gather and present separate evidence to tell their own "story." Presumably, each party hopes that the jury will listen to their story and disregard their opponent's; in our language, it is likely that pieces of one story are substitutes for pieces of the other.

Again, an example may be helpful. Consider the case of a crime committed in a particular neighborhood of New York City. The prosecution might present a parking violation incurred by the defendant in that neighborhood, along with other potentially damning evidence, to develop a story to suggest the defendant's guilt. The defense may present evidence that the defendant visited a family member living in that neighborhood, in conjunction with other potentially exculpatory evidence, to develop a story to suggest the defendant's innocence. These two pieces of evidence would be substitutes – they both establish the defendant's presence in the neighborhood in question – but they are part of opposing stories.

The question, then, is whether Proposition 2 can be extended to a setup in which the entire information system is not necessarily complementary. The answer is negative: we show that when the information system is made up of two competing stories, it may be optimal to exclude parts of a story. So, the fact that pieces of evidence in a story are complements does not guarantee that it is necessarily optimal for *all* pieces to be admitted. Put differently, the fact that a piece of information is complementary to a story which gets told at the trial does not necessarily guarantee that it is optimal to admit that piece.

To make our point formally, we need to define what we mean by "story." Intuitively, a story is a collection of pieces of evidence which are all complementary with each other. The formal definition follows.

Definition 4. *A subset* $\mathbf{S}$ *of an information system* $\mathcal{S}$ *is said to be a* story *if, for all* $\mathbf{a}_1, \mathbf{a}_2 \subset \mathbf{S}$ *and* $\mathbf{b} \subset \mathcal{S} \backslash \mathbf{S}$,

$$V\left(\left(\mathbf{a}_1 \cup \mathbf{a}_2\right) \cup \mathbf{b}\right) + V\left(\left(\mathbf{a}_1 \cap \mathbf{a}_2\right) \cup \mathbf{b}\right) \geq V\left(\mathbf{a}_1 \cup \mathbf{b}\right) + V\left(\mathbf{a}_2 \cup \mathbf{b}\right).$$

According to this definition, a story $\mathbf{S}$ is composed of pieces of evidence which are all complements with each other regardless of what evidence $\mathbf{b}$ may exist outside of $\mathbf{S}$.

Proposition 3. *It may be optimal to exclude part of a story.*

Proof. We will show that the property holds in an example with a Bayesian decision maker. Consider a Bayesian decision problem in which the unknown θ can take one of two values, guilty or innocent, with equal probability. The action is binary: convict or acquit. The loss function is equal to -1 if convicting the innocent or acquitting the guilty and 1 otherwise, so that $V(\emptyset) = 0$. Let the plaintiff's story be a singleton $\mathbf{S}_P = \{E_1\}$, and let the defendant's story be composed of two complementary pieces of evidence, $\mathbf{S}_D = \{E_2, E_3\}$. Suppose for simplicity that the cost function is additive, with $C(E_2) = C(E_3) = 0$. $C(E_1)$ will be determined below.

Suppose E_1 and (E_2, E_3) are substitutes, and that neither story is perfectly informative, so that

$$1 > V(E_1) \geq V(E_1, E_2, E_3) - V(E_2, E_3) \ \equiv \ \epsilon \geq 0. \tag{3}$$

Moreover, suppose that processing of the bundle (E_1, E_3) is socially preferred to the bundle (E_2, E_3), though neither bundle is perfectly informative, so that

$$1 > V(E_1, E_3) - V(E_2, E_3) \ \equiv \ \delta > 0. \tag{4}$$

Finally, suppose that E_3 has very little probative value, while E_1 is very informative, so that

$$V(E_1, E_3) - V(E_3) \ = \ 1 - \eta \tag{5}$$

for some $0 \leq \eta < 1 - \max\{\epsilon, \delta\}$. Then for any $\max\{\epsilon, \delta\} < C(E_1) < 1 - \eta$, the following inequalities are true:

$$V(E_2, E_3) \quad > \quad V(E_1, E_2, E_3) - C(E_1) \tag{6}$$

$$V(E_2, E_3) \quad > \quad V(E_1, E_3) - C(E_1) \tag{7}$$

$$V(E_1, E_3) - C(E_1) \quad > \quad V(E_3). \tag{8}$$

These three inequalities imply, respectively, that (i) $C(E_1)$ is sufficiently large that the juror would not choose to process all three pieces of evidence, (ii) $C(E_1)$ is sufficiently large that the juror prefers (E_2, E_3) to (E_1, E_3), and (iii) $C(E_1)$ is sufficiently small that the juror prefers to (E_1, E_3) to just E_1. Given our assumption that $V(E_1, E_3) > V(E_2, E_3)$, it follows immediately that the judge would optimally exclude E_2, even though it is complementary to E_3. ∎

6   Discussion of Modeling Assumptions

In this section, we discuss several of our modeling assumptions. The first discussion is technical; we establish conditions under which our analysis of a single-juror model would extend to a multi-juror setup. The second discussion is practical; we present a variety of evidence from research in cognitive psychology, behavioral decision-making, and even psychophysiology to support our assumptions behind the cognitive costs of processing information and the strategic behavior of jurors in selecting which evidence to process.

6.1   Multiple Jurors

In the benchmark model, we assumed that there is only a single fact-finder. This was done mainly for expositional ease. We now establish conditions under which the analysis carried out in the previous sections applies verbatim to juries of any size. To that end, suppose that there are $J > 1$ jurors that are homogeneous with respect to

their preferences, their ability to process information, and the manner in which they update beliefs. Let us assume further that, once the effort of evaluating the significance of a piece of evidence has been incurred, a juror can communicate his conclusions to the other jurors immediately and at zero cost. Finally, let us assume that the cost function is additive.

Since jurors have common values, they will want to share fully the outcome of whatever evidence they have evaluated, and therefore all jurors will have the same beliefs after information has been shared. Moreover, since jurors have identical preferences and beliefs, they will naturally agree on the optimal decision. Operatively, this means that if $\mathbf{s}$ represents the evidence collectively evaluated by all members of the jury, then all jurors share the same function $V(\mathbf{s})$.

It remains to be determined, however, *who* among the jurors is responsible for evaluating the various pieces of evidence. In this respect jurors face a free-riding problem, since each juror would rather that someone else evaluate the information and report the outcome to all. However, consider the strategy of a single juror when considering whether or not to evaluate a piece of evidence E_n, taking as given the subset $\mathbf{s}$ of evidence being processed by other jurors. The private benefit of evaluating E_n is given by

$$V(\mathbf{s} \cup E_n) - V(\mathbf{s})$$

and the private cost is given by $C(E_n)$. In this sense, the "marginal" conditions that dictate whether a juror evaluates E_n in a multi-juror model are identical to those conditions in a single-juror model. Consequently, if it is optimal to acquire the configuration $\mathbf{s}^*$ in the single-juror setup, then $\mathbf{s}^*$ is also a Nash equilibrium in the multi-juror case[22]

Note that this observation is silent on how cognitive costs will be distributed across the jurors. This distributional question is immaterial because we have assumed that the cost function is linear. It is possible, in particular, that all the evaluating is performed

---

[22] The absence of any possibility of payments (in kind or otherwise) for effort expended among the jury motivates the noncooperative equilibrium concept.

by just one juror, or that it is distributed equitably among all jury members. If we had a cost function C with non-decreasing returns to scale, as assumed in Section 5, then there would be efficiency gains from assigning all cognitive costs to a single juror. Then, again, the configuration $\mathbf{s}^*$ that is optimal in the single-juror setup is an equilibrium in the multi-juror case. We record these observations in the following proposition.

Proposition 4. *Suppose all jurors share identical functions V and C, the function C has non-decreasing returns, and jurors can share effortlessly the result of their evaluation of evidence. Then if in the single-juror setup it is optimal to acquire a configuration $\mathbf{s}^*$, then $\mathbf{s}^*$ is also a Nash equilibrium in the multi-juror case.*

If the cost function had decreasing returns to scale, then it might be optimal to distribute the effort among jury members. In that case, a large jury might perform better than a single-person jury because it would be able to allocate effort more efficiently. If there was heterogeneity across jurors, then the analysis would have to be adapted to deal with the problem of aggregating the disparate preferences of the jurors. In such a setting the voting rules (simple majority, unanimity, etc.) will presumably matter. Nevertheless, we expect the key results—that exclusion is a way of providing the proper incentives for jurors to exert mental effort, and that delegation to a judge may be preferable to mandatory exclusion rules—to carry over to such environments.

6.2 Cognitive Capabilities of Jurors

We make three crucial assumptions that underlie our analysis of juror behavior: that evidence must be processed in order to learn its informational content, that such processing requires some cognitive cost, and that jurors can strategically select which pieces of evidence they will process and which pieces they will not. We now discuss each of these assumptions in order, and provide evidence to support that our specification of juror behavior is consistent with research on learning, cognitive capabilities, and decision-making.

The first crucial assumption is that there is a distinction between a piece of evidence and the fact that is trying to be established. For example, a piece of evidence might be a footprint of a particular size and shape near the scene of the crime, and the fact that is trying to be established is that the defendant was at the scene of the crime. The footprint alone - prior to any consideration by a juror - cannot establish the defendant's presence at the scene. Instead, the juror must listen to the various arguments, compare the footprint to the defendant's footprint, and so forth in order to conclude whether the defendant was, indeed, at the scene of the crime. This distinction between evidence and fact is common in the legal literature; Loh (1985) speaks to this directly when he states that "proof involves drawing inference from the evidence... [since] no conclusion can be drawn from facts without some step of inductive inference."

The second crucial assumption is that there is a cost associated with absorbing, processing, and drawing an inference from a piece of evidence. Such costs have been studied and verified at various stages of the learning and decision-making process. At the first stage, attention is required to simply observe or listen to a new piece of evidence. As Broadbent (1958) and Kahnemann (1973) document, such attention requires effort and the use of limited cognitive resources.[23] At the second stage, both time and effort are required to transform new material into long-term memory that can later be used in combination with other knowledge to draw inferences (see Craik and Lockhart 1972 and Lindsay and Norman 1977).

At the last stage, reasoning itself requires cognitive strain. This is a widely accepted fact in the literature on cognition and decision-making, though a universally accepted metric of this strain remains elusive. There are various approaches to measuring cognitive effort (see O'Donnell and Eggemeier 1986). One approach is to simply ask the subjects in an experiment to rate their own expended effort. A basic finding is that

---

[23] Further documentation and discussion of limited cognitive resources can be found in Moray (1967), Norman and Bobrow (1975), and Navon and Gopher (1979, 1980). Experiments by Wickens and Kramer (1985) and Pashler (1988) lend additional support to our assumptions on the mind's limitations; they find considerable interference when subjects are asked to perform two tasks simultaneously.

individuals consistently rate tasks that require more elaborate and precise tactics as more costly in terms of effort (Payne et al. 1993). Another approach is to give a second task to subjects, and observe how much this task interferes with the original reasoning task. Reasoning tasks or methods that are considered harder are indeed commonly found to be subject to more interference by a second task. Third, researchers can observe the physiological effects of working through reasoning tasks of different levels of difficulty. They have found that cognitive strain can be measured physiologically, documenting increases in heart rate, blood pressure, and glucose metabolism much like physical exertion.[24] Bettman et al. (1990) review previous attempts at measuring cognitive strain, and propose a universal metric by establishing a relationship between the amount of effort expended on certain tasks and the number of "elementary information processes"[25] they require. They find that the total portfolio of elementary information processes is indeed a good predictor of the time needed to solve a problem and of the self-reported cognitive effort.

These studies provide clear evidence of the cognitive limits on reasoning and the effort costs associated with it. Indeed, the limitations of the human mind and the cost of cognitive processing are acknowledged in a wide range of psychological literatures,[26] and even serve as *maintained assumptions* in many lines of research. For example, a large body of work explores methods that ameliorate the strain of processing infor-

[24] Mulder and Mulder (1987), Aasman et al. (1987) and Backs and Seljos (1994) document the effect of reasoning on pulse, heart rate variability, and blood pressure, respectively. Jonides et al. (1997), Fibiger et al. (1986) and Lund-Anderson (1979) document the effect of cognitive strain on glucose metabolism. One key aspect of understanding cognitive strain is recognizing that reasoning requires the use of working memory, which has limited capacity and involves significant exertion. See, for example, Miller (1956), Miyaki and Shah (1999), and De Neys et al. (2005). On a more illustrative level, Leedy and Dubeck (1971) find that master chess players lose a significant amount of energy and body weight during matches.

[25] The approach of splitting cognitive processes up into elementary parts goes back at least to Newell and Simon (1972). Examples of elementary information processes are: reads, additions, differences, products, elimination, comparisons.

[26] See, for example, Operario and Fiske (1999) and Kruglanski and Orehek (2007) for the field of social cognition, and Hastie (2001) and Mellers et al. (1998) for perspectives on judgement and decision-making research.

mation in order to extract the best cognitive performance.[27] Typical questions are concerned with the dependence of cognitive processes on the manner in which information is presented, as well as the effectiveness of additional tools, such as note-taking, on cognitive performance. In sum, we conclude that there is ample support for the assumption that processing evidence is costly.

Having established that evidence must be processed, and that such processing is costly, our third and final crucial assumption is that the juror has a choice of whether or not to incur these costs when presented with a piece of evidence. Treating the juror as a strategic decision-maker in gathering costly information is the foundation of research on the selection of decision strategies. In this literature, the trade-off between cognitive effort and accuracy is studied carefully.[28] For example by changing the payoffs associated with a correct answer, Payne et al. (1993) find that changes in the benefits of accuracy lead to changes in the amount of cognitive effort expended. They also attempt to quantify the effort expended and the accuracy attained and conclude that "people exhibit intelligent, if not optimal, responses to changes in ... task and context variables." (p. 249) This adaptivity is evidence that decision makers have a choice with regards to incurring a cognitive cost or not, and make this choice strategically. There is also physiological evidence of conscious "executive control" over cognitive processes. Baker et al. (1996) isolate those areas of the brain that choreograph these executive decisions, noting that they become activated when people are weighing choices.

Again, the ability to decide on cognitive effort is so widely accepted that it is often taken as a *premise* in the cognitive sciences. In current research on information processing, persuasion and social cognition, an almost paradigmatic metaphor for the decision maker is that of a "motivated tactician," who Fiske and Taylor (1993) describe

---

[27] The "engineering psychology" literature (e.g. Gopher and Kimchi (1989) and Wickens and Kramer (1985)) focuses on the optimal presentation of information to human operators.

[28] Sometimes this is referred to as "cost-benefit theory" in the field of decision making. Early papers that consider this or similar trade-offs are Yates and Kulick (1977), Beach and Mitchell (1978), Christensen-Szalanski (1978, 1980), Payne (1982), Johnson and Payne (1985), Russo and Dosher (1983) and Shugan (1980).

as "a fully engaged thinker who has multiple cognitive strategies available and chooses among them based on goals, motives, and needs." (p. 13) According to Molden and Higgens (2005), this "tactician" balances the benefits of reasoning - increased accuracy, for one - with the costs of effort.[29] Empirical research has found that individuals respond to raising the stakes of accuracy by expending more effort to be accurate (Kruglanski and Freund 1983, Freund et al. 1985), for example by considering more alternatives and coming up with more complicated explanations (e.g. Tetlock and Kim 1987). Maheswaran and Chaiken (1991) found experimental evidence that pieces of information continue to be processed until a desired level of accuracy is reached, a policy they call the "sufficiency principle", which goes back to the notion of "satisficing" by Simon (1955).

There is support for our assumptions in the jury literature as well. To a large extent, this literature focuses more on practical issues, and less on empirical tests of the "best" cognitive model of juror behavior.[30] However, often while attacking more practical issues, these authors find supporting evidence for our assumption that jurors trade off the benefits of a more accurate verdict with cognitive costs. A number of examples are mentioned below. Forsterlee et al. (1997) and Fosterlee et al. (2005) find that note-taking encourages more complicated evidence and more complex arguments to be considered by a jury, which points directly to the cognitive constraints the jury faces. As stated above, Maheswaran and Chaiken (1991) find that jurors process pieces of information until the desired level of accuracy is reached, where this level is responsive to the importance that the subject attaches to it. Weinstock and Flaton (2004) find more support for the optimization under cognitive constraints: they document that those jurors who are more certain of their verdict are less apt to process additional information, while those that are less certain will likely utilize supplementary pieces of evidence. Other indicative support comes from jury responses to changes in the

[29]See also an early overview by Kunda (1990).

[30]Some common practical issues in this literature are stereotyping, the impact of expert witnesses, the use of jury instructions, and the impact of sentencing options. See Devine et al. (2000) for a survey.

presentation of evidence. For example, Bourgeois et al., (1993) find that complex evidence that is explained in non-technical terms is taken into account more often and more thoroughly. Petty and Cacioppo (1986) find the same when evidence is presented twice.

In sum, we have presented a model of learning and decision-making that is consistent with the general literature in cognitive psychology, and more specifically consistent with a variety of studies of juror behavior.

7 Conclusion

We have presented a formal model of an important principle of evidence in common-law legal systems: exclusion by reason of undue prejudice. The key novelty of the model is that the fact-finders (jurors) have a cognitive cost of processing evidence. We have shown that this assumption is well grounded in the psychological literature.

Within this framework, the judge excludes evidence in order to incentivize the jury to focus on other, more probative evidence. Exclusion is not, therefore, a countermeasure to irrational updating on the part of the jury; rather, it is a way to incentivize jurors that are "cognitive misers." We studied the comparative statics properties of this model, and we have shown that some fairly intuitive properties do not always hold. For example, making evidence more probative may lead it to be optimally excluded. Similarly, a judge may optimally choose to exclude more evidence when faced with a more competent jury. Finally, we have shown that the decision to exclude evidence is a contextual one: optimal exclusion of one piece of evidence requires taking into account its relationship with all other pieces of evidence. We interpret these counterintuitive properties and complexities as evidence that general rules mandating exclusion are unlikely to be optimal. In our model, optimal exclusion is achieved straightforwardly by giving the judge broad exclusionary powers. This is, of course, the arrangement that prevails in current procedure.

We also provided sufficient conditions under which exclusion is not helpful. This is the case when, roughly speaking, all the available evidence fits together tightly into one coherent story (formally, when all pieces of evidence are complementary with each other). This configuration of evidence is arguably more likely to arise in the inquisitorial systems typical of continental law, in which evidence-gathering is carried out by one agent (typically, a judge). In adversarial systems, the evidence is gathered by two opposing parties, and so it is unlikely to all fit together into a coherent story. Rather, the evidence gathered by each party is likely to fit together into a coherent story, but the two stories need not fit together with each other. In this case, we have shown that it may be optimal to exclude some part of a story. We interpret this property as consistent with the regularity that exclusion by reason of prejudice is found almost exclusively in adversarial (common law) systems.

Beyond the specific contributions mentioned above, this paper can be seen in a broader context. A broader contribution of this paper, in our view, is to introduce evidence rules as a a new and potentially important area of application for the theory of endogenous information acquisition. Information theorists have analyzed other aspects of judicial decision-making as optimally designed schemes to incentivize information acquisition. The rich area of evidence rules, however, has hitherto received little attention. If, as we propose here, some rules of evidence can profitably be interpreted as devices to induce more accurate decision-making by the fact finder, then a whole body of rules potentially opens up for analysis.

8 Appendix

8.1 Derivation of Values in Table 1

The entries in Table 1 are simple to calculate. For example, consider the expected payoff to the juror from processing E_1. First, the juror incurs processing cost -5. With probability $\frac{1}{5}$ the signal reveals that the driver's speed was in the interval $[0, 20]$, and the juror can rule with certainty that the car was traveling at less than fifty miles per hour. His payoff from this correct ruling is 0. On the other hand, with probability $\frac{4}{5}$ the signal reveals that the driver's speed was in the interval $[20, 100]$. Conditional on this realization, the juror can deduce that the car was traveling less than fifty miles per our with probability $\frac{3}{8}$ and more than fifty miles per our with probability $\frac{5}{8}$. Therefore, the juror's expected payoff from ruling that $x \leq 50$ is $\left(\frac{3}{8}\right)(0) + \left(\frac{5}{8}\right)(-100)$, while the expected payoff from ruling that $x > 50$ is $\left(\frac{5}{8}\right)(0) + \left(\frac{3}{8}\right)(-100)$. Clearly, then, the juror optimally rules that the car's speed exceeded fifty miles per hour. The ex-ante expected payoff from processing E_1 is thus

$$-5 + \left(\frac{1}{5}\right)(1)(0) + \left(\frac{4}{5}\right)\left(\frac{3}{8}\right)(-100) = -35.$$

The payoff to society is the same, except that the processing costs are ignored. All other entries are calculated in an identical fashion.

8.2 Extending the Simple Example

In this section, we generalize the simple example that has been employed throughout the paper in order to further explain the intuition behind our results. To that end, again let the decision and the payoffs of the juror, as well as the properties of x and E_2, remain as specified in Section 2. However, let E_1, with cost $c_1 = 5$, reveal whether or not x lies in the region $[0, \xi]$, for some value of $\xi \in [0, 50]$. Figures 2 and 3 illustrate the expected payoffs to the juror and society, respectively, as the value of ξ varies from 0 to 50. Note than when ξ is small E_1 has little informational content, and as ξ grows

it becomes more informative.[31]

[Insert Figure 2 Here]

[Insert Figure 3 Here]

When $\xi \leq 5$, E_1 is sufficiently uninformative that the juror will not incur the processing costs. In this region, the juror chooses only to process E_2, and an increase in the informational content of E_1 has no effect. When $\xi \in [5, 15]$, the juror chooses to process both E_1 and E_2, so that the first best is achieved. However, when $\xi \in [15, 40]$, the juror chooses to process only E_1 and the payoff to society decreases. Finally, when $\xi \in [40, 50]$, E_1 is the more informative signal. In this region, the juror processes only E_1, and the payoff to society is increasing in the informational content of E_1. Consistent with proposition 1, it is easy to see that an increase in the informational content of E_1 can lead to either better or worse outcomes.

In Figure 4 below, we plot the expected payoff to society conditional on the optimal exclusion policy being implemented. Notice first that E_1 and E_2 are complements when $\xi \leq 10$ and substitutes when $\xi > 10$. Therefore, E_1 switches from a complement to a substitute for E_2 as it becomes more informative. Also notice that E_1 is allowed for all $\xi \leq 10$, reflecting that complementarity is a sufficient condition for evidence to be admitted, as summarized in proposition 5.2. However, that $\xi \in [10, 15] \cup [40, 50]$ is also optimally allowed illustrates that complementarity is *not* a necessary condition for evidence to be admitted, and indeed the usefulness of proposition 5.2 may be limited. Figure 4 further highlights the absence of general principles guiding exclusion as it illustrates that the optimal decision rule of the judge is non-monotonic in the informational content of the information system (as summarized in proposition 1), as

---

[31] Consider signal E_1^B with corresponding value ξ^B and signal E_1^A with corresponding value ξ^A, such that $\xi^B \geq \xi^A$. Strictly speaking, E_1^B is not more informative than E_1^A in the sense of Blackwell, since E_1^B is not necessarily more informative than signal E_1^A for **all** decision problems. However, **for this decision problem** E_1^B is more informative than E_1^A, and thus the example that follows is a perfectly adequate mechanism to illustrate the effects of evidence becoming more informative.

are the expected payoffs to society (as summarized in proposition 1).

[Insert Figure 4 Here]

References

AASMAN, J., G. MULDER, AND L. MULDER (1987): "Operator effort and the measurement of heart-rate variability," *Human Factors*, 29(2), 161–170.

BACKS, R., AND K. SELJOS (1994): "Metabolic and cardiorespiratory measures of mental effort: the effects of level of difficulty in a working memory task," *International journal of psychophysiology*, 16(1), 57–68.

BAKER, S., R. ROGERS, A. OWEN, C. FRITH, R. DOLAN, R. FRACKOWIAK, AND T. ROBBINS (1996): "Neural systems engaged by planning: a PET study of the Tower of London task," *Neuropsychologia*, 34(6), 515–526.

BEACH, L., AND T. MITCHELL (1978): "A Contingency Model for the Selection of Decision Strategies," *The Academy of Management Review*, 3(3), 439–449.

BENTHAM, J. (1827): *Rationale of Judicial Evidence, Specially Applied to English Practice*. ed. by J.S. Mill, 5 volumes, Hunt and Clarke.

BORGERS, T., A. HERNANDO-VECIANA, AND D. KRAHMER (????): "When are Signals Complements or Substitutes," Manuscript, University of Michigan, July 15 2007.

BROADBENT, D. (1958): *Perception and communication*. Pergamon Press New York.

CALDER, B., C. INSKO, AND B. YANDELL (1974): "The Relation of Cognitive and Memorial Processes to Persuasion in a Simulated Jury Trial," *Journal of Applied Social Psychology*, 4(1), 62–93.

CHEN, S., AND S. CHAIKEN (1999): "The heuristic-systematic model in its broader context," *Dual-process theories in social psychology*, pp. 73–96.

CHRISTENSEN-SZALANSKI, J. (1978): "A Mechanism for Decision Strategy Selection and Some Implications.," *Organizational Behavior and Human Performance*, 22, 307–323.

CHRISTENSEN-SZALANSKI, J. (1980): "A further examination of the selection of problem-solving strategies: The effects of deadlines and analytic aptitudes," *Organizational Behavior and Human Performance*, 25, 107–122.

CRAIK, F., AND R. LOCKHART (1972): "Levels of processing: A framework for memory research," *Journal of Verbal Learning and Verbal Behavior*, 11(6), 671–684.

DAMASKA, M. (1997): *Evidence Law Adrift*. Yale University Press.

DE NEYS, W., W. SCHAEKEN, AND G. D'YDEWALLE (2002): "Semantic Memory Retrieval During Conditional Reasoning: Every Counterexample Counts," *Proceedings of the 24th Cognitive Science Conference. Mahwah: Erlbaum*.

DEVINE, D., L. CLAYTON, B. DUNFORD, R. SEYING, AND J. PRYCE (2001): "Jury decision making: 45 years of empirical research on deliberating groups," *Psychology, Public Policy, and Law*, 7(3), 622–727.

EAGLY, A., AND S. CHAIKEN (1993): *The psychology of attitudes*. Harcourt Brace Jovanovich College Publishers Fort Worth.

FEDDERSEN, T., AND A. SANDRONI (2006): "Ethical Voters and Costly Information Acquisition", *Quarterly Journal of Political Science*, 1: 287–311

FIBIGER, W., O. EVANS, AND G. SINGER (1986): "Hormonal responses to a graded mental workload," *European Journal of Applied Physiology*, 55(4), 339–343.

FISKE, S. (1993): "Social Cognition and Social Perception," *Annual Review of Psychology*, 44(1), 155–194.

FORSTERLEE, L., AND I. HOROWITZ (1997): "Enhancing Juror Competence in a Complex Trial," *Applied Cognitive Psychology*, 11(4), 305–319.

FORSTERLEE, L., I. HOROWITZ, AND M. BOURGEOIS (1994): "Effects of notetaking on verdicts and evidence processing in a civil trial," *Law and Human Behavior*, 18(5), 567–578.

FORSTERLEE, L., L. KENT, AND I. HOROWITZ (2005): "The cognitive effects of jury aids on decision-making in complex civil litigation," *Applied Cognitive Psychology*.

FREUND, T., A. KRUGLANSKI, AND A. SHPITZAJZEN (1985): "The Freezing and Unfreezing of Impressional Primacy: Effects of the Need for Structure and the Fear of Invalidity," *Personality and Social Psychology Bulletin*, 11(4), 479.

GERARDI, D., AND L. YARIV (2008a): "Information acquisition in committees," *Games and Economic Behavior*, 62, 436–459

——— (2008b): "Costly Expertise," Discussion paper, Forthcoming, *American Economic Review, Papers and Proceedings*.

GERSBACH, H. (1995): "Information efficiency and majority decisions," *Social Choice and Welfare*, 12(4), 363–370.

GERSHKOV, A., AND B. SZENTES (2008): "Optimal Voting Schemes with Costly Information Acquisition," *Journal of Economic Theory (Forthcoming)*.

GLAESER, E., AND A. SHLEIFER (2002): "Legal Origins," *Quarterly Journal of Economics*, 117, 1193–1230.

GOPHER, D., AND R. KIMCHI (1989): "Engineering Psychology," *Annual Review of Psychology*, 40(1), 431–455.

HASTIE, R. (1993): *Inside the Juror: The Psychology of Juror Decision Making*. Cambridge University Press.

——— (2001): "Problems For Judgment And Decision Making," *Annual Review of Psychology*, 52(1), 653–683.

HEUER, L., AND S. PENROD (1994): "Trial complexity," *Law and Human Behavior*, 18(1), 29–51.

——— (1994a): "Juror notetaking and question asking during trials," *Law and Human Behavior*, 18(2), 121–150.

JAIN, S., AND D. MAHESWARAN (2000): "Motivated Reasoning: A Depth-of-Processing Perspective," *The Journal of Consumer Research*, 26(4), 358–371.

JOHNSON, E., AND J. PAYNE (1985): "Effort and Accuracy in Choice," *Management Science*, 31(4), 395–414.

JONIDES, J. (1997): "Verbal Working Memory Load Affects Regional Brain Activation as Measured by PET," *Journal of Cognitive Neuroscience*, 9(4), 462–475.

KAHNEMAN, D. (1973): *Attention and effort*. Prentice-Hall Englewood Cliffs, NJ.

KRUGLANSKI, A., AND T. FREUND (1983): "The freezing and unfreezing of lay-inferences: Effects on impressional primacy, ethnic stereotyping, and numerical anchoring," *Journal of Experimental Social Psychology*, 19(5), 448–468.

KRUGLANSKI, A., AND E. OREHEK (2007): "Partitioning the Domain of Social Inference: Dual Mode and Systems Models and Their Alternatives," *Annual Review of Psychology*, 58, 291–316.

KUNDA, Z. (1990): "The case for motivated reasoning," *Psychological Bulletin*, 108(3), 480–498.

LEEDY, C., AND L. DUBECK (1971): "Physiological changes during tournament chess," *Chess Live and Review*, 26, 708.

LOH, W. (1985): "The evidence and trial procedure: The law, social policy, and psychological research," *The psychology of evidence and trial procedure*, pp. 13–39.

LUND-ANDERSEN, H. (1979): "Transport of glucose from blood to brain," .

MAHESWARAN, D., AND S. CHAIKEN (1991): "Promoting systematic processing in low-involvement settings: Effect of incongruent information on processing and judgment," *Journal of Personality and Social Psychology*, 61(1), 13–25.

MARTINELLI, C. (2006): "Would Rational Voters Acquire Costly Information?," *Journal of Economic Theory*, 129, 225-251.

MARTINELLI, C. (2007): "Rational Ignorance and Voting Behavior," *International Journal of Game Theory*, 35, 315-335.

MELLERS, B., A. SCHWARTZ, AND A. COOKE (1998): "Judgment and Decision Making," *Annual Review of Psychology*, 49(1), 447–477.

MILLER, G. (1956): "The magic number seven, plus or minus two," *Psychological Review*, 63(2).

MIYAKE, A., AND P. SHAH (1999): *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*. Cambridge University Press.

MOLDEN, D., AND E. HIGGINS (2005): "Motivated thinking," in *The Cambridge Handbook of Thinking and Reasoning*, ed. by K. Holyoak, and R. Morrison. Cambridge University Press.

MUKHOPADHAYA, K. (2003): "Jury Size and the Free Rider Problem," .

MULDER, L., AND G. MULDER (1987): "Cardiovascular reactivity and mental workload," in *The Beat-by-Beat investigation of cardiovascular function*, ed. by O. R. . R. Kitney, pp. 216–253.

NAVON, D., AND D. GOPHER (1979): "On the Economy of the Human-Processing System.," *Psychological Review*, 86(3), 214–55.

NEWELL, A., AND H. SIMON (1972): *Human problem solving*. Prentice-Hall.

NORMAN, D., AND D. BOBROW (1975): "On Data-limited and Resource-limited Processes.," *Cognitive Psychology*, 7(1), 44–64.

OPERARIO, D., AND S. FISKE (1999): "Social Cognition Permeates Social Psychology: Motivated Mental Processes Guide the Study of Human Social Behavior," *Asian Journal Of Social Psychology*, 2(1), 63–78.

PASHLER, H. (1998): *The Psychology of Attention*. MIT Press.

PAYNE, J. (1982): "Contingent Decision Behavior: A Review and Discussion of Issues.," DTIC Research Report ADA111655.

PAYNE, J., J. BETTMAN, AND E. JOHNSON (1993): *The Adaptive Decision Maker*. Cambridge University Press.

PERSICO, N. (2004): "Committee Design with Endogenous Information," *Review of Economic Studies*, 71(1), 165–191.

PETTY, R., AND J. CACIOPPO (1986): "Elaboration likelihood model," *Advances in experimental social psychology*, 19, 123–205.

POSNER, R. (1999): "An Economic Approach to the Law of Evidence.," *Stanford Law Review*, 51(6).

RUSSO, J., AND B. DOSHER (1983): "Strategies for multiattribute binary choice," *Journal of experimental psychology. Learning, memory, and cognition*, 9(4), 676–696.

SANCHIRICO, C. (2001): "Character Evidence and the Object of Trial," *Columbia Law Review*, 101(6), 1227–1311.

SCHRAG, J., AND S. SCOTCHMER (2002): "Crime and Prejudice: The Use of Character Evidence in Criminal Trials," *Journal of Law, Economics, and Organization*, 10(2), 319–342.

SIMON, H. (1955): "A Behavioral Model of Rational Choice," *The Quarterly Journal of Economics*, 69(1), 99–118.

SHALLICE, T., AND P. BURGESS (1991): "Higher-order cognitive impairments and frontal lobe lesions in man," *Frontal lobe function and dysfunction*, pp. 125–138.

SHUGAN, S. (1980): "The Cost of Thinking," *The Journal of Consumer Research*, 7(2), 99–111.

STEPHENSON, M. (2007): "Bureaucratic Decision Costs and Endogenous Agency Expertise," *Journal of Law, Economics, and Organization*, 23(2), 469.

STUSS, D., D. BENSON, ET AL. (1986): *The Frontal Lobes*. Raven Press New York:.

TULLOCK, GORDON (1996): Legal Heresy: Presidential Address to the Western Economic Association, 1995. Economic Inquiry 34(1), p. 1-9.

TURNER, J., AND D. CARROLL (1985): "Heart rate and oxygen consumption during mental arithmetic, a video game, and graded exercise: Further evidence of metabolically-exaggerated cardiac adjustments," *Psychophysiology*, 22(3), 261–267.

VINCENT, A., F. CRAIK, AND J. FUREDY (1996): "Relations among memory performance, mental workload and cardiovascular responses," *International Journal of Psychophysiology*, 23(3), 181–198.

WEINSTOCK, M. (2005): "Cognitive bases for effective participation in democratic institutions: Argument skill and juror reasoning," *Theory and Research in Social Education*, 33(1), 73–102.

WEINSTOCK, M., AND R. FLATON (2004): "Evidence coverage and argument skills: cognitive factors in a juror's verdict choice," *Journal of Behavioral Decision Making*, 17(3), 191–212.

WICKENS, C., AND A. KRAMER (1985): "Engineering Psychology," *Annual Review of Psychology*, 36(1), 307–348.

YATES, J., AND R. KULICK (1977): "Effort control and judgments," *Organizational Behavior and Human Performance*, 20(1), 54–65.

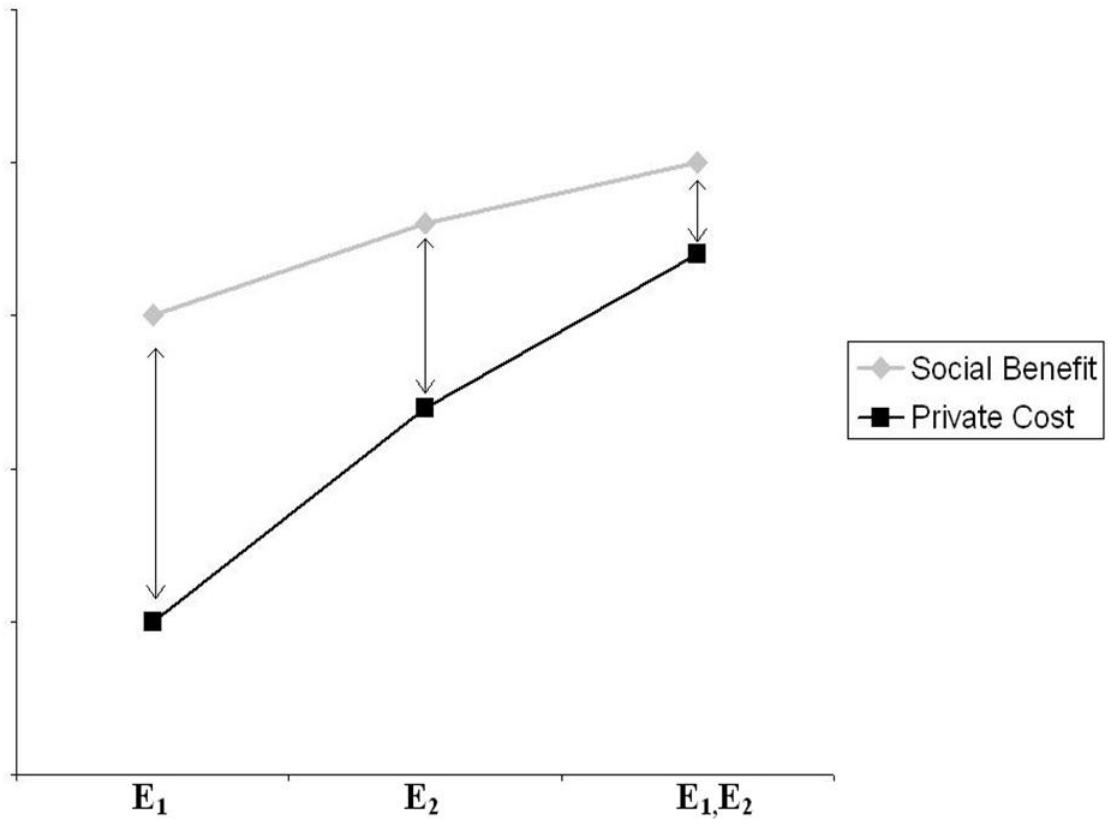Figure 1: Social Benefits and Private Costs
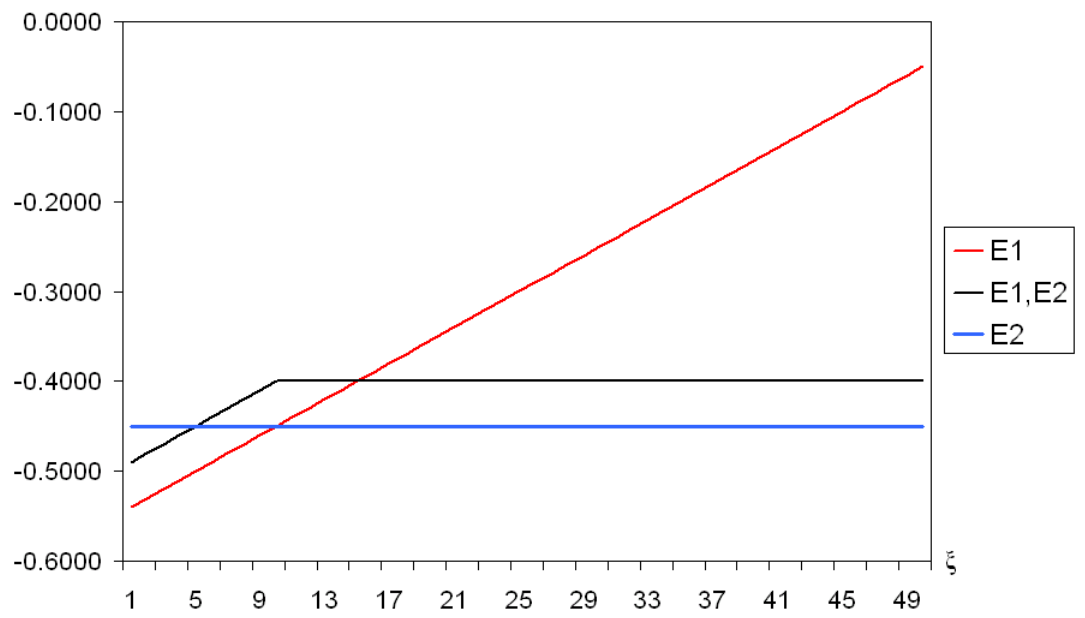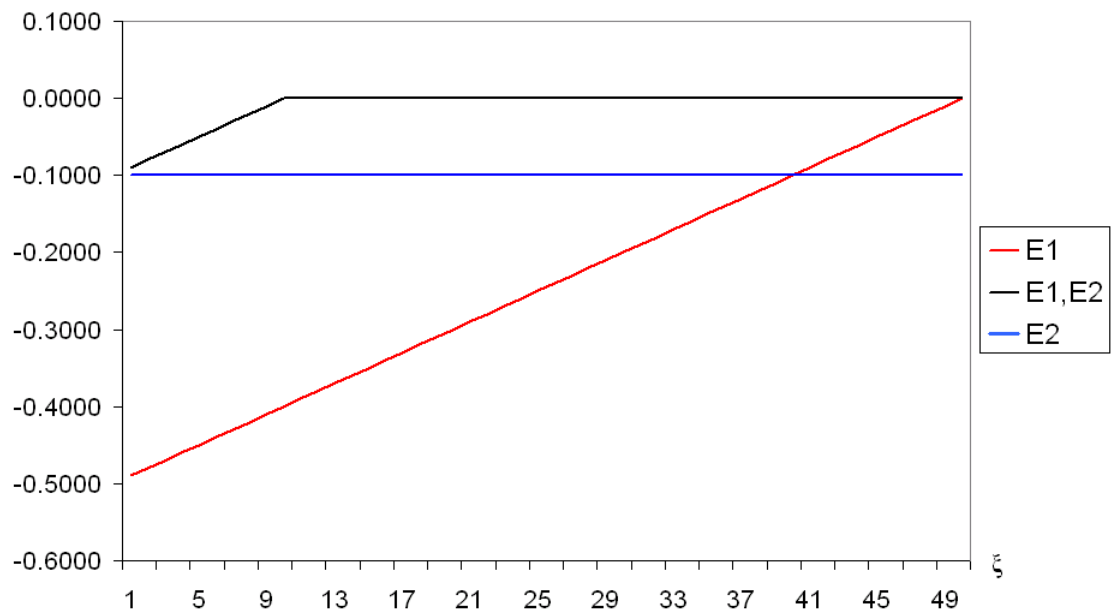
Figure 2: Juror's Expected Payoffs

Figure 3: Society's Expected Payoffs

Figure 4: Optimal Policy and Payoffs