Judicial Consistency: Legal Evidence*

Andrew Caplin[†], Andrei Gomberg[‡] and Joyce Sadka[§]

October 17, 2018

Abstract

Consider a decision-maker (a judge) who repeatedly issues binary (Guilty or Innocent) verdicts in a series of ex ante similar cases. Suppose we observe, besides the verdict, a record of evidence, testifying to the facts of the case, based on which the determination is made. Such a record would allow us to associate with each verdict an event observed by the decision-maker. We provide a theoretical characterization of restrictions on such data consistent with a simple evidence-based Bayesian model of judicial decision-making and use the data to reveal the underlying parameters of the model. We then apply our results to an empirical data set derived from a large sample of case files of firing lawsuits in a Mexican labor court. Very preliminary results show consistency within but not across judges in mapping legal facts to the verdict, and inconsistencies within judge and across judges in mapping observed fact patterns to legal facts.

1 Introduction

"What has the court done?" In this paper we shall be asking this question in a very specific setting, considering decisions of a small group of judge's clerks, resolving cases in a relatively obscure legal institution: a labor court on the outskirts of Mexico City. At the same time, we attempt to conduct this specific inquiry in a manner that, we believe, would be of use in approaching judicial decision-making in general. By taking as evidence the corpus of decisions from a homogeneous group of cases, we believe we may reveal both the preferences and biases of judges, and the legal rules they implement.

While our court might lack on factual intrigue and legal brilliance of more famous chambers, its very routine and the repetitiveness of its task have combined to create an environment tractable and transparent enough to make feasible an almost literal application of basic decision-theoretic concepts to its analysis, while the nature of Mexican legal system has made it possible to collect data that would be difficult to obtain elsewhere. In particular, we exploit the fact that authors of decisions are not present at hearings, but base their decisions on the same written file that we have access to. At the same time, by law they have to produce motivated decisions, listing both the facts as they observe them and legal determinations on which they base their verdicts. Collecting this information gives us a unique look into the court decision process.

*EXTREMELY PRELIMINARY AND INCOMPLETE. Please do not circulate. We would like to thank Diego Jimenez, Felipe Gomez, Alejandro Hernandez Gonzalez, Misael Ramirez Rangel, Pablo Mendoza, and Diego Villegas for research assistance. We would like to thank participants in the FUR 2014 conference in Rotterdam, 2015 NYU Alumni conference, 2016 APET conference, and seminars in HSE in Moscow and ITAM for valuable comments. Financial support from the Associación Mexicana de Cultura and Interamerican Development Bank is gratefully acknowledged.

[†]NYU

[‡]ITAM

[§]ITAM

By developing our theoretical tools to attack this simple but detailed empirical environment we believe we are able to ask questions that have not, till now, been asked that, in the end, may be fundamental to understanding the nature of what courts - including those of greater fame and sophistication - do. In particular, we can extend the model of judicial decision-making to account for different stages of the process in a manner that creates testable implications for our data set. Furthermore, by considering these data in light of our model we may be able to disentangle various motivations that drive judicial decision-making. In this paper we concentrate on the role played by legal rules the judges are to implement, as well as on their individual beliefs and/or preferences, which may be crucial to how they establish the facts of the case.

The general question of "what the courts do" is the subject of Cameron and Kornhauser (2017, henceforth CK) synthesis of modern developments in the formal modeling of courts. They summarize the task of a court as, firstly, fact-finding: observing and aggregating evidence in the case into a collection of legal facts - and, secondly, adjudication - applying legal rules to the set of observed facts in order to have the case properly disposed. In the latter task, presumably, the courts are more constrained by law and precedent. Disparity in legal outcomes based on the same set of legal facts would strike at the very heart of the notions of due process and legal treatment inherent in the administration of justice. In contrast, fact-finding may be a task opened to a greater degree of discretion, with different reasonable decision-makers coming to different conclusions in establishing what really happened from noisy evidence presented in the case, based perhaps on preconceived beliefs about the likelihood of various events. It is, perhaps, understandable that, as acknowledged by CK, the bulk of the literature has concerned itself with the cleaner task of adjudication.

Nevertheless, even though the fact-finding stage might allow a somewhat greater role for judicial discretion, this discretion, in the words of Chief Justice Roberts, "is not whim, and limiting discretion <...> helps promote the basic principle of justice that like cases should be decided alike." (Martin vs. Frank Capital Corp., 2005). Hence, a crucial importance that has always been ascribed by legal research to the study of patterns that relate judicial outcomes to observed factual evidence. The difficulties involved, however, are substantial. To begin with, in most environments an outside researcher's ability to observe the same evidence as perceived by the judge may be limited. Even if this problem can be overcome (as we shall claim is the case in our environment), we may not be sure which combination of evidentiary facts a judge takes to be relevant for her decision. Furthermore, patterns of such facts observed in actually resolved cases naturally represent only a selection from the universe of all possible factual combinations, so that it may not necessarily be clear how the existing observations would relate to the proper resolution of similar cases with somewhat different factual combinations (Kort 1963). Finally, even if we could perfectly observe facts of the case as presented to the judge and establish the appropriate legal rule the particular jurisdiction had previously applied, we would still face the possibility that the judge's perception of facts may be distinct from that of an outside observer. Judges are human, and like all human decision-makers they may commit errors, perhaps due to failure to observe or interpret available evidence.

Thus, if we, as outside observers, note apparent inconsistency between judicial outcomes in apparently similar cases, we have to consider distinct possibilities that call for distinct responses. It could be that the judges differ in their interpretation of legal rules applicable in a given case: if it is observed, such inconsistency might call for limiting future discretion either through legislation or precedent . It could also be that the disagreements arise from differences in what the judges infer from imperfectly observed evidence based on their personal beliefs and biases: in a certain sense these constitute the substance of the fact-finding task performed by judges or juries. Or else the apparent inconsistencies might arise from outright judicial mistakes, "lapses of attention" as they may be and should, perhaps, be dealt with by improving judicial incentives for avoiding them.

It may be argued, however, that judicial decision-making naturally generates extremely rich data sets, that could be used to distinguish between different sources of judicial inconsistency. Courts keep detailed evidence records (and, in the case of the court we shall be talking about in this study, it is this written evidence record is the only basis on which judges may be basing their decisions). Furthermore, they typically provide written decisions, which contain not merely eventual disposition of the case, but also multiple findings of legal facts and reasoning. Finally, courts typically deal with large number of similar but distinct cases. In fact, an output of almost any court is much richer than whatever is available to a student of most other real-life decisions, be those investment decisions by a firm, individual consumer choices, electoral outcomes or even diagnostic determinations by doctors. It is therefore natural to attempt to use this data to empirically determine the distinctions we have postulated here.

The idea of using this data to answer questions of judicial consistency is, of course, not new. As early as Lawror (1963) and Kort (1957, 1963) proposes emerged for deriving the legal rules judges use based on the logical patterns of data in the observed cases. A number of studies concentrated on applying statistical techniques to precedent-setting decisions by leading courts. Thus, Segal (1984) used established legal facts of the case to construct a logit model of the US Supreme Court decisions in search and seizure cases and, more recently, Kastellec (2010) refined this work by applying non-parametric decision-tree techniques to essentially the same data set. In another recent study Niblett (2013) analyses the content of the written decisions of the California Court of Appeals to document inconsistencies between its decisions on contract enforceability. These studies, like ours, concentrate on the nature of legal cases involved, and attempt to exploit the actual data on the content of the legal controversy. More broadly, this literature is a part of a more general agenda establishing and measuring and explaining the sources of legal inconsistencies between judges and/or courts (see Fischman 2013 for a recent survey). This literature has included both large statistical analyses of administrative courts (e.g. Ramji-Nogales 2007), experiments in which judges were polled about their opinions of hypothetical cases (e.g. Van Koppen and Ten Kate 1984) and studies of disagreements between judges of multi-member courts (e.g. Alarie and Green 2007).

It is our belief that the natural richness of data sets generated by court actions has so far been underexploited as means of understanding of judicial inconsistencies. A natural starting point here could be found in the stochastic choice theory, which tries to explain what appear to be inconsistencies in usual choice data that economists deal with. A typical proposed observation in this framework is formulated as a probability an alternative would be chosen out of a given set (something that would empirically require a large number of observations from homogeneous environments). Going back to the early work by Luce (1958) and Block and Marshak (1960), testable implications on such a data set of the Random Utility Model have been studied, in which variability of choices was related to that of individual preferences. While, arguably, this approach would be somewhat unsatisfactory in a legal environment, in which consistent application of deterministic legal rules may be viewed as desirable, stochastic variability may obtain from shocks to beliefs and/or attentional constraints. This has, in fact, been the focus of recent work such as Caplin and Dean (2015), Caplin and Martin (2015) and Matejka and McKay (2015). What has emerged, in part, from hits literature is the importance of expanding the notion of observable data to include other features, such as "true" state of the world ("state-contingent stochastic choice), individual beliefs, self-corrected errors, etc. We propose here that information contained in legal records provides one such extension that is both empirically feasible and useful to answering the questions of legal consistency referred above. At the same time, both the data and the questions here considered are novel for the choice literature itself. In particular, the question of empirical identification of legal rules has not been properly considered in a manner consistent with the individual choice model. Therefore, in this, the first paper of this study we shall address that issue, before building it into the stochastic choice model which may allow us to incorporate issues of judicial attention.

In practical terms, we propose a simple model of judicial decision and concentrate on deriving its implications for observables that are generated by the court. Analysis of a judge's legal output using our theoretical framework allows us to empirically derive both the legal rules that judges adhere to, and individual biases that influence their rulings. We believe that modeling judicial decision-making is essential to be able to attempt to distinguish between disagreements and inconsistencies based on the former and the latter. Furthermore, our approach allows us to ask novel questions that, until now, may not have been naturally formulated. Thus, it has long been recognized that identifying inconsistencies implies taking a stand on what constitutes meaningful distinctions between cases: by making finer and finer distinctions any case may be presented as unique and, hence, deserving a disposition unconstrained by any prior precedent. However, to the best of our knowledge it has never been shown that making additional distinctions between cases cannot introduce what would appear to be new inconsistencies in the legal output of a judge or a court. Within our model we demonstrate that in fact, such monotonicity requires a certain "product" structure of the case space. Once we impose this structure, we can discuss what constitutes a minimal set of case distinctions that makes a given collection of cases internally consistent.

It should be reiterated that our approach is not merely theoretical. It is, in fact, strongly motivated by the features of an actual institution to which we have access and from which we can extract a novel data set that we are using to illustrate and test our results. As part of a long-term field project on administration of justice in Mexican labor courts conducted by one of us, we have obtained access to complete legal files, containing both admitted evidence and first instance judgments in a large collection of cases involving employer-employee disputes. Crucially, we observe the entire evidence seen by the decision-maker: the people who prepare these decisions are never present at court and decide based on the same case file that we have In addition, of course, we have actual judicial decisions, which besides an eventual disposition ("the verdict ") contain determination of legal facts used in reaching the final decision. By concentrating on a homogeneous subset of these cases, we compile a data set that corresponds to the one we propose in this paper. In order to abstract from the "mistakes" inherent in analyzing complicated legal files by the judges (something we intend to address in a sequel to this paper) we also conducted a lab-in-field experiment in which the same judges answered questions about a set of hypothetical cases we formulated. Finally, we were able to administer a field experiment in which we exogenously improved readability of case files, allowing for random variation of the cost of judicial attention. Data from all these sources, still being compiled, is uniquely suited to figuring out the sources of judicial inconsistency.

It should also be stressed that at this point we are very much presenting a work in progress. Our data effort is highly ambitious, and, at this point, given resources available to us so far, even coding of the initial data set will take close to another year. Furthermore, as noted above, in this paper we concentrate on identifying from data the structure of the legal rules used by the judges and possible individual biases they bring into their decision-making. The issue of identifying judicial (in)attention, which was our original motivation for this project, will be explored in a sequel to the present paper.

The rest of the paper is organized as follows. Section 2 discusses the legal environment and the data extracted from case files and court decisions. Section 3 presents the model we use to analyze the data . Section 4 provides some simulation results using randomly generated data.. Section 5 presents some of the preliminary empirical results. Section 6 concludes and discusses possible extensions of this research agenda.

2 Legal environment and data

As part of a set of experiments aimed at testing policies to improve the administration of justice, we obtained access to a large data set of decisions coming from administrative labor courts in Mexico. These are administrative tribunals that belong to the executive branch of government at the state or federal levels. Labor law is federal in Mexico, so the legal environment is the same across the country. Most labor lawsuits are firing related. The law classifies firing in only two ways: fair or unfair. Fair firing does not lead to any severance pay. A finding of unfair firing means substantial severance pay, and the specific amount of money award is determined in the court's decision. Each decision establishes a number of legal facts, such as whether there was, in fact, an employee/employer relationship between the plaintiff and the defendant, whether the firing occurred, and if it did whether it was fair or unfair; it also declares whether the decision is in favor of the firm, in favor of the worker, or "mixed", and gives a specific number for the total award (while detailing how the amount is arrived at) in cases of unfair firing.

It is important to note that the draft decisions are not produced by the judge who participates in the hearings process. Rather, each "court location" has a number of draft decision writers (we will call them clerks) who have the same formal rank as judges' assistants, but are administratively independent from each individual labor court. Once the judge's assistant declares a case file "closed", the case file is sent to the "draft decisions area" where one of the clerks reads the case file and produces a draft decision. According to the law, judges must review all draft decisions; in practice we believe very few are reviewed, and we know that in only around 5% of all cases does a judge ask for any change or correction in the draft decision. Finally an important institutional feature of these courts is the built-in random assignment. To begin with when case files enter the "court location" they are assigned in a round-robin (quasi-random) fashion to the individual courts. Secondly, when case files arrive at the decision writing area, they are also assigned to clerks randomly.

The clerks assigned to produce written decisions effectively act in the capacity of a standard first-instance court (jury/judge) in that they decide issues of fact, based on the evaluation of evidence, and issues of law, as they apply legal statutes and jurisprudence (accepted body of interpretations of labor law made by the Mexican Supreme Court). Jurisprudence is particularly important in the application of Mexican labor law because the law itself does not clearly state how burdens of proof are determined in different legal controversies, but jurisprudence contains at least some guidelines, which, as we will see later, are not always applied consistently by the judges.

Due process principles imply that the clerks should consider all evidence (and only the evidence) which is offered by one of the parties and legally admitted by court. In practice, we commonly observe a number of violations of due process. For the purposes of our study it is important that written decisions frequently ignore existence or admission of evidence. Supervision of the clerks by the judges, however imperfect, provides incentives for clerks writing the decisions to avoid mistakes. More importantly, mistakes may lead to successful appeals by the parties to the case. A granted appeal generally results in the a new judgment having to be written, and is assigned to the same clerk who wrote the first decision.

For each clerk, we have multiple observations of decisions. In addition to issuing a verdict (essentially, a series of binary determinations of legally established facts, followed by binary conclusion on whether there was an illegal firing, which is then followed by a separate determination of damages), the clerk must provide a written summary of the evidence contained in the file as part of the legal motivation (they may be assumed to be incentivized to do this properly, as failure to do so may result in an appeal, responding to which would increase their own workload). It should be noted that case files vary substantially in their volume, the number of issues considered, clarity, etc. The content of each written decision is determined by law. In particular, article 840 of Mexico's *Federal Labor Law* specifies that judge's opinions should

i. restate facts and claims of the case.

ii. list evidence admitted and viewed.

iii. state, per item of evidence, whether it is beneficial to the party that presented it.

iv. formally determine the legal facts of the case (i.e, state whether the "labor relationship existed", "worker resigned", etc.)

v. issue a verdict.

vi. quantify an award, if any.

The recorded and observed events are coded by encoding the binary answers (or lack thereof) to a series of 75 questions, reflecting the most common evidence types presented in such cases. Given the availability of the full case file, we can do this both for the evidence admitted by the court and for the decision written by the clerk/judge. For the purposes of showing preliminary results from coded casefiles, in this paper we will often focus on a collection of legal cases in which the main legal issue is a factual determination of existence of a contractual labor relationship between the parties. This allows us to concentrate on a homogeneous collection of around a dozen evidence bits that are common in making such a determination. In general, however, our data set contains many more binary evidence and outcome variables that could be used for similar analysis.

Note that each column, under perfect information, would contain a 0 or 1, indicating that the answer is NO or YES, to the question posed by the column. In reality, however, there may be no mention of certain evidence, or no question of a particular type put to a party or witness, so that the column would contain an "x". The database, then, contains long strings of "0"s, "1"s, and "x"s, while the true state of the world is a string of only "0"s and "1"s.

Our initial effort concentrated on 1200 cases from 6 clerks/judges. At this point we have succeeded in coding 250 of these, with 2 of the 6 clerks/judges completed, allowing for preliminary work on this data. Additionally, we have now obtained close to 7000 scanned case files from a different and larger labor court, with correspondingly larger legal workforce. While we are continuing to compile and code the data on actual court cases, we have also used implemented a small-scale lab-in-field experiment with 6 clerks, of which 5 are the same as those for whom we collected real data. This consisted of a questionnaire with

47 simplified case scenarios based on typical fact patterns, which the clerks had to answer individually (without comparing notes). These 47 cases covered 4 distinct case types, each involving a different "legal controversy" (such as a claim for indemnity versus the defense that the worker resigned voluntarily from the job). The subjects were asked to make legal factual determinations and provide a corresponding verdict for each of the cases. Incentives were provided based on an evaluation of their responses by a judge in a different court within the labor court system.

In order to analyze the data thus produced we propose a simple model that we present in the next section.

3 The Model

A judge has to decide on the merits of the case, issuing the verdict $v \in \{1, -1\}$, where 1 stands for $guilty$ and -1 for $innocent$. We shall propose that the judge uses a two-stage procedure for coming up with the verdict. First, she establishes a number of *legal facts*, based on evidence presented to and accepted by the court (for the moment, we shall assume that observing this evidence is costless and the judge does this perfectly). In our environment a typical legal fact could be whether there existed an actual labor relationship between the alleged employer and the employee or whether the employee was actually fired. Once this is done, she utilizes a legal rule to come up with a verdict either in favor of the plaintiff or the defendant.

In that second stage, when deciding on the verdict, she faces a universe of possible cases F, which, following Kornhauser (1992) we shall call the case space. Each state $f \in F$ may be thought of as a length $n \in \mathbb{N}$ string of zeros and ones: $f = (f_1, f_2, ..., f_n)\,, f_i \in \{-1, 1\}$. Each element f_i of the string, shall be called a *legal fact*, the case space is $F = \{-1, 1\}^n$ (its cardinality is $N = 2^n < \infty$). The case space F is partitioned into two disjoint payoff-relevant events: guilty and innocent, $G \cup I = F, G \cap I = \varnothing$. This parititioning of the state space may be interpreted as a *legal rule*. We shall assume throughout that the legal rule is fixed. Once the set of legal facts and hence the true case $f^* \in F$ is determined (that is, the judge establishes the value of f_i for each $i = 1, 2, ...n$) she will apply the legal rule, to issue the verdict $v = 1$ if $f^* \in G$ and $v = -1$ otherwise.

Prior to this, of course, the judge has to make a binary determination for each of the legal facts. We shall assume that she employs a similar procedure, by aggregating a number of evidentiary *bits* into a legal factual determination. Specifically, we shall assume that for each legal fact i the judge considers a space of possible true states of the world Ω^i, in which each state $\omega^i \in \Omega^i$ is described by a string (of length n_i) of possible answers to a series of true/false questions, such as "has a formal contract been presented to the court?" or "did the employer allege the employee resigned voluntarily?" Therefore, $\omega^i = \left(\omega_1^i, \omega_2^i, ..., \omega_{n_i}^i\right), f_i \in \{-1, 1\}$. Each element ω_j^i of the string, shall be called an *evidentiary bit* and the corresponding case space is the case space is $\Omega^i = \{-1, 1\}^{n_i}$. As before, we shall assume that the judge has in her mind a partition $G^i \cup I^i = \Omega^i, G^i \cap I^i = \varnothing$, constituting a legal rule.

However, unlike in the second stage, by which all legal facts have been established and no uncertainty exists as to where the case lies with respect to the partition, in the first stage the judge might be forced to make her decisions based on incomplete evidence, without observing the true state $\omega^{i*} \in \Omega^i$. Correspondingly, we may define the *evidence record* to be a string, $e^i = \left\{e_1^i, e_2^i ... e_n^i\right\}$, $e_j^i \in \{0, 1, x\}$, where $e_j^i = x$ denotes a failure of the record to reflect potential evidence. We shall assume that every $e_j^i \in \{\omega_j^{i*}, x\}$ - that is, that the record always allows for the truth.. Any such string e^i corresponds to an event $E^i \in 2^{\Omega^i}$. The true state is denoted $\omega^{i*} \in \Omega^i$ and may belong either to G^i or I^i.

The two stages we define are, of course, quite similar. We could have, in fact, merged them into a single act of judicial decision-making. Our decision to separate the two is based on the nature of our data: we actually observe both the admitted evidence and the legal fact determinations that, by law (in this case, it is article 840 of the Mexican *Federal Labor Law*) the judges must make in their written decisions. We thus have an opportunity to establish both the legal rules, implicit in the observed mapping between established legal facts and verdicts, and the rules the judges use to establish the legal facts themselves based on their observation of the empirical facts of the case as described by the evidence admitted by the court. The structure of the two problems, however, is similar enough that it will be convenient to treat the two together. We shall return to treating them as distinct when analyzing the data. For the rest of this section we shall concentrate on a single problem: that of a Bayesian judge aggregating observed data into a single binary verdict.

3.1 The judge

We shall now consider a simple model of a single-stage decision by the judge, who has to issue a single binary verdict based on evidence.. It shall be convenient, for the moment, to abstract from the structure of the evidence she faces. It is sufficient for our first result that she knows that the true state ω^* is an element of some finite measurable set $(\Omega, 2^\Omega)$ We shall assume that she has a prior belief β, which is a probability measure over the finite measurable space $(\Omega, 2^\Omega)$. The set of all possible beliefs shall be denoted as $-(\Omega) = \Delta^{2^n}$ As the evidence is presented, she is observing an event $\omega^* \in E \subset \Omega$ and updates (in a Bayesian way) the probability of guilt.

As before, our judge, can take one of two actions $v \in \{-1, 1\}$ the former standing for innocence and the latter for guilt. We shall assume her *ex post* utility of each action depends only on whether the true state ω^* belonging either to the event G or the event I into which the state space is partitioned by the legal rule. Let $U_T(v)$ denote the utility of action v if the truth lies in $T \in \{G, I\}$. In everything that follows we shall assume that $U_G(1) = U_I(-1) = 0, U_I(1) = -q$ and $U_G(-1) = -(1-q)$ for some $q \in (0,1)$, which may be naturally interpreted as aversion to convicting the innocent (see for instance, Coughlan 2000). Clearly, with this functional form, the judge would want to convict whenever the $\tau(G) > q$ and acquit whenever $\tau(G) < q$ (for simplicity we shall ignore the possibility of indifference). We may interpret q as the judge's *conviction bias.*

The true state of the world, $\omega^* \in \Omega$ may or may not be fully observable by the court. Indeed, at a trial only partial evidence may be presented and or allowed, producing a recorded event $\omega^* \in E \subset \Omega$. Typically, only the evidence admitted by the court may be considered in issuing the verdict.Hence, assuming the judicial decision is based on the evidence record (ignoring the possibility of indifference), $\tau(G|E) = \frac{\beta(E \cap G)}{\beta(E)} > q$ implies a verdict of g and a verdict of i follows from $\frac{\beta(E \cap G)}{\beta(E)} < q$.

Hence, the *case record* can be denoted as a pair $r = (E, v)$, consisting of the evidence admitted by court E and and the verdict v The set $\mathcal{C}$ of all observed case records r produced by a given judge we shall call his *legal output.* Notably, legal output $\mathcal{C} = \{r_1, r_2, ..., r_M\}$ is observable and can be used as data in analyzing the behavior of the judge. In particular, it can be used to reveal the parameters determining individual choice, such as G, β and q, which may be subjective and not directly observed by the researcher.

3.2 Rationalizing the legal output

The first question that we shall ask is whether the *legal output* of a judge is consistent with her basing decisions only on evidence formally admitted by the court. It turns out that even this very stylized model, which remains agnostic on the causes of incompleteness in the court record, imposes a testable restriction on what $\mathcal{C}$ may contain, even if we do not know the judge's prior β, her conviction bias q or her guilty set G. In fact, assuming the data satisfy this restriction, we may use it to "reveal" the unobserved parameters of the model.

We shall say that a triple (β, q, G) *rationalizes the legal output* if for any $r = (E, v) \in \mathcal{C}$ such that $f = g$ we have $\tau(G|E) = \frac{\beta(E\cap G)}{\beta(E)} > q$ and for any $r' = (E', v') \in \mathcal{C}$ such that $v' = i$ we have $q > \tau\frac{\beta\left(E'\cap G\right)}{\beta(E')} = (G|E')$.

We shall denote the set of all recorded events corresponding to the legal output of a judge that result in conviction as $A(\mathcal{C})$ and the set of all recorded events resulting in acquittal as $B(\mathcal{C})$. It is straightforward to see that the two sets $A(\mathcal{C}), B(\mathcal{C}) \in 2^{\Omega}$ are disjoint. Furthermore, it is easy to show that a union of any two disjoint events from $A(\mathcal{C})$ (or, respectively,$B(\rfloor)$), if the corresponding verdict is ever observed, must be in the same set. Indeed, for any $E_1, E_2 \in A(\mathcal{C})$ such that $E_1 \cap E_2 = \varnothing$, $\tau(G|E_1 \cup E_2) = \frac{\beta((E_1\cup E_2)\cap G)}{\beta(E_1\cup E_2)} = \frac{\beta(E_1\cap G)+\beta(E_2\cap G)}{\beta(E_1)+\beta(E_2)} = \frac{\frac{\beta(E_1\cap G)}{\beta(E_1)}\beta(E_1)+\frac{\beta(E_2\cap G)}{\beta(E_2)}\beta(E_2)}{\beta(E_1)+\beta(E_2)} > q\frac{\beta(E_1)+\beta(E_2)}{\beta(E_1)+\beta(E_2)} = q$ (and, correspondingly, for any $E_1, E_2 \in B(\mathcal{C})$ such that $E_1 \cap E_2 = \varnothing, \tau(G|E_1 \cup E_2) < q$). This, as can be easily seen from the examples below, implies a clear testable restriction on $\mathcal{C}$.

A1 (additivity) For any $E_1, E_2 \in A(\mathcal{C})$ such that $E_1 \cap E_2 = \varnothing$, it follows that $E_1 \cup E_2 \notin B(\mathcal{C})$ and for any $E_1, E_2 \in B(\mathcal{C})$ such that $E_1 \cap E_2 = \varnothing$, it follows that $E_1 \cup E_2 \notin A(\mathcal{C})$

Notably, as long as additivity is satisfied we have the following natural corollary: if, whatever the realization of a certain bit of evidence, given the other bits observed the verdict is unchanged, then not observing the same bit shall not affect the verdict either (this, in fact, could be considered as a consequence of the standard Blackwell (1953) information ranking). This can be easily illustrated in the following pair of simple examples.

Example 1 *Let $N = 4$ and suppose we observe that $A(\mathcal{C}) = \{\{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}\}$, $B(\mathcal{C}) = \{\{\omega_1, \omega_3\}, \{\omega_2, \omega_4\}\}$. Clearly this is impossible, as it would necessitate that, no matter what G, q and β, the verdict corresponding to the uninformative $E = \Omega$, if it ever were to be observed, would have to simultaneously be in $A(\mathcal{C})$ and $B(\mathcal{C})$.*

Example 2 *Let $N = 4$, $A(\mathcal{C}) = \{\{\omega_1, \omega_2\}, \{\omega_1, \omega_3\}, \{\omega_4\}\}$ and $B(\mathcal{C}) = \{\Omega, \{\omega_1\}\}$ are inconsistent following repeated application of the same "Blackwellian" argument. Indeed, if ω_2 were to result in the verdict of innocent, since $\omega_1 \in B(\mathcal{C})$ it would imply, by the argument of the previous example, that, contrary to the record $\{\omega_1, \omega_2\} \in B(\mathcal{C})$. Hence, we know that ω_2 can only correspond only to the verdict of guilt. Analogously, ω_3 may only be consistent with guilt. However, that, together with the fact that $\{\omega_4\} \in A(\mathcal{C})$ implies that $\{\omega_2, \omega_4\}$ and $\{\omega_3, \omega_4\}$ can only correspond to the verdict of guilt, which, noting that $\{\omega_1, \omega_3\} \in A(\mathcal{C})$, would require a verdict of guilt based on the uninformative event Ω - contradiction.*

Unfortunately, additivity is insufficient for rationalizability of a legal output, as can be seen from the following example

Example 3 *Let* $N = 3$, $A(\mathcal{C}) = \{\{\omega_1, \omega_2\}, \{\omega_2, \omega_3\}, \{\omega_1, \omega_3\}\}$ *and* $B(\mathcal{C}) = \{\Omega\}$ *As there are no two observed cases with empty intersection, additivity has not bite here. However, whatever the prior* β *the judge may have and whatever the legal rule* (G, I) *and evidence standard* q *she may use these decisions may not be rationalized. Indeed, for each state* ω_i *we may define* $\alpha_i = \beta(\omega_i)$ *if* $\omega_i \in G$ *and* $\alpha_i = 0$ *otherwise. Clearly, the three observed guilty verdicts imply that* $\alpha_1 + \alpha_2 > q(\beta(\omega_1) + \beta(\omega_2))$, $\alpha_2 + \alpha_3 > q(\beta(\omega_2) + \beta(\omega_3))$, $\alpha_1 + \alpha_3 > q(\beta(\omega_1) + \beta(\omega_3))$ *so that, summing them up we obtain* $2(\alpha_1 + \alpha_2 + \alpha_3) > 2q(\beta(\omega_1) + \beta(\omega_2) + \beta(\omega_3)) = 2q$. *But the innocent verdict imples that* $\alpha_1 + \alpha_2 + \alpha_3 < q$ *- contradiction.*

The exact strengthening of additivity that is necessary if our attention model is to be consistent with the observed data turns out to be well-known. Originally introduced in Kraft *et al.* (1959), this condition may, following Fishburn (1970), be presented as follows. Consider two collections of recorded events $\mathbf{A} = (A_1, A_2, ..., A_{m_1})$ and $\mathbf{B} = (B_1, B_2, ..., B_{m_2})$ (repetitions of events in a collection allowed). Denote as $n_\omega(\mathbf{E})$ the number of events in the collection $\mathbf{E}$ that state ω is included in. We say that the two collections are equivalent $\mathbf{A} \approxeq \mathbf{B}$ if for each event $\omega \in \Omega$ $n_\omega(\mathbf{A}) = n_\omega(\mathbf{B})$.

A2 *(strong additivity)*: For any $\mathcal{C}$ there does not exist equivalent collections $\mathbf{A}$ consisting of elements of $A(\mathcal{C})$ (possibly repeated) and $\mathbf{B}$ of elements of $B(\mathcal{C})$ (also possibly repeated).

The assumption A2 cannot be violated with a single observation of a decision by a judge, and becomes increasingly harder to satisfy, as the number of observations grows. As the following theorem shows, the

Proposition 4 *There exist the partition of the state space into* G *and* I, *the prior* β *and the conviction threshold* q *rationalizing the legal output* $\mathcal{C}$ *if and only if A2 holds. Furthermore, without loss of generality, the conviction bias may be normalized to* $q = \frac{1}{2}$.
Proof. *See appendix* ■

Notably, as part of the proof of proposition 1 we recover the parameters of the model. In fact, the partition of the state space Ω into G and I would be unique if the judge's legal output were sufficiently complete: this is trivial, assuming the evidence were fully recorded (i.e., $E = \{\omega^*\}$ in each case) and a decision corresponding to every possible $\omega \in \Omega$ were on record. In contrast, the prior distribution β will, in general, be non-unique (we cannot hope to get uniqueness in general, unless we get the infinite divisibility of the state space). These parameters, in fact, correspond to a solution of a system of linear inequalities

$$Ax > 0$$

with the matrix A given by the legal output $\mathcal{C}$. Each row $j = 1, 2, ..., M$ of the matrix would correspond to a case record $r_j = (E_j, v_j)$, while each column $i = 1, 2, ...N$, would correspond to a state ω_i. The corresponding element of the matrix is $a_{ji} = v_j 1_{E_j}(\omega_1)$. For each solution $x = (x_1, x_2, ...x_N)$ to such a system such that $\sum_{i=1}^{N} |x_i| = 1$ we may interpret the absolute value of the coordinate as the prior probability $\beta(\omega_i) = |x_i"|$, while the sign of the x_i would indicate whether ω_i is in G (if it is positive) or I (if it is negative).

Example 5 *Let $N = 4$, $A(\mathcal{C}) = \{\{\omega_1, \omega_2\}, \{\omega_1, \omega_3\}\}$ and $B(\mathcal{C}) = \{\{\omega_2, \omega_4\}, \{\omega_3, \omega_4\}\}$. The parameters of the model can be discovered from the solution to the following system of inequalities:*

$$\begin{aligned} x_1 + x_2 &> 0 \\ x_1 + x_3 &> 0 \\ -x_3 - x_4 &> 0 \\ -x_2 - x_4 &> 0 \end{aligned}$$

One (normalized) solution to the system could be $x = \left(\frac{1}{3}, \frac{1}{6}, \frac{1}{6}, -\frac{1}{3}\right)$ corresponding to $G = \{\omega_1, \omega_2, \omega_3\}$,$I = \{\omega_4\}$ and the prior distribution $\beta = \left(\frac{1}{3}, \frac{1}{6}, \frac{1}{6}, \frac{1}{3}\right)$.

3.2.1 Minimal rationalizing facts: attention

A natural question to ask here is, what proportion of possible legal records of this sort could be rationalized. Indeed, if legal records of a certain dimension are nearly always rationalizable, our ability to use actual data to test the theory using real data should be considered suspect, as failure to reject would likely reflect the lack of "power " in the test we propose. Indeed, if we consider state spaces that are large enough, we should be always able to rationalized whatever data we consider, since pretty much any pair of cases is likely to be substantially different in many minor aspects of the case. Consequently, the question is not whether the data can be rationalized if we consider a large number of possible case distinctions: it, probably, can. Rather, the question is what is the minimal collection of such distinctions that would make the legal output of a judge rationalizable. We could, in fact, interpret such "coarsest" state space as what the judge had to, at a minimum, pay attention to in order for the data to be consistent with our theory. Posing this question in this way is somewhat complicated, however since it is not, in fact, obvious that rationalizability is monotonic in how detailed is the judge's observation of the case. Indeed, in general, as the following example shows, it is not true.

Example 6 *Consider an example with 4 states,$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega 6\}$, and 4 cases $E_1 = \Omega$, $E_2 = \{\omega_2, \omega_5\}$, $E_3 = \{\omega_1, \omega_4\}$ and $E_4 = \{\omega_1, \omega_6\}$. Suppose when nothing is known, $E_1 = \Omega$ we observe the verdict verdict g, while otherwise is i. Clearly*

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 > 0$$

$$x_2 + x_5 < 0$$

$$x_1 + x_4 < 0$$

$$x_3 + x_6 < 0$$

Which cannot be rationalized. However, suppose the judge cannot distinguish between states ω_1 and ω_2 so that whenever one of the two obtains, she says that either is possible. Effectively, this results in a merger of the two states *into a single state which we will call ω_{12}. We know get $\Omega = \{\omega_{12}, \omega_3, \omega_4, \omega_5, \omega_6\}$, $E_1 = \Omega$, $E_2 = \{\omega_{12}, \omega_5\}$, $E_3 = \{\omega_{12}, \omega_4\}$ and $E_4 = \{\omega_3, \omega_6\}$ implying the system*

$$x_{12} + x_3 + x_4 + x_5 + x_6 > 0$$

$$x_{12} + x_5 < 0$$

$$x_{12} + x_4 < 0$$

$$x_3 + x_6 < 0$$

which is solved, for instance, by $x = \left(-\frac{1}{3}, -\frac{1}{8}, \frac{1}{4}, \frac{1}{4}, \frac{1}{24}\right)$

It turns out however, that our original view of evidence, as a sequence of answers to binary questions, the monotonicity we conjectured is restored. Fortunately, this evidence structure is indeed exactly what we obtain from actual legal data.

Indeed, we shall now recall that we have originally defined $\omega \in \Omega$ as a length-$n, n \in \mathbb{N}$ string of zeros and ones: $\omega = (\omega_1, \omega_2, ..., \omega_n), \omega_i \in \{0, 1\}$. Each element ω_i of the string, shall be called a *bit*, the state space is $\Omega = \{0, 1\}^n$ (its cardinality is $2^n < \infty$). As before the state space is partitioned into two disjoint payoff-relevant events: guilty and innocent, $G \cup I = \Omega, G \cap I = \varnothing$. The true state is denoted $\omega^* = (\omega_1^*, \omega_2^*, ..., \omega_n^*) \in \Omega$. Each bit in our string may be interpreted as referring to a binary determination of the truth of a relevant statement (whether the defendant ate in a McDonald's an hour before the murder or not; whether he was due a certain payment or not, *etc.*). The even G corresponds to the collection of states, which unambiguously (in the eyes of a decision-maker) correspond to defendant's guilt.

Correspondingly, we may define the *evidence record* to be a string, $e = \{e_1, e_2 ... e_n\}$, $e_i \in \{0, 1, x\}$, where $e_i = x$ denotes a failure of the record to reflect potential evidence. We shall assume that every $e_i \in \{\omega_i^*, x\}$ - that is, that the record always allows for the truth. We shall denote as $k(e) = \{i : e_i \neq x\}$ the total number of elements in e that are different from x (i.e., the number of bits contained in the record). For every observed case record we may define the corresponding *recorded event*

$$E_e = \{\omega \in \Omega : e_i \neq x \text{ implies } e_i = \omega_i \text{ for all } i = 1, 2, ..., n\}$$

Likewise , we may define the recorded motivation to be $\varpi = (\varpi_1, \varpi_2, ..., \varpi_n)$, $\varpi_i \in \{0, 1, x\}$, where $\varpi_i = x$ denotes failure to mention a bit of evidence, either because it is unavailable in the original record, or for any other reason. If the judge is completely uninformed, we shall denote her observation as $\varpi = \varpi^0$. Similarly to the *recorded event*, we may define the corresponding *described event*

$$E_{\varpi} = \{\omega \in \Omega : \varpi_i \neq x \text{ implies } \varpi_i = \omega_i \text{ for all } i = 1, 2, ..., n\}$$

In what follows we shall assume that $\varpi_i \in \{e_i, x\}$, that is that the only "mistake" that the judge may make is not to consider available evidence. We shall denote as $k(\varpi) = \{i : \varpi_i \neq x\}$ the number of observed bits

In a legal record with m different cases we shall denote the case number with a superscript, so that $\left(e^j, f^j\right) \in \{0,1,x\}^n \times \{0,1\}$ and a decision is $\left(\varpi^j, f^j\right) \in \{0,1,x\}^n \times \{0,1\}$. Thus, the legal record is $\mathcal{C} = \left\{\left(e^j, f^j\right)\right\}_{j=1,2,\ldots m}$ and the collection of decisions is $\mathcal{D} = \left\{\left(\varpi^j, f^j\right)\right\}_{j=1,2,\ldots m}$.

We can easily translate the examples of the previous section into this language, as we illustrate below.

Example 7 *Let $n = 2$ and suppose we observe that $A(\mathcal{C}) = \{1x, 0x\}$, $B(\mathcal{C}) = \{x1, x0\}$. Clearly this is impossible, as it would necessitate that, no matter what G, q and β, the verdict corresponding to a string xx, if it ever were to be observed, would have to simultaneously be in $A(\mathcal{C})$ and $B(\mathcal{C})$.*

Example 8 *Let $n = 2$, $A(\mathcal{C}) = \{1x, x1, 00\}$ and $B(\mathcal{C}) = \{xx, 11\}$ are inconsistent following repeated application of the same "Blackwellian" argument. Indeed, if 10 were to result in the verdict of innocent, since $11 \in B(\mathcal{C})$ it would imply, by the argument of the previous example, that $1x \in B(\mathcal{C})$. Hence, we know that 10 can only correspond only to the verdict of guilt. Analogously, 01 may only be consistent with guilt. However, that, together with the fact that $00 \in A(\mathcal{C})$ implies that $0x$ and $x0$ can only correspond to the verdict of guilt, which, noting that $1x \in A(\mathcal{C})$, would require a verdict of guilt based on xx - contradiction.*

Clearly, the characterization of the previous section apply here as well. Of course, the number of states here is large: $\#\Omega = 2^n$. Hence, even for a relatively small number of observable bits n we may expect a large probability of data being spuriously rationalizable (due to irrelevant but observable distinctions between cases). Fortunately, however, a judge consistently observing an extra bit of information effectively splits each in two not one, but every state $\omega \in \Omega$. Consequently, it actually turns out that observing finer distinctions can never hurt rationalizability, which allows us to meaningfully ask the question: what is the smallest (by inclusion) collection of evidence such that if we assume the judge is basing her decisions on the bits in this collection, her legal output is rationalizable. Formally, if $K = \{1, 2, ...K\}$ is the set of observable bits, we shall define a non-empty subset $\varnothing \neq S \subset N$ as the *attention set* of the judge, with corresponding $\Omega(S)$ being the judge's *subjective state space*. Clearly, we can use the proposition 1 above to check the rationalizability given $\Omega(S)$. We then obtain the following proposition

Proposition 9 *Consider two possible attention sets such that $\varnothing \neq S \subset T \subset N$ If the data can be rationalized with $\Omega(S)$ then it can be rationalized with $\Omega(T)$*

Proof. *See Appendix* ■

4 Some simulations

As noted above, if data sets are very detailed, we may be running the risk of trivially rationalizing any legal outputs. In order to get a feeling for how much data we may need to test our theory, we ran a series of simulations in which randomly legal outputs of a certain dimension were constructed 1000 times by randomly drawing $1s$, $0s$ and xs in the bit sequences describing individual cases and $1s$ and $-1s$ corresponding to verdicts from a uniform distribution. We present results of this simulation in Table 1, which shows the proportion of legal outputs of various dimension (i.e., varied by the number of cases and factual bits). Though useful, this exercise should, of course, be treated with a caution, as random and real data are likely to be

Table 1: Proportion of Sets Consistent with Rationality: Randomly Generated Data.

	Bits								
Obs.	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
20	21.1%	63.5%	86.8%	96.8%	98.6%	99.6%	99.8%	100%	100%
40	0%	10.4%	55.8%	81.1%	94.2%	98.6%	99.6%	99.5%	100%
60	0%	0.4%	22%	66.6%	88.3%	96.1%	98%	99.9%	99.7%
80	0%	0%	6.5%	42.4%	79.4%	91.5%	97.2%	99.2%	99.9%
100	0%	0%	0.6%	29.7%	69.2%	87.1%	96%	99%	99.6%
120	0%	0%	0.1%	16.3%	55.1%	83.6%	94.4%	98.6%	99.2%
140	0%	0%	0%	6.5%	43.3%	78.6%	91.3%	98%	99.2%

Table 2: Example: Non-rationalizable set with indirect conflict.

Bit 1	**Bit 2**	**Bit 3**	**Bit 4**	**Bit 5**			**Verdict**	
1	X	X	X	1			1	
1	0	X	X	0			1	
1	X	X	X	X			-1	
1	1	X	X	0			1	

very different in that not all binary sequences are equally likely to obtain in the field, if for no other reason than the logical relationship between legal issues and evidence. In fact, in the preliminary results shown below, we will find conflicts far more frequently in both lab in the field experimental data and real data from case files, as compared to what we find in randomly generated sets.

It should be noted that most non-rationalizabilities we encounter in this simulation come from "direct conflicts": identical cases with opposite verdicts. However, we also observe a number of longer non-rationalizable collection of cases. Table 2 presents one such example from our simulation in a 5-bit data set. In this case, all four cases must be used to generate a conflict; removing any one of the rows eliminated the conflict, leaving a set of three rationalizable data points.

5 First Analysis Using Data

Here we report a few preliminary results from two sources of field data. First, a lab in the field experiment we ran with 6 of the judges (judge/clerks who will here just call judges) whose decisions we document from real casefiles, and second, evidence from around 200 casefiles of 2 of the judges whose decisions we observe.

In both analyses, we will separate cases into 4 distinct groups, based on the main legal controversies. On the claim side, the only two possible claims in a firing lawsuit are a claim for indemnity and a claim for reinstatement, which almost always includes indemnity as a subsidiary claim should reinstatement not be achieved. On the defense side, there are 3 relevant answers to firing lawsuit claims: denying the existence of the labor relationship, a counter-claim that the worker resigned voluntarily from the post as opposed to being fired, and a denial of having fired the worker along with an offer to reinstate her in the same position. These 3 defenses cover almost 90% of all firing lawsuits, while firing lawsuit comprise about 95% of all the cases handles by these courts.

When one of the two first defenses is used by the firm, burdens of proof are distributed in the same way whether the worker claims indemnity or reinstatement as his main claim, so we consider a one single "controversy" either indemnity or reinstatement, versus denial of labor relationship, and as another single "controversy" either indemnity or reinstatement, versus voluntary resignation. For the first of these two, the worker has the burden of proving the labor relationship existed, and if this burden is discharged correctly, he should win the case. On the other hand when the firm's defense is voluntary resignation, the labor relationship is assumed to exist and the firm has the burden of proof to show the worker resigned.

The legal controversies that involve an offer of reinstatement made by the firm are more complicated in terms of burden of proof, because jurisprudence states that if the worker refuses an offer of reinstatement made in good faith, then she must carry the burden of proof to show that unfair firing happened, rather than the firm carrying the burden of proof to show it did not. However, while the jurisprudence does not predicate this on the main claim made by the worker, in fact judges seem to apply this jurisprudence more strictly when the worker's main claim is reinstatement. In fact, when a reinstatement offer is made by the firm in its answer to the lawsuit, some judges consider that if the worker and her lawyers do not give an explicit answer to this offer, their "implicit" answer is in the affirmative if and only if the principal claim was reinstatement. For this reason, we separate controversies with an offer of reinstatement by the firm, according to the main claim of the worker.

As we have mentioned, the lab in the field experiment was administered to 6 judges for whom we observe or will observe real casefile data. The questionnaire they answered included 47 hypothetical cases spread over the 4 legal controversies described above. Each hypothetical included a simple fact pattern, claims and counter-claims, and evidence submitted by the parties. The hypotheticals were constructed to represent simple and typical cases, however, with an emphasis on cases relatively close to "lines of disagreement" we anticipated based on our previous field work with judges' decisions in the same court. This, and the fact that we have questionnaire data coded and analyzed for all 6 judges (rather than 2 in the case of the real casefile data), may be why so far we find more direct conflicts in the lab in the field data. Before explaining a sample of the current preliminary results, we would like to point out that we are not, yet, in a position to use the data to test or identify the parameters of our model. The main reason for this that we have only been able to finish coding a small part of the real casefile data that has been collected. We therefore concentrate on demonstrating that our data are sufficiently internally consistent, but also contain enough substantive disagreements between judges to make application of our methodology of potential interest. In part to advance this agenda, while we progress with data coding, we administered the incentivized questionnaire just explained. Besides giving us an easy to collect dataset, the questionnaire minimizes the attentional issues we chose not to address in the present paper. We present only some of the preliminary results here. However, we believe that these are of sufficient interest to warrant continuation of our efforts.

Now to a few of the results, first questionnaire data for 6 judges, then real data for 2 judges. In the questionnaire data and in the second mapping, from legal facts to verdict, we find rationalizability (barring what appear to be 2 mistakes by one judge) for both of the first two "simpler" controversies, involving denial of the labor relationship or counter-claim of resignation.

Mapping 2 in the two more complicated controversies yields several conflicts, and analyzing these leads us to believe that judges are using different legal rules. Figures 1 and 2 shows conflicts for these two controversies. In these figures we graph only each two-way direct conflict as a line, and each shape shown in the figures is one conflict, either of a single question, across the 6 judges, or of multiple questions but which have exactly the same legal facts determined by all judges, but different verdicts across judges. Some of the conflicts shown which are just one case by one judge conflicting with many others, appear to be mistakes. But all the conflicts with more than one judge disagreeing with others, when analyzing them, point to different legal rules, and the differences are along the lines that the law does not determine very clearly. For example, a couple of judges seem to think that if a worker's main claim is reinstatement, and she refuses a reinstatement offer, then, even if she proves that unfair firing occurred, she is should not win on a subsidiary indemnity claim. In another example, the judges split equally on the verdict because they have different opinions on whether their role is to "counteract" or "fix" mistakes made by the trial court, or not. In that hypothetical case, the worker claims indemnity, not reinstatement, and the firm offers reinstatement. The trial court requests an answer from the worker but states that no explicit answer from the worker will imply tacit acceptance of reinstatement. The worker's side provides no explicit answer, a reinstatement is scheduled by the court, and the worker does not show up, All judges state that the trial court's presumption was wrongly managed. But 3 decide to let the worker win, maybe as a result of the trial court's mistake, and 3 take the formalist view and worker loses since she rejected a reinstatement but did not then carry the burden of proof successfully to establish that unfair firing took place.

Figure 1: Conflicts between all judges: mapping legal facts to verdict for Indemnity Claim vs. Reinstatement Offer

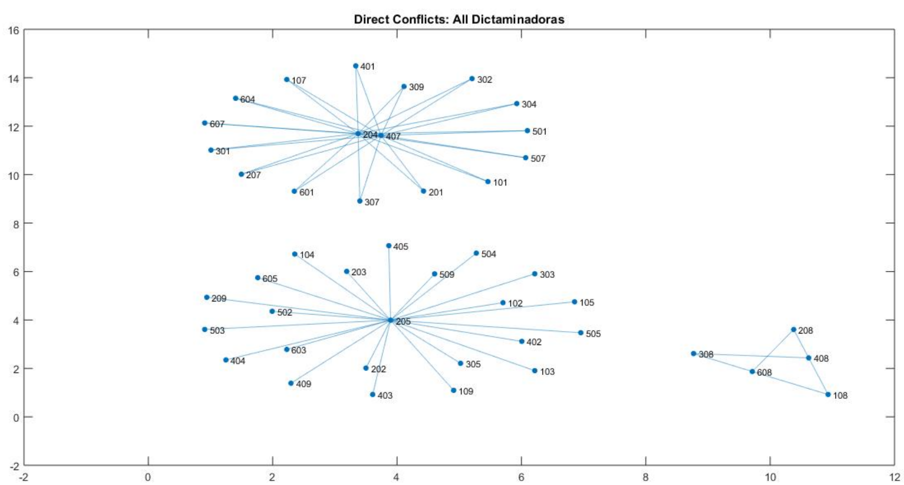
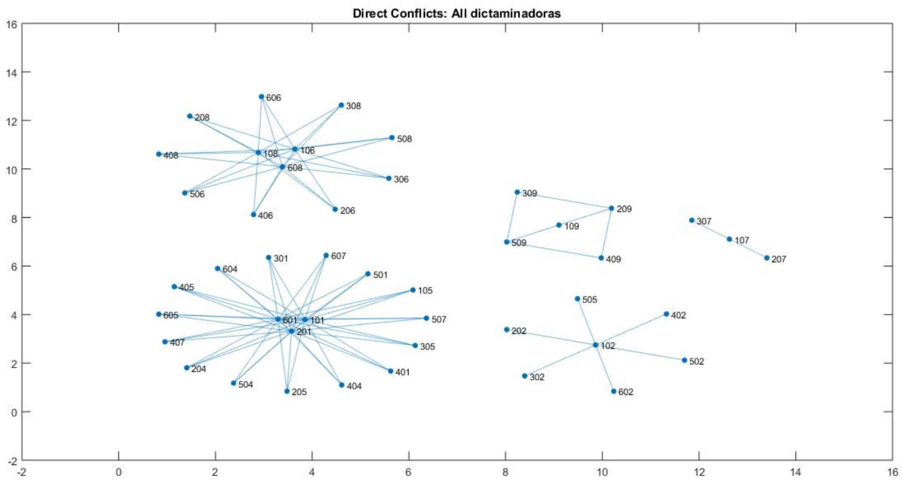


Figure 2: Conflicts between all judges: mapping legal facts to verdict for Reinstatement Claim vs. Reinstatement Offer



On the first mapping, from evidence to legal facts, we only show here the results from conflicts across all 6 judges for the mapping from all the evidence provided to the legal fact "the labor relationship existed".

Figure 3: Conflicts between all judges: mapping evidence to legal fact "labor relationship existed"

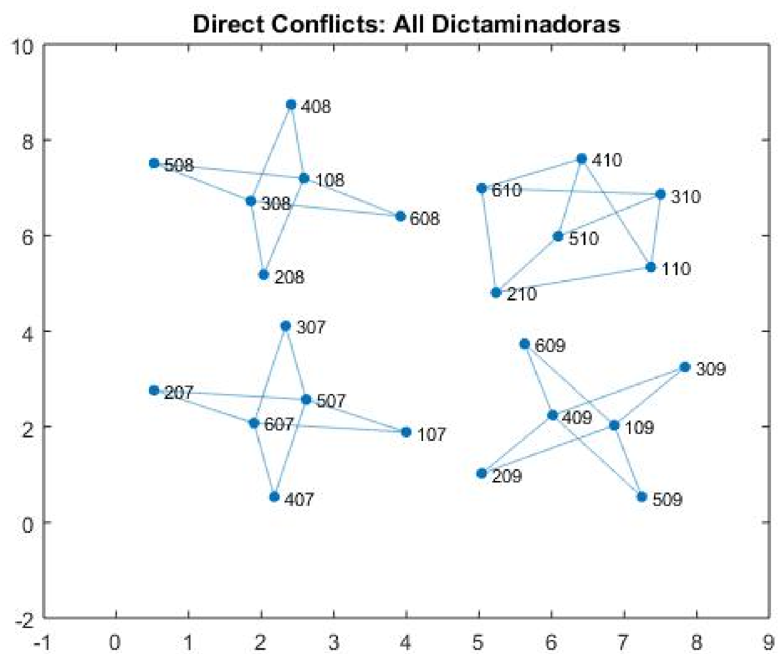


Table 3: Inconsistencies in evidence to legal facts mapping, lab in the field data

	Litis: Indemnity / Denial of labor relationship		
	Scenario	Judge	Relationship exists?
question 7, day 2	Worker offers inspection	1	YES
	of labor contract,	2	YES
	to be produced by firm;	3	YES
	firm does not show up to inspection;	4	YES
	firm provides no evidence	5	NO
	of labor contract,	6	NO
question 10, day 2	Worker offers inspection	1	NO
	of labor contract,	2	NO
	to be produced by firm;	3	YES
	firm shows up ut does not produce contract,	4	NO
	claiming it does not exist;	5	NO
	firm provides no evidence	6	NO

Table 3 shows more detail about the discrepancies in Figure 3. Clearly there is significant variation in what evidence the judges consider sufficient to establish a labor relationship, even in stripped down hypothetical cases. Since casefiles are randomly assigned to these judges when they enter the draft decisions division, parties would randomly be subject to these variations in valuation of evidence.

In data from real casefiles, we have now worked on two judges, with around 200 cases for both, spread over the same 4 controversies we analyze for the lab in the field data. Generally, what we have found so far is rationalizability at the level of mapping 2, and conflicts, within judge, at the level of mapping 1. In other words, in the real data looked at so far, apparently each judge uses a stable legal rule and these are consistent across the two, but there are non-trivial inconsistencies within each of the judges in the mapping of evidence to legal facts. This is our most preliminary data work, but it does indicate that looking seriously at the mapping of evidence to legal facts is promising.

6 Further research and conclusions

6.1 Field Experiment on Attention

While so far we have concentrated primarily on the issue of judicial consistency, the starting point for this agenda was the issue of judge's attention. Early on we were able to implement a field experiment that started as part of an effort to provide case summaries to the clerks in order to reduce error incidence which had been detected in a previous quality control experiment carried out in 2012 - 2013. In the quality control experiment, all evidence from the case file was coded, as well as all evidence mentioned in the draft decision. Discrepancies between the two were written down on an "observations sheet", and for 50% of the casefiles, the observations sheet was given to the clerk responsible for decision, and she was given an opportunity to correct or rewrite her decision.

A fairly high incidence of clear procedural errors was found, including not mentioning items of evi-

Table 4

Year	Type of entry to judgments division of the court				
	Regular process	Granted appeal requiring a new judgment	Granted appeal requiring new procedure and judgment	Other appeal	Cases with no employer answer
2013	1225	85	14	5	
2014	974	112	61	58	162
2015	988	172	78	153	125
2016	731	69	56	108	3

dence that had been viewed by the trial court and quantifying the amount owed to the plaintiff based on erroneous facts. However, we found that only about one-third of the time clerks made changed to their draft decision after receiving an observation sheet with discrepancies. When we interviewed clerks after the experiment had ended, they stated that on many occasions although the observations sheet pointed out real discrepancies, they could not rewrite decisions since this was too time consuming. They suggested instead a mechanism for error prevention, and this gave rise to the idea of providing case summaries that the clerks could read before working on a casefile, and that would provide them with information about what items of evidence had been admitted and viewed at trial, including where to locate these items in the file.

The case summaries experiment was carried out between August 2013 and September 2015. During the first stage, which ended in October 2014, casefiles entering the draft decisions division of the court as part of the regular process (non-appeals) were randomized at 50% between receiving a case summary and not receiving a case summary. In the second stage, covering October 2014 to September 2015, we focused on creating greater variation in the information provided to the clerks, in order to later test hypotheses based on models of inattention. As such, we provided all casefiles with summaries, but varied the number of items of evidence included in the summary, as well as whether page numbers were provided or not. Clerks were notified clearly of which type of summary they were being provided for in each casefile, so that they knew that only up to a certain number of items of evidence were included, and therefore knew to look for additional items not coded in the summary sheet.

Table 4 shows the population of casefiles that entered the draft decisions division of the court, from administrative records kept by the manager of the division. These records register each case file that arrives at the division, and classifies the entry as "regular process", a granted appeal that requires a new draft decision but may or may not require new hearings, other appeals, and cases with no employee answer, which were supposed to be decided by the judges themselves, but of which some proportion were sent to the clerks during 2014 and 2015 due to heavy case loads of the judges.

The experiment worked as follows: only casefiles that entered as part of the regular process (non-appeals) were allowed into the experiment. These were randomized at 50% in the 1st stage and at 25% in the second stage into the control and treatment groups. All casefiles, of control or any treatment group, were detained for a few days while the casefile was coded, and the summary sheets produced, so that

Table 5: Experiment population

(a) Field experiment 1st stage

		Did not receive summary sheet	Received summary sheet
Casefiles with first entry as regular process	905	464	441
After first entry as regular process, subsequent entry as non-regular	215	109	106

(b) Field experiment 2nd stage

		Treatment			
		0	1	2	3
Casefiles with first entry as regular process	819	190	211	202	216
After first entry as regular process, subsequent entry as non-regular	143	29	40	40	34

Notes: Second stage treatment groups correspond to the following:

0=up to 1 of each type of evidence, without page numbers

1=up to 1 of each type of evidence, with page numbers

2=up to 5 of each type of evidence, without page numbers

3=up to 5 of each type of evidence, with page numbers

when the clerk received the casefile, the summary sheet was stapled to the front. From then on, the clerk had control of the casefile and produced the draft decision, later returning the file with its draft decision to the division manager, who would return the file to the judge of the individual labor court. Table 5 shows the population of the experiment in each of the two stages. The first row in each panel shows the number of casefiles whose first entry to the division was a non-appeal, and then for these same case files, the second row shows the numbers that entered as non-appeals for the first time but subsequently entered as an appeal.

Tables 6 and 7 show descriptive statistics of treatment effects, and regressions of treatment effects. Since not enough data has yet been collected from subsequent entries of casefile in the second stage of the experiment, we only conduct treatment effects regressions for the first stage. We find statistically significant effects in the following sense: of those files that entered as non-appeals for the first time, those that received a case summary were strictly less likely to enter on a subsequent date as an appeal.

These treatment effects mean that the quality of the draft decisions, measured by subsequent appeals, increased when the responsible clerk received a case summary before working on the file. What could explain this results? The most natural explanation is that clerks do not pay full attention to every part of the casefile, because attention is costly. In a model of costly attention, information that helps focus the attention of the decision maker on relevant bits of information is likely to raise the quality of her decision.

Table 6: Descriptive statistics of treatment effects

(a) Field experiment 1st stage

	Did not receive summary sheet	Received summary sheet
Second or later entry as regular process	65	79
Second or later entry not as regular process	66	38

(b) Field experiment 2nd stage

	Treatment			
	0	1	2	3
Second or later entry as regular process	25	35	34	32
Second or later entry not as regular process	5	9	11	8

Table 7: Treatment effects regression

	Subsequent entry is an appeal	
	(1)	(2)
Summary provided	-0.047**	-0.051**
	(0.023)	(0.023)
Constant	0.128***	0.525***
	(0.016)	(0.217)
Observations	732	732
Controls	NO	YES
Log-likelihood	-172.63	-165.845
Akaike Inf. Crit.	349.26	367.689

Notes: Controls are month/year of first entry.

6.2 Preliminary conclusions

This paper explores implications of treating the motivations contained in legal decisions as evidence for (in)attention on the part of the decisions' author. We derive the testable implications of a simple inattention model on such data and show how to recover the underlying parameters of the model. We than apply this model to real data from a labor court deciding a dichotomous issue in a relatively simple setting, but in which high caseloads and high levels of appeals may imply rational inattention by the clerks who draft first-instance decisions. Data work is now well designed but very far from finished: we have still to code most of the data from full casefiles, and have still to determine rationalizability or conflicts with and across judges for the full data set.

Our aim is to find whether judges are using the same legal rule or not across each of the possible controversies, and whether they aggregate evidence to legal facts in the same way. When this is not the case, we would like to determine whether the conflicts within and across judges are the result of rational inattention, and would like to verify the impact of the field experiment using case summaries, as an intervention that randomly changes the cost of acquiring information from the case file, for the decision maker responsible for the draft ruling.

7 Appendix A: Proofs of propositions

Proposition 10 *There exist the partition of the state space into G and I, the prior β and the conviction threshold q rationalizing the legal output $\mathcal{C}$ if and only if A2 holds. Furthermore, without loss of generality, the conviction bias may be normalized to $q = \frac{1}{2}$.*
Proof. *See appendix* ■

The necessity of this condition follows from the observation that for any two equivalent collections $\mathbf{A} \approxeq \mathbf{B}$ it must be that, if we define $1_E(\omega) = \begin{cases} 1, \text{if } \omega \in E \\ 0, \text{if } \omega \notin E \end{cases}$ to be the indicator function for $\omega \in E$, then, whatever the G

$$\sum_{i=1}^{m_1} \beta(A_i \cap G) = \sum_{\omega \in \Omega} n_\omega(\mathbf{A}) \beta(\omega) 1_G(\omega) = \sum_{\omega \in \Omega} n_\omega(\mathbf{B}) \beta(\omega) 1_G(\omega) = \sum_{i=1}^{m_2} \beta(A_i \cap G)$$

and

$$\sum_{i=1}^{m_1} \beta(A_i) = \sum_{\omega \in \Omega} n_\omega(\mathbf{A}) \beta(\omega) = \sum_{\omega \in \Omega} n_\omega(\mathbf{B}) \beta(\omega) = \sum_{i=1}^{m_2} \beta(B_i)$$

However, $\sum_{i=1}^{m_1} \beta(A_i \cap G) = \sum_{i=1}^{m_1} \frac{\beta(A_i \cap G)}{\beta(A_i)} \beta(A_i) > q \sum_{i=1}^{m_1} \beta(A_i) = q \sum_{i=1}^{m_2} \beta(B_i) > \sum_{i=1}^{m_2} \frac{\beta(B_i \cap G)}{\beta(B_i)} \beta(B_i) = \sum_{i=1}^{m_2} \beta(B_i \cap G)$, implying a contradiction.

In order to show sufficiency, first note that for an arbitrary recorded events $A \in A(\mathcal{C})$. Note that for

$$\beta(A \cap G) - q\beta(A) > 0$$

while for any $B \in \mathbf{B}(\mathcal{C})$

$$\beta(B \cap G) - q\beta(B) > 0$$

Define $a(\omega) = 1_A(\omega)$ for any $A \in \mathbf{A}(\mathcal{C})$ and $a(\omega) = -1_B(\omega)$ for any $B \in B(\mathcal{C})$. The above inequalities can now be rewritten as

$$\sum_{\omega \in \Omega} (1_G(\omega) - q)\,\beta(\omega)\,a(\omega) > 0$$

Notably, the $a(\omega)$ are observed. Our task is to use the data to identify G, q and β.

Denote cardinality of the legal output $\#\mathcal{C} = M$ and $\#\Omega = N$. Indexing the elements of $\mathcal{C}$ by i and those of Ω by j we can rewrite the conditions the unobserved parameters of the model must satisfy as a system of M inequalities in N variables

$$\sum_{i=1}^{N} x_i a_i^j > 0$$

where $a_i^j = 1_{E_j}(\omega_i)$ if $E^j \in A(\mathcal{C})$ or $-1_{B_j}(\omega_i)$ if $E^j \in B(\mathcal{C})$, while $x_i = (1_G(\omega_i) - q)\,\beta(\omega_i)$.

By the theorem of the alternative (see Theorem 4.2 in Fishburn 1970), and since all the coefficients a_i^j are all rational, it can be shown that this system has a solution $x \in \mathbb{R}^N$ if and only if A2 holds. In fact, suppose no solution exists. Then, there must exist a collection of non-negative numbers $r_{k,}$ $k = 1, 2, ..., M$, not all of them zero, so that for every $j = 1, 2, ..., N$

$${}_{k=1}^{M} r_k a_j^k = 0$$

Which can be rewritten as

$${}_{E^k \in \mathbf{A}(\mathcal{C})} r_k 1_{E^k}(\omega_j) - {}_{E^k \in \mathbf{B}(\mathcal{C})}\, r_k 1_{B^k}(\omega_j) = 0$$

In fact, since all a_j^k are rational by construction, all r_k may be chosen to be integers. Consider now two event collections $\mathbf{A}$ and $\mathbf{B}$ such that each event $E^k \in \mathbf{A}(\mathcal{C})$ is repeated r_k times in $\mathbf{A}$ and each event $E^j \in \mathbf{B}(\mathcal{C})$ is repeated r_j times in $\mathbf{B}$. From the preceding equation it follows that the number of times each ω_j is included in events in each collection is

Hence

$$n_j(\mathbf{A}) =_{E^k \in \mathbf{A}(\mathcal{C})} r_k 1_{E^k}(\omega_j) =_{E^k \in \mathbf{B}(\mathcal{C})} r_k 1_{B^k}(\omega_j) = n_j(\mathbf{B})$$

and, hence $\mathbf{A} \approxeq \mathbf{B}$. But by construction we have $A^k \in \mathbf{A}(\mathcal{C})$ and $B^k \in \mathbf{B}(\mathcal{C})$ for all $k = 1, 2, ...M$. Hence, A2 is violated.

We now know that the A2 is the necessary and sufficient condition for the linear system above to have a solution. Furthermore, obviously zero cannot be such a solution and every positive scalar multiple of a solution will also be a solution.

Since $(1_G(\omega_i) - q)$ is positive for any $\omega_i \in G$ and negative otherwise each solution to the system corresponds nearly uniquely to a putative $G \subset \Omega$ (the only ambiguity is where to assign those ω_i for which the corresponding $x_i = 0$; of course, this would imply that the prior probability $\beta(\omega_i) = 0$; for convenience, for such zero prior probability ω_i we shall always assume $\omega_i \in I$). It remains to reconstruct the q and the β and show that these would generate the observed behavior. To do this, we will sum up the x_i over i to obtain

$$\sum_{i=1}^{N} x_i = \sum_{i=1}^{N} \left(1_G\left(\omega_i\right) - q\right) \beta\left(\omega_i\right) = \sum_{x_i>0} \beta\left(\omega_i\right) - q = \beta\left(G\right) - q$$

Furthermore, since for every $\omega_i \in G$ we have $\beta\left(\omega_i\right) = \frac{x_i}{1-q}$ it follows that

$$\beta\left(G\right) = \frac{1}{1-q} \sum_{x_i>0} x_i$$

Hence, if a q is to be the conviction threshold of the model, it must solve the following quadratic equation, with coefficients determined by a solution of the linear system x:

$$q^2 - \left(1 - \sum_{i=1}^{N} x_i\right) q - \sum_{x_i \leq 0} x_i = 0$$

As $\sum_{x_i \leq 0} x_i \leq 0$ the equation may, in general, not have real roots. However, note that multiplying an arbitrary solution $x \in \mathbb{R}^n \setminus \{\varnothing\}$ by a positive scalar $\alpha > 0$ we obtain another solution to the linear system $x' = \alpha x$. In fact, by setting $\alpha = \frac{1}{2\sum_{i=1}^{N}|x_i|}$ we may guarantee that one of the roots $q = \frac{1}{2}$! Hence, we may always normalize $q = \frac{1}{2}$. In other words, whatever bias the judge might feel in favor or against conviction can always be ascribed to his or her prior probabilities of the states of the world, which in this case are $\beta\left(\omega_i\right) = \frac{1}{2}x_i > 0$ for $x_i > 0$ and $\beta\left(\omega_i\right) = -\frac{1}{2}x_i > 0$ for $x_i \leq 0$.

It remains to show that there will exist a solution of the linear system x such that, in fact, $\sum_{i=1}^{N} \beta\left(\omega_i\right) = 1$. By construction, that implies that

$$\frac{1}{2}\sum_{x_i>0} x_i - \frac{1}{2}\sum_{x_i\leq 0} x_i = \sum_{i=1}^{N} |x_i| = 1$$

which, of course, is implied by the normalization $\alpha = \frac{1}{2\sum_{i=1}^{N}|x_i|}$ above.

We have, thus, used the data recover the set G, and a collection of the prior distributions β consistent with the normalized $q = \frac{1}{2}$. It remains to show, that at least one of those will, in fact, generate the original choice pattern. This, however, is straightforward, since by construction for every event $E \in \mathbf{A}\left(\mathcal{C}\right)$

$$\sum_{\omega\in\Omega} \left(1_G\left(\omega\right) - q\right) \beta\left(\omega\right) 1_E\left(\omega\right) > 0$$

and for every $E \in \mathbf{B}\left(\mathcal{C}\right)$

$$\sum_{\omega\in\Omega} \left(1_G\left(\omega\right) - q\right) \beta\left(\omega\right) 1_E\left(\omega\right) < 0$$

Q.E.D.

Proposition 11 *Consider two possible attention sets such that $\varnothing \neq S \subset T \subset N$ If the data can be rationalized with*

$\Omega(S)$ then it can be rationalized with $\Omega(T)$

Proof. *We shall prove the equivalent statement that if the data cannot be rationalized with $\Omega(T)$ it cannot be rationalized with $\Omega(S)$. Consider the evidence record $\mathcal{C}$ of $M \in \mathbb{N}$ cases which may not be rationalized by observation of a particular collection T of bits It is sufficient to prove that droppipng any one bit from T may not restore rationalizability. Therefore, WLOG we shall assume that $T = \{1, 2...T\}$ and $S = \{1, 2...T-1\}$. From the proof of proposition 1 we know that for the $M \times 2^T$-dimensional matrix $A(T)$ defined by the observed legal output $\mathcal{C}$ for the state space $\Omega(T)$ there does not exist a solution to the system of inequalities $A(T)x > 0$. Hence, by the theorem of the atlernative, there exists a vector $r_T \in \mathbb{R}^M_+ \setminus \{0\}$ which solves $r'_T A(T) = 0$. We shall prove that there likewise exists a vector $r_S \in \mathbb{R}^M_+ \setminus \{0\}$ which solves the equation $r'_S A(S) = 0$ for the $M \times 2^{T-1}$-dimensional matrix $A(S)$ defined by the same evidence for $\Omega(S)$ Consider all pairs of 2 states that are indistinguishable on S but distinguishable on T: that is, they are described by the bit strings that are identical except in the Tth bit. Without loss of generality, if the first element of such a pair is the ith state ω_{2i-1} let the second element be ω_{2i}; we have 2^{T-1} such pairs Since $r'_T A(T) = 0$ we know that $\sum_{j=1}^M r_{Tj} 1_{E_j}(\omega_{2i-1}) v_j = \sum_{j=1}^M r_{Tj} 1_{E_j}(\omega_{2i}) v_j = 0$. In the reduced state space $\Omega(S)$ both of these sates will correspond to the single state $\omega'_i \in \Omega(S)$. Clearly, $1_{E_j}(\omega'_i) = \max\{1_{E_j}(\omega_{2i-1}), 1_{E_j}(\omega_{2i})\}$ and the verdict v_j is unaffected by the reduction of the state space. Note further that for a given case j the realization of the Tth bit we are eliminating is either the same for ω_{2i-1} and ω_{2i-1} or it is different. In the former case $1_{E_j}(\omega_{2i-1}) = 1_{E_j}(\omega_{2i})$ and their sum is equal to either 2 or 0 and in the latter case $1_{E_j}(\omega_{2i-1}) \neq 1_{E_j}(\omega_{2i})$ and their sum is equal to 1. Consequently if we take the columns of the matrix $A(T)$ corresponding to ω_{2i-1} and ω_{2i} and add them up, we will obtain a matrix B in which for each row j the corresponding matrix element can take only one of 2 values, $b_{jk} \in \{0, 1\}, \{0, -1\}$ (cases 1 and 2) or $b_{jk} \in \{0, 2\}$, $\{0, -2\}$ (cases 3 and 4). Furthermore, the corresponding element of $A(S)$, $a_{jk}(S) = b_{jk}$ for the first 2 of these cases and $a_{jk}(S) = \frac{1}{2} b_{jk}$ for the last 2 cases. By construction $r'_T B = 0$. But this implies that by setting $r_{Sj} = r_{Tj}$ for cases 1 and 2 and $r_{Sj} = \frac{1}{2} r_{Tj}$ we shall construct the vector $r_S \in \mathbb{R}^M_+ \setminus \{0\}$ which solves the equation $r'_S A(S) = 0$. Q.E.D.* ■

8 Appendix B: Sample Hypothetical Case for Lab Treatment

Valentina Ramirez, aged 45, claimes 9 years of tenure at a daily wage of 350.00 pesos at a manufacturing firm Manufacturera Oliver S.A.de C.V. She alleges unfair dismissal and claims severance pay consisting of 90 days standard legal severance, as well as overtime of 10 hours per week over her entire tenure. Employer answers lawsuit alleging that it never hired the worker. Evidence submitted is the following:

Worker:

1. Provides photocopy of labor contract. No expert testimony (cotejo) to verify that the copy was made from a true original takes place.

2. Offers inspection of the firm's bank records to prove regular salary deposits made to the worker's bank account.

Firm:

1. Provides original of its employment lists, to show worker was never included in these lists during the alleged period of employment.

2. Does not attend the inspection hearing in which it was summoned to provide its official bank records.

3. Provides testimony by two current workers with tenure covering the alleged employment period of the worker, who attend the testimony hearing and answer the question: “Did the firm Manufacturera Oliver S.A.de C.V. ever hire Valentina Ramirez?” in the negative.

References

[1] Alarie, Benjamin RD, and Andrew Green. "The Reasonable Justice: An Empirical Analysis of Frank Iacobucci's Career on the Supreme Court of Canada." U. Toronto LJ 57 (2007): 195.

[2] Blackwell, D. (1953), "Equivalent comparisons of experiments", *Annals of Mathematical Statistics* **24**: 265-272.

[3] Block, Henry David, and Jacob Marschak. Random orderings and stochastic theories of response. No. 66. Cowles Foundation for Research in Economics, Yale University, 1959.

[4] Cameron, C. and Kornhauser, L. (2017), "What do the Courts Do... And how to Model It ", NYU Law and Economics Research Paper

[5] Caplin, Andrew, and Daniel Martin. "A testable theory of imperfect perception." The Economic Journal 125.582 (2015): 184-202.

[6] Caplin, Andrew, and Mark Dean. "Revealed preference, rational inattention, and costly information acquisition." American Economic Review 105.7 (2015): 2183-2203.

[7] Coughlan, P. J. (2000). "In defense of unanimous jury verdicts: Mistrials, communication, and strategic voting," *American Political Science Review*, **94**: 375-393

[8] Fishburn, P. (1970), *Utility Theory for Decision-Making*, John Wiley and Sons : 1970

[9] Fischman, Joshua B. "Measuring inconsistency, indeterminacy, and error in adjudication." American law and economics review 16.1 (2013): 40-85.

[10] Kastellec, Jonathan P. "The statistical analysis of judicial decisions and legal rules with classification trees." Journal of Empirical Legal Studies 7.2 (2010): 202-230.

[11] Kornhauser, Lewis A. "Modeling collegial courts. II. Legal doctrine." JL Econ. & Org. 8 (1992): 441.

[12] Kort, Fred. "Predicting Supreme Court decisions mathematically: A quantitative analysis of the "right to counsel" cases." American Political Science Review 51.1 (1957): 1-12.

[13] Kort, Fred. "Simultaneous equations and boolean algebra in the analysis of judicial decisions." Law and Contemporary Problems 28.1 (1963): 143-163.

[14] Kraft, C., J. Pratt. and A. Seidenberg (1959), Intuitive Probability on Finite Sets, *Annals of Statistics* **30**: 408-419

[15] Lawlor, Reed C. "What computers can do: Analysis and prediction of judicial decisions." American Bar Association Journal (1963): 337-344.

[16] Luce, R. Duncan. "A probabilistic theory of utility." Econometrica: Journal of the Econometric Society (1958): 193-224.

[17] Matějka, Filip, and Alisdair McKay. "Rational inattention to discrete choices: A new foundation for the multinomial logit model." American Economic Review 105.1 (2015): 272-98.

[18] Niblett, Anthony. "Tracking inconsistent judicial behavior." International Review of Law and Economics 34 (2013): 9-20.

[19] Ramji-Nogales, Jaya, Andrew I. Schoenholtz, and Philip G. Schrag. "Refugee roulette: Disparities in asylum adjudication." Stan. L. Rev. 60 (2007): 295.

[20] Segal, Jeffrey A. "Predicting Supreme Court cases probabilistically: The search and seizure cases, 1962-1981." American Political Science Review 78.4 (1984): 891-900.