Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

Alexander Lundberg	Date

The Economics of Judge and Jury Decision Making

By

Alexander Lundberg
Doctor of Philosophy

Economics

David Jacho-Chávez, Ph.D
Advisor

Hugo Mialon, Ph.D
Advisor

Kim Huynh, Ph.D
Committee Member

Ian McCarthy, Ph.D
Committee Member

Paul Rubin, Ph.D
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D
Dean of the James T. Laney School of Graduate Studies

Date

The Economics of Judge and Jury Decision Making

By

Alexander Lundberg

Advisor:
David Jacho-Chávez, Ph.D
Hugo Mialon, Ph.D

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Economics
2017

Abstract

The Economics of Judge and Jury Decision Making

By Alexander Lundberg

The courts have a powerful incentive to reach a correct decision in criminal trials because the social costs of both convicting the innocent and letting free the guilty are high. This dissertation derives two hypotheses of judge and jury decision making from economic theory. Both are tested empirically using federal trial data. Results are consistent with judges applying a higher burden of proof to more serious crimes. The other hypothesis, that judges will apply lower burdens of proof and lower sentences during periods of greater discretion, is partially supported by the data. Lastly, a final chapter examines a theory of group (or jury) interaction in decision making.

The Economics of Judge and Jury Decision Making

By

Alexander Lundberg

Advisor:
David Jacho-Chávez, Ph.D
Hugo Mialon, Ph.D

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Economics
2017

Acknowledgement

Many people have been essential to my development as an economic researcher. With the understanding that every list of acknowledgements is incomplete, I would like to thank several people who stood out over the course of my dissertation:

- My advisor, David Jacho-Chávez, who rendered the phrase "above and beyond" obsolete on a regular basis.

- My advisor, Hugo Mialon, who provided invaluable advice on research in law and economics.

- My committee members, Kim Huynh, Ian McCarthy, and Paul Rubin, for their many helpful comments on the direction of my research.

- Sara Markowitz, director of graduate studies, de facto committee member, and frequent job market advisor.

- Stephanie Gray, tireless graduate program coordinator, who made the logistics of my graduate studies an afterthought.

- Luyao Zou for creating the LaTeX template for this document.

A special thanks to my wife, Lindsay Allen, whose unflagging support has been a continual source of inspiration for me.

Table of Contents

List of Figures

List of Tables

Chapter 1   Introduction

Every society must contend with the economics of crime. The benefits of mutual interaction are constrained by rules of fair play, and individuals who break those rules incur a negative externality on their fellow citizens. Some rules of play are so fundamental as to be codified in criminal law by the government, but law enforcement is an expensive social agenda because crime is a temptation for many individuals.

Although law enforcement is a complicated task involving many participants, the courts are the heart of the justice system. As informed by the legislature, they set the reference point for criminal sanctions, and they choose when to apply them in a world of uncertainty. Knowledge of guilt is rarely if ever absolute, which leaves the courts with an economic decision to make. Should they acquit a defendant and risk setting free the guilty, or should they convict a defendant and risk imprisoning the innocent? According to the US legal system, the latter risk outweighs the former. The presumption of innocence is given procedural deference. While that deference has obvious appeal, the optimal strategy for risk management is not so obvious when the cost of a false acquittal is high. A fully economic adjudicator would, case by case, balance the risk of a false acquittal against the risk of a wrongful conviction, but such a policy would be forbidden by US law.

To support a conviction in a criminal trial, the government must prove the defendant to be guilty beyond reasonable doubt (*In re Winship*, 397 U.S. 358 [1970]). The same burden of proof applies for any crime, whether petty or serious. Economic models, however, reject a uniform standard—crimes with different social cost should carry different burdens of proof (Andreoni 1991; Davis 1994; Kaplow 2011, 2012). The heart of the economic model is the tradeoff between the risk of a wrongful conviction and false acquittal. The optimal burden of proof is a statistical rule minimizing those

expected costs. Given the tension between constitutional and economic prescriptions for criminal jurisprudence, an immediate question is whether or not triers-of-fact do vary their burden of proof.

Although a trier-of-fact might vary his burden of proof for any number of reasons, including whim and prejudice, the focus here is on two economic arguments. First, in standard economic models of reasonable doubt, a more serious crime raises the cost of both a wrongful conviction and a false acquittal. The former increases the optimal burden of proof while the latter reduces it, posing something of a "conviction paradox" (de Keijser and van Koppen 2007). However, the model here shows the burden of proof should rise with the gravity of the crime unless agents are fixated on punishing the guilty. Second, if the same party is entrusted with both the verdict and the sentence, the party might be tempted to "compromise" on the verdict by convicting but imposing a lenient sentence when the facts do not fully establish guilt. In a version of the current model, Lundberg (2016) shows an economic agent should respond to sentencing discretion by engaging in what looks to be a compromise. The best strategy is to reduce the potential cost of a wrongful conviction by choosing a punishment lighter than what fits the crime; the burden of proof then drops in response to the lower cost of error.

The next chapter of this dissertation describes the theoretical model, while the following tests the two hypotheses with data on federal bench trials over the last decade. Although the Sixth Amendment guarantees the right to a trial by jury, criminal defendants routinely waive that right in favor of a bench trial, in which a single judge decides both the verdict and, conditional on a conviction, the sentence to be imposed. Bench trials are a large and understudied part of the justice system. Across 23 states in 2014, just under 44% of criminal trials were conducted before the bench (LaFountain et al. 2016). In the subset of felony trials, that figure was 33%. Judicial verdicts are common at the federal level too. Across district courts in 2010,

14% of criminal trials were held before the bench, a figure below the historical average of roughly 30% (Bureau of Justice Statistics 2010).

Their simple volume is not the only reason to study bench trials. Judges are some of the most experienced practitioners of the law, and they appear to reason economically more often than lay jurors (Viscusi 2001). Therefore, any economically motivated variation in burdens of proof should manifest in judges before juries. Another reason to focus on the bench is the potential for compromise verdicts. Critics normally cite the risk of compromise in regard to jury sentencing (Lillquist 2004), but a bench trial creates the same incentive structure for a judge. Lastly, federal trials are attractive because they offer a rich data set. Information on defendant characteristics, trial outcomes, and sentencing outcomes can be linked across multiple government agencies.

Although a personal burden of proof is not directly measurable, it may be possible to infer changes in burdens of proof from changes in conviction decisions. The first hypothesis, that judges will apply a higher standard in more serious cases, is given support by the data. After controlling for a number of judicial and defendant characteristics, judges are less likely to convict when the charge is serious (as determined by the possible sentence). Of course, evidentiary strength might be correlated with criminal severity. While the results cannot rule out a story whereby changes in conviction probabilities are driven entirely by changes in evidentiary strength, the lack of a similar result for jury trials would seem to imply otherwise. Judges are familiar with sentencing. Lay jurors are usually not, and they are instructed to ignore sentencing in their verdict. Without a convincing explanation for why the distribution of evidentiary strength should vary across bench and jury trials, the results offer preliminary evidence that judges do apply a heightened burden of proof to more serious crimes.

The finding is not without precedent. In an experimental setting, Simon and

Mahan (1971) find subjects apply a higher standard of proof when the crime is more serious (cf. Koch and Devine 1999, Kerr 1978). Andreoni (1995) also documents a negative correlation between criminal penalties and conviction rates, and Snyder (1990) finds that conviction rates in antitrust trials declined after Congress raised the antitrust penalty from a misdemeanor to a felony.

The second hypothesis, that a trier-of-fact will respond to discretion by compromising on verdicts, is tested through an empirical two-part model. A judge first decides whether to convict and then decides the sentence to impose. If judges are compromising on verdicts, a joint test of whether they both convict more frequently and impose more lenient sentences during periods of greater discretion should be statistically significant. Federal sentencing occurs within a set of guidelines developed by the United States Sentencing Commission. The guidelines are a table containing a fairly narrow range of sentences for a defendant with a given criminal history and offense level. Judges are expected, though not required, to choose a sentence from the relevant table range. Over the last decade, several laws and Supreme Court rulings have effectively varied the sentencing discretion of federal judges, mainly by strengthening and relaxing review standards for appeals on sentences outside of the guidelines calculation. Judges prefer not to have their decisions overturned, and the prospect of appellate review is a meaningful constraint, according to empirical evidence (Freeborn and Hartmann 2010; Fischman and Schanzenbach 2011).[1] Within the two-part model, judges do convict more frequently during periods of greater discretion, but their reduction in sentence length is not statistically significant. Likewise, a joint significance test is not significant, providing limited overall evidence of compromise.

The following chapters go through the model, the empirical results, and an extension to the group decision setting. A number of proofs and sensitivity analyses

---

[1]In a sample of a state jurisdiction, Bushway et al. (2012) exploit random human calculation error in sentencing ranges to show judges are meaningfully constrained even by purely advisory guidelines.

are contained in the Appendices. Please note the first chapter is reproduced from the *International Review of Law and Economics* with permission by Elsevier (doi: http://www.sciencedirect.com/science/article/pii/S0144818816000090).

Chapter 2 Compromise Verdicts

Note: this section reproduced from the *International Review of Law and Economics*, Vol. 46, Lundberg, Alexander, 34–42, 2016, with permission from Elsevier (doi: http://www.sciencedirect.com/science/article/pii/S0144818816000090).

Sentencing is a mainstay of criminal justice. The aims of the justice system—rehabilitation, deterrence, retribution, incapacitation, moral expression—are all balanced through it. Who, then, should be trusted with its discretion? The common answer is judges. In the US, judges retain discretion both federally and in the majority of states. When a defendant is convicted, a judge sets a sentence within a range prescribed by law. However, several states employ jury sentencing, and the question of whether the judge or jury should sentence is a perennial debate.[1]

Currently, six states employ jury sentencing in non-capital cases.[2] That number is down from thirteen in 1960 (Iontcheva 2003), and the decline follows a long-standing distrust of the practice among legal scholars (e.g. Kerr 1918, Jouras 1952, Betts 1956, Webster 1960, Jackson 1999). Critics argue that expert judges offer more principled sentencing than untrained jurors (Roberts and de Keijser 2014). Specifically, jurors might give harsh, disparate sentences or even succumb to "compromise verdicts," where they return a guilty verdict but a light sentence when they are uncertain about the facts of a case.

This final criticism of jury sentencing is well summarized in the words of Betts (1956).

---

[1]It should be noted that judges and juries are neither the first nor the last exercisers of discretion in the criminal justice system. Legislators, prosecutors, parole boards, and patrol officers all have the authority to sentence in some capacity. Prosecutors in particular hold great authority in the age of plea bargaining (see both Schulhofer 1988 for a treatment of prosecutorial discretion and Kessler and Piehl 1998 for empirical support of its prominence). Nevertheless, the sentences imposed by judges and juries are of central importance since they largely inform the choices of other authorities in the system.

[2]The states include Arkansas, Kentucky, Missouri, Oklahoma, Texas, and Virginia (King and Noble 2004).

"Consideration of a light sentence, suspended sentence, or a quick parole may exert just the degree of influence necessary to persuade the doubtful juror to agree to a verdict of guilty." The fear is that such consideration will result in more wrongful convictions. Lending empirical support to this concern, Kaplan and Kupra (1986) find mock jurors more likely to convict when they know they will control the level of punishment. Lillquist (2004) further motivates the concern from a behavioral economics perspective. When people are given a choice between A and B, the introduction of a third choice, C, can lead people to the middle option through either the "compromise effect" or the "decoy effect."[3] The anticipation of regret might also encourage people to choose a middle option perceived as a safer bet (Zeelenberg 1999).

This article shows that even under an expected utility model, *any* fact finder—juror or judge—should react to sentencing discretion by imposing a lighter sentence and lowering his burden of proof beyond reasonable doubt. That is, sentencing discretion should lead to behavior that is observationally equivalent to a compromise verdict. The fact finder chooses a lighter sentence than the punishment that "fits the crime" because he wants to mitigate the potential cost of a wrongful conviction. His burden of proof is responsive to the sentence he chooses because a severe penalty increases the cost of a wrongful conviction. Consequently, when he chooses the lighter sentence, he also reduces his standard of proof.

Although critics of jury sentencing have cause to be wary of the practice, they frequently overlook the same temptation to compromise for a judge in a bench trial, where a judge returns both the verdict and the sentence. This omission is noteworthy since bench trials are the standard format in many countries and remain widespread even in countries employing juries (see Leib 2008 for a comparative review). If the temptation to compromise is present in any trial format in which a single party decides

---

[3]Let A be not guilty, B guilty with a light sentence, and C guilty with a harsh sentence. By the compromise effect, a juror will be more likely to choose B from the set {A,B,C} than from {A,B}. Switching the definition of B and C yields the same result by the decoy effect.

both the verdict and the sentence, then the conventional jury trial, where the jury returns the verdict and the judge returns the sentence, offers a possible solution. If a single party would be tempted to compromise, then dividing responsibility between two parties appears wise. The jury concerns itself with the facts, while the judge handles the sentence.

The focus here is on the temptation of a juror (or judge) to compromise on a verdict, but a related literature addresses jury sentencing more generally. A recent minority of advocates for the practice emphasize the importance of the jury as the conscience of the community (Lanni 1999, Hoffman 2003, Iontcheva 2003, Bibas 2012, Dzur 2012). Jurors are deemed more able to express the moral judgment of the public, and if their sentences are severe or disparate, those outcomes may not be inferior. Severity might reflect community outrage, and disparity might reflect a careful tailoring of sentences to individual cases (Iontcheva 2003). Furthermore, evidence is mixed on whether juries give relatively severe sentences. Robinson et al. (2010) find that respondents in surveys choose sentences far lower than those actually imposed in many cases, and Diamond and Stalans (1989) find that jurors and college students generally favor more lenient sentences than judges. On the other hand, in empirical studies of states with jury sentencing, Weninger (1994) and King and Noble (2005) find that juries do give both relatively harsh and disparate sentences.

Setting aside these broader questions in favor of a narrower focus on compromise verdicts, the rest of the article is outlined as follows. Section 2 presents the basic model of a fact finder in the absence of sentencing discretion. Section 3 introduces discretion and shows that, under reasonable assumptions, discretion should lead to lighter sentences and lower burdens of proof. This result offers a rationale for the common prohibition on providing juries with sentencing information, which is discussed in Section 4, along with the more general welfare implications of discretion. Section 5 briefly addresses empirical implications of the model for conviction rates,

and Section 6 concludes.

2.1 The Basic Model

A criminal trial serves two primary functions. One is the determination of guilt. The other is the imposition of a sentence if a defendant is found guilty. The authority tasked with the first function is called a fact finder. This authority may have no duty beyond returning a verdict but may, in some situations, also be tasked with choosing a sentence if the defendant is found guilty. The former case, where the fact finder must accept the sentence of another party, is the focus of this section. The latter case is addressed in the following section.

Let a fact finder in a criminal trial seek to return a correct verdict. She wants to convict when the defendant is guilty and acquit when the defendant is innocent. Consider two possible scenarios. In the first, the fact finder weighs the consequences of her verdict. In the second, she ignores them and simply compares the evidence at trial to a predefined threshold of reasonable doubt.

2.1.1 *Consequence of Verdict*

Suppose the fact finder weighs the consequences of her verdict, including the possible sentence to be imposed.[4] She incurs a cost $c_1(\ell)$ from a wrongful conviction and a cost c_2 from a false acquittal. The cost of a wrongful conviction depends on the sentence severity ℓ. For convenience, refer to severity as length. The longer the sentence, the more grieved is the fact finder for imposing it unjustly, and $c_1'(\ell) > 0$.[5] If

---

[4]This model is similar in spirit to that of Andreoni (1991). See also Davis (1994), Feess and Wohlschlegel (2009), and Dharmapala et al. (2014) for models involving a fact finder who balances the costs of wrongful conviction and false acquittal.

[5]Andreoni (1991) and Miceli (1990) also make the cost of a false acquittal depend on sentence length since the psychological cost of letting free the guilty is worse the more harmful the crime. Here no correlation is assumed between sentence length and criminal harm. The purpose of the model is to consider a change in sentence length for a *given* crime. Still, even for a given crime, the cost of letting free the guilty might depend on the sentence. Incapacitation, general deterrence, and specific

the conviction is rightful, punishment instead affords utility, which might come from a sense of retribution or personal justice. Let t be the sentence the fact finder would choose for a certainly guilty defendant, i.e., the punishment that "fits the crime." The utility from punishment $u(\ell; t)$ is a function of the target sentence t and the actual sentence ℓ, satisfying

$$\text{Assumption 1. } \frac{\partial u(\ell; t)}{\partial \ell} \begin{cases} > 0 & \text{if } \ell < t \\ < 0 & \text{if } \ell > t. \end{cases}$$

When the sentence is perceived as too lenient $(\ell < t)$, a longer sentence increases utility, and when the sentence is perceived as too harsh $(\ell > t)$, a shorter sentence increases utility. Normalizing the utility of a correct verdict to zero, $u(\ell; t)$ represents the additional utility from punishing the guilty. The importance of Assumption 1 will become clear in the following section.

The fact finder summarizes all evidence at trial in a probability p that a defendant is guilty. Without sentencing discretion, the fact finder must take ℓ as given, at least in expectation. The expected utility of conviction is then $\text{EU}_\text{c} = p \cdot u(\ell; t) - (1-p) \cdot c_1(\ell)$, while the expected utility of acquittal is $\text{EU}_\text{a} = -p \cdot c_2$. Whenever $\text{EU}_\text{c} \geq \text{EU}_\text{a}$, the fact finder will vote to convict. The inequality defines a threshold of reasonable doubt

$$p(\ell) = \frac{c_1(\ell)}{c_1(\ell) + c_2 + u(\ell; t)}. \tag{2.1}$$

For any $p \geq p(\ell)$, guilt is proved beyond reasonable doubt and the fact finder will convict. If $p < p(\ell)$, the evidence is not convincing enough and she will acquit.

---

deterrence are all potential costs, but since they are not privately borne, Feess and Wohlschlegel (2009) suggest that jurors might not fully internalize them. In any case, while c_2 and ℓ may not be independent, the results of the model are qualitatively unaltered with only minor assumptions on a function $c_2(\ell)$.

2.1.2 *Simple Threshold*

Assume now that the fact finder ignores the consequences of her verdict. She still summarizes all evidence in a probability p that the defendant is guilty, but now she compares p to a predetermined threshold of reasonable doubt, $\bar{p}$, say 90% or 95%. If $p \geq \bar{p}$, she will convict, and if $p < \bar{p}$, she will acquit.

2.2 Sentencing Discretion

Suppose in the model above that the fact finder is free to choose the sentence length ℓ. In this case, she necessarily evaluates the consequences of her verdict and sentence together. She anticipates that upon conviction she will choose the sentence that maximizes her expected utility.[6] The expected utility from conviction, $\mathrm{EU_c} = p \cdot \mathrm{u}(\ell; t) - (1 - p) \cdot c_1(\ell)$, is maximized according to the first-order condition:

$$\frac{\partial \mathrm{u}(\ell^*; t)}{\partial \ell} = \frac{(1 - p)}{p} c_1'(\ell^*), \tag{2.2}$$

where ℓ^* is the optimal sentence length. Since $c_1'(\ell) > 0 \ \forall \ell$, we have that $\partial \mathrm{u}(\ell^*; t)/\partial \ell > 0$, which by Assumption 1 implies $\ell^* < t$. Instead of choosing t, the punishment that fits the crime, the fact finder chooses a more lenient sentence to mitigate the potential cost of a wrongful conviction.

Note the importance of Assumption 1. The cost of making an incorrect decision cannot be the only thing motivating the fact finder, or else, if she chose to convict, she would always choose the lower bound on sentence length.[7] Clearly, she must derive some utility from imposing a just punishment, and Assumption 1 fills that role.

For (2) to define the optimal sentence length, the second-order condition must be

---

[6]The choice of ℓ is presumably bounded, and information gathering may occur between the verdict and sentence, but neither of these features affect the main results below.

[7]This statement would remain true if the cost of a false acquittal depended on the sentence length.

satisfied:

$$p\frac{\partial^2 u(\ell; t)}{\partial \ell^2} - (1-p)c_1''(\ell) < 0. \tag{2.3}$$

Assuming that $\partial^2 u(\ell; t)/\partial \ell^2 \le 0$, meaning that deviations from the target sentence length become more upsetting the farther away they are from the target, a sufficient condition for (3) is that the fact finder is risk-averse in the sentence of a wrongful conviction. (Since $c_1(\ell)$ is a loss, risk-aversion means that $c_1''(\ell) > 0$.) The fact finder can also be risk-loving provided she is especially sensitive to deviations from the target sentence.

After choosing ℓ^*, the threshold of reasonable doubt is

$$p(\ell^*) = \frac{c_1(\ell^*)}{c_1(\ell^*) + c_2 + u(\ell^*; t)}.$$

The key question is how this threshold compares to the one the fact finder would adopt without sentencing discretion, if she were required to impose the target sentence t. Section 2 presents the two main benchmarks for comparison.

First consider the case of Section 2.1, where the fact finder weighs the consequences of her verdict both with and without sentencing discretion. Denoting by $p(t)$ the threshold of reasonable doubt she would choose when required to impose the target sentence, the Appendix shows that $p(\ell^*) < p(t)$ if and only if

$$\frac{c_2 + u(t; t)}{c_2 + u(\ell^*; t)} < \frac{c_1(t)}{c_1(\ell^*)}. \tag{2.4}$$

A sufficient condition for (4) to hold is $c_1(t)/c_1(\ell^*) > u(t; t)/u(\ell^*; t)$; that is, the growth in the cost of a wrongful conviction in going from ℓ^* to t is bigger than the growth in utility from imposing a punishment closer to the target.

Proposition 1: Under condition (4), a fact finder with sentencing discretion will choose a lower burden of proof and a lighter sentence than a fact finder required to impose the target sentence t.

According to *Proposition 1*, the fact finder engages in what is essentially a compromise verdict. Instead of imposing the punishment that fits the crime, she imposes a lesser sentence and lowers her burden of proof. She may not be motivated by a psychological desire to compromise, but the result is identical. She chooses a lenient sentence to mitigate the potential cost of a wrongful conviction, and the lower cost of a wrongful conviction leads her to reduce her standard of proof.

While condition (4) is a reasonable assumption, it must be noted that if the cost of a false acquittal is low and the desire to punish greater than the fear of doing so improperly, the burden of proof could be higher when the juror is given discretion. However, there seems to be an inherent tension between a low cost of a false acquittal and a strong desire to punish, and condition (4) is unlikely to be violated in practice.

Next consider the case of Section 2.2, where, without sentencing discretion, the fact finder ignores the consequences of her verdict but adopts a predetermined threshold of reasonable doubt, $\bar{p}$. If $\bar{p}$ is relatively high, then $p(\ell^*) < \bar{p}$, and she again engages in a compromise verdict. If $\bar{p}$ is relatively low, then $p(\ell^*) \geq \bar{p}$, and discretion could actually lead to both a lower sentence and a higher burden of proof.

To summarize, if fact finders always weigh the consequences of their verdict, entrusting them with sentencing discretion is very likely to result in what are effectively compromise verdicts. If they ignore the consequence of their verdict but instead adopt a predetermined threshold of reasonable doubt, $\bar{p}$, the result depends on how high $\bar{p}$ is. If $\bar{p}$ is relatively high, as might be hoped and expected, sentencing discretion again yields compromise verdicts. Alternatively, if $\bar{p}$ is relatively low, discretion yields lower sentences and higher burdens of proof, to the complete benefit of defendants.

As a final note, this model describes a bifurcated trial, where guilt and punishment are determined sequentially. One complaint of jury sentencing concerns unitary trials, where guilt and punishment are determined simultaneously. In a unitary trial,

jurors might confuse facts relevant to sentencing, such as prior convictions, with facts relevant to the determination of guilt. In order to avoid such confusion, five of the six US states with jury sentencing adopt bifurcated trials, and the last (Oklahoma) adopts them for murder and repeat offender cases (Iontcheva 2003 and MO. REV. STAT. § 557.036 2004). Overall, bifurcation is the dominant form of trial.

2.3   Welfare and Discretion in Practice

To most people, the idea that an innocent defendant might be found guilty because she faces a light sentence is disturbing. From a social perspective, that idea is potentially less troubling (as noted by Lillquist 2004). If compromise verdicts lead to more convictions but lower average sentences, society must weigh that trade-off. In a justice system already characterized by error, an increase in error might be tolerable if the benefits, which include lower sentences for those wrongfully convicted, outweigh the costs.

In addition to balancing the risk of erroneous verdicts, an important aim of any sentencing policy is deterrence. To formalize the welfare analysis, suppose that a population of potential offenders must choose whether to commit a harmful act. Both the basic framework and the notation here follow Kaplow (2011, 2012). Each harmful act yields an external social cost h and a benefit b to the offender, where benefits come from the density function $f^H(b)$. Law enforcement is imperfect, and a fraction π of all individuals—whether they commit the act or not—are scrutinized and brought to court.[8] When an individual is scrutinized, the investigation produces both a set of evidence and a target sentence. The set of evidence can be summarized in the ostensible probability p that the individual is guilty. The target sentence t depends on the evidence because the sentencing authority might find aggravating or

---

[8]As described in Kaplow 2011, the results are largely unchanged if the fraction π differs for the guilty and innocent.

mitigating circumstances in either the facts or the disposition of the defendant.

The ostensible probability of guilt p and target sentence t are distributed via density function $g^H(p,t)$ for those who commit the act and $g^B(p,t)$ for those who do not (where H stands for "harmful" and B stands for "benign"). For any given t, the distribution function $G^H(p,t)$ first-order stochastically dominates $G^B(p,t)$ in p. That is, for any target sentence, and for any quantity of evidence, it is more likely that a guilty individual generates at least that much evidence. Under this assumption, stronger evidence makes the fact finder more inclined to convict. Lastly, all potential offenders are risk-neutral, and they choose to commit the harmful act if and only if their personal benefit outweighs the expected cost.

2.3.1 *Deterrence*

Consider social welfare in the absence of sentencing discretion. If the fact finders weigh the consequences of their decision, then a determinate sentencing scheme imposing $\bar{t}$ yields total welfare equal to

$$
\begin{aligned}
W_{\bar{t}} = &\int_T \int_{p(\bar{t})}^1 u(\bar{t};t)\, g^H(p,t)dpdt - \int_T \int_{p(\bar{t})}^1 c_1(\bar{t})\, g^B(p,t)dpdt \\
&- \int_T \int_0^{p(\bar{t})} c_2\, g^H(p,t)dpdt + w \int_{\pi\Delta G(\bar{t})\bar{t}}^\infty (b-h)\, f^H(b)db,
\end{aligned}
\tag{2.5}
$$

where T is the support of the target sentence, w is the relative weight placed on deterrence, and $\Delta G(t) \equiv G^B(p(t)) - G^H(p(t))$. The sentence is assumed costless to society for simplicity.[9] The first three terms capture the average utility from punishing the guilty, the cost of wrongful convictions, and the cost of false acquittals, respectively. For example, in the second term, since a fact finder chooses her burden of proof $p(\bar{t})$ for sentence $\bar{t}$ via (1), innocent defendants with evidence $p \geq p(\bar{t})$ are convicted, incurring cost $c_1(\bar{t})$. The last term is the benefit from deterrence.[10]

---

[9]The addition of a cost function on $\bar{t}$ modifies the results in predictable ways.

[10]Assume deterrence is incomplete ($\pi\Delta G(\bar{t})\bar{t} < h\, \forall\, \bar{t}$).

Potential offenders commit the harmful act if and only if their expected gain from the act, b, exceeds the expected cost, which is the difference between the expected punishment when breaking and obeying the law, $\pi\bar{t}[1 - G^H(p(\bar{t}))] - \pi\bar{t}[1 - G^B(p(\bar{t}))]$. Therefore, the last term integrates over those personal benefits where $b \geq \pi\Delta G(\bar{t})\bar{t}$.

Assuming the second-order condition holds, the first-order condition for (5) with respect to $\bar{t}$ conveys the intuition behind choosing the optimal determinate sentencing scheme. Rearranging the first-order condition gives

$$\int_T \int_{p(\bar{t})}^1 \frac{\partial \mathrm{u}(\bar{t};t)}{\partial \bar{t}} \ g^H(p,t)dpdt + \frac{\partial p(\bar{t})}{\partial \bar{t}} \int_T c_1(\bar{t}) \ g^B(p(\bar{t}),t)dt + \frac{\partial p(\bar{t})}{\partial \bar{t}} \int_T c_2 \ g^H(p(\bar{t}),t)dt$$

$$= \frac{\partial p(\bar{t})}{\partial \bar{t}} \int_T \mathrm{u}(\bar{t};t) \ g^H(p(\bar{t}),t)dt + \int_T \int_{p(\bar{t})}^1 \frac{\partial c_1(\bar{t})}{\partial \bar{t}} \ g^B(p,t)dpdt$$

$$+ w\left[\pi\Delta G(\bar{t}) + \pi\frac{\partial \Delta G(\bar{t})}{\partial \bar{t}}\bar{t}\right][\pi\Delta G(\bar{t})\bar{t} - h]f^H(\pi\Delta G(\bar{t})\bar{t}).$$

The expression might look complicated, but its message is simple. If a marginal increase in $\bar{t}$ raises the burden of proof, then a tougher sentence has six effects. An increase in $\bar{t}$ changes the average utility from punishing the guilty and raises the cost of a wrongful conviction. Since the burden of proof is now higher, more guilty individuals are falsely exonerated, fewer guilty are punished, and fewer innocents are wrongfully convicted. Finally, deterrence may go up or down, depending on whether a tougher sentence matters more to a potential offender than the new difference in conviction probabilities coming from the higher burden of proof. An optimal $\bar{t}$ requires each of these factors to be weighed accordingly.

Whether or not $\bar{t}$ is optimal, social welfare under a determinate sentencing scheme may be higher or lower than welfare under discretion, which is given by

$$\mathrm{W}_{\ell(t)} = \int_T \int_{p(\ell(t))}^1 \mathrm{u}(\ell(t);t) \ g^H(p,t)dpdt - \int_T \int_{p(\ell(t))}^1 c_1(\ell(t)) \ g^B(p,t)dpdt$$

$$- \int_T \int_0^{p(\ell(t))} c_2 \ g^H(p,t)dpdt + w \int_{\pi E_T[\Delta G(\ell(t))\cdot\ell(t)]}^\infty (b - h) \ f^H(b)db. \tag{2.6}$$

Here $\ell(t)$ denotes the sentence the fact finder would choose via equation (2) given a

target sentence t. Discretion creates the risk for compromise verdicts, which, again, may or may not be socially desirable themselves, but it also provides leeway in matching the punishment to the target sentence. Even though a fact finder will not choose the punishment that fits the crime (recall *Proposition 1*), this greater flexibility could still be a benefit on average. Of course, depending on whether discretion raises or lowers the expected penalty for the crime, it can either raise or lower deterrence as well.

Proposition 2: If society aims to deter offenders, punish the guilty, and avoid erroneous verdicts, then discretion is optimal if and only if (6) $\geq$ (5).

The result is conditioned explicitly on the aims of society because in some cases society might have additional objectives. Perhaps the most important relates to the "chilling" of desirable behavior.

2.3.2 *Chilling*

In some situations, the possibility of legal sanction deters not only harmful acts but also benign ones, a phenomenon which Kaplow (2011) names the "chilling" of desirable behavior. For example, a firm might avoid a promotional price if antitrust law enforcement might mistake the promotion for predation, or a doctor might avoid a risky surgery in fear of a medical malpractice suit. Suppose now that in addition to the population of potential offenders, another population of individuals chooses whether to commit a benign act with no social externality but with private benefit b, where benefits come from the density function $f^B(b)$. This population also chooses to commit the act if and only if the expected benefit exceeds the expected cost. However, in this set-up, only individuals who commit some type of act, whether harmful

or benign, are subject to possible scrutiny; inaction is not scrutinized.[11] Under a determinate sentencing scheme, social welfare contains an extra cost of chilling,

$$W_{\bar{t}}^c = \mathrm{W}_{\bar{t}} + \gamma \int_{\pi[1-G^B(p(\bar{t}))]\bar{t}}^{\infty} b \, f^B(b)db, \tag{5'}$$

where $\mathrm{W}_{\bar{t}}$ refers to equation (5), and γ is the weight placed on chilling (w and γ can also be thought of as relative population sizes).[12] Similarly, under discretion, social welfare is now

$$W_{\ell(t)}^c = W_{\ell(t)} + \int_{\pi E_T[[1-G^B(p(\ell(t)))]\cdot\ell(t)]}^{\infty} b \, f^B(b)db. \tag{6'}$$

Whether discretion is welfare-improving now also depends on whether the expected penalty under discretion is higher or lower than $\bar{t}$, since a higher (lower) expected penalty will chill more (less) desirable behavior.

Proposition 2': If society aims to punish the guilty, deter harmful acts but not chill benign ones, and avoid erroneous verdicts, then discretion is optimal if and only if (6') $\geq$ (5').

If the model is applied to a civil dispute between two private parties, the costs of a "wrongful conviction" and "false acquittal" need reinterpretation. A transfer of wealth from one party to the other is not necessarily a welfare cost at all.

2.3.3 *Trial Formats*

Depending on its relative costs and benefits, discretion may or may not be socially desirable.[13] Common trial formats shed light on the positions different societies have

---

[11] A firm will not be flagged by authorities if it does not offer a promotional price, nor will a doctor be flagged if she refuses to perform a surgery.

[12] There is an abuse of notation here because when only those who commit some type of act are scrutinized, the lower integral limit in the deterrence term of (5) should be $[1 - G^H(p(\bar{t}))]$ and for (6) it should be $E_T[[1 - G^B(p(\ell(t)))] \cdot \ell(t)]$.

[13] One issue not addressed here is that adjudicators might have different preferences from society at large (see, e.g., Shavell 2007).

taken, but those positions often appear inconsistent. Specifically, where discretion is given, its tendency to promote compromise verdicts is not fully addressed.

In a conventional jury trial, as adopted to varying degrees in common law countries, the jury returns the verdict and the judge sets the sentence. This trial format is not exclusive, and Table 2.1 presents the four basic combinations of responsibility for judges and juries.[14] Column (a) describes the jury trial, and column (d) has little practical relevance, but columns (b) and (c) are both common in practice. In the US, the Sixth Amendment guarantees the right to a trial by jury, but this right can typically be waived in favor of a bench trial, in which a judge is both fact finder and sentencing authority. In jury-sentencing states, the jury is both fact finder and sentencing authority. In either (b) or (c), the same judge or jury typically returns both the verdict and the sentence.

If the power to sentence can tempt a fact finder to compromise on a difficult verdict, and compromise verdicts are undesirable on average, then format (c) is clearly problematic, and critics of jury sentencing are right to be wary. Lost in the debate, however, is the question of why format (b) is not equally concerning. It is not clear why jurors but not judges would be tempted to compromise. Perhaps greater expertise makes judges more attentive to the danger. In a survey of Dutch judges asked how they would rule in different cases, de Keijser and van Koppen (2007) find no evidence of compromise verdicts. Still, the recognition that neither judges nor jurors are immune to temptation offers a basic argument for trial format (a). If entrusting a single party to choose both the verdict and the sentence creates a tendency to compromise, then one solution is to divide responsibility. Let the jury decide the facts and the judge decide the sentence.

The success of trial format (a) in stopping compromise verdicts is conditional

---

[14]In some countries, a panel of judges or a mixed panel of judges and jurors may decide both guilt and the possible sentence. The point of the division in Table 1 is not to ignore such cases but to highlight the general importance of assigning fact finding and sentencing responsibilities to either one party or two.

Table 2.1: Fact Finder and Sentencing Authority

	(a)	(b)	(c)	(d)
Verdict	Jury	Judge	Jury	Judge
Sentence	Judge	Judge	Jury	Jury

on the requirement that jurors ignore the consequence of their verdict. Under the expected utility model above, if the jury and judge have similar preferences and information, the jury will anticipate the sentence the judge will impose, and the outcome will be exactly the same as if the jury had been given discretion. The jury will lower its burden of proof in anticipation of the lenient sentence to come. The only difference is that the judge is the one imposing the sentence.

A large body of evidence suggests that juries do condition their verdicts on the expected sentence. Cassak and Heumann (2007) refer to the widely publicized Michigan Felony Firearm Statute of 1977 as an example of jury regard for punishment. The statute automatically added a sentence of two years to any felony committed by a person in possession of a firearm. Of the forty-three felonious assault cases carrying a charge under the "Gun Law," only three resulted in the mandatory two-year sentence. The consensus among judges, prosecutors, and jurors involved in the trials was that juries were aware of the law but circumvented it because they deemed the penalty too harsh. Kalven and Zeisel (1966) also explain the relaxation of a mandatory drunk driving penalty in the state of Indiana through the unwillingness of juries to convict at high sentences. A final, classic example is the "pious perjury" committed by juries in medieval England, who would value stolen goods just below a threshold mandating the death penalty.

According to US case law, juries are generally denied sentencing information in non-capital cases. Cassak and Heumann (2007) describe this prohibition as "a hard and fast rule with few exceptions." In the absence of punishment information, jurors will be pushed toward the adoption of a simple threshold of reasonable doubt, as in

Section 2.2. Thus, the system of trial format (a) coupled with a prohibition on providing jurors with sentencing information has a clear rationale. It protects innocent defendants from compromise verdicts.

On the other hand, the system can also harm defendants. Jurors might acquit more frequently if they were made aware of sentences they perceive as harsh and unfair. Moreover, the denial of punishment information to jurors might be an especially bad arrangement if jurors vote with their own, incorrect expectations of punishment in mind, and a minority of proponents believe juries should be more informed (Cassak and Heumann 2007, Bellin 2010).

If the denial of punishment information can in fact push jurors to adopt a simple threshold of doubt, then a determinate sentencing scheme becomes relatively more attractive. In equation (5), welfare can be maximized over both $\bar{t}$ and the burden of proof, which is no longer constrained by equation (1). However, even in that case, a determinate scheme may still be inferior to discretion if allowing the sentence imposed to vary with the target sentence is important for welfare. It is worth repeating that total welfare goes beyond compromise verdicts to include factors like deterrence.

2.4 Empirical Implications - Conviction Rate Puzzle

Ceteris paribus, if discretion leads to a lower burden of proof, it should lead to higher conviction rates. Since the model in Section 3 is directly applicable to a judge in a bench trial, a simple empirical test is to examine judicial rulings. Might verdicts in bench trials reveal any prima facie indication of compromise?

Figure 2.1 shows the annual conviction rates for federal juries and judges in US district courts from 1946-2010 (Bureau of Justice Statistics 2010). Over the last two decades, juries convicted at an average rate of 86%, while judges convicted at an
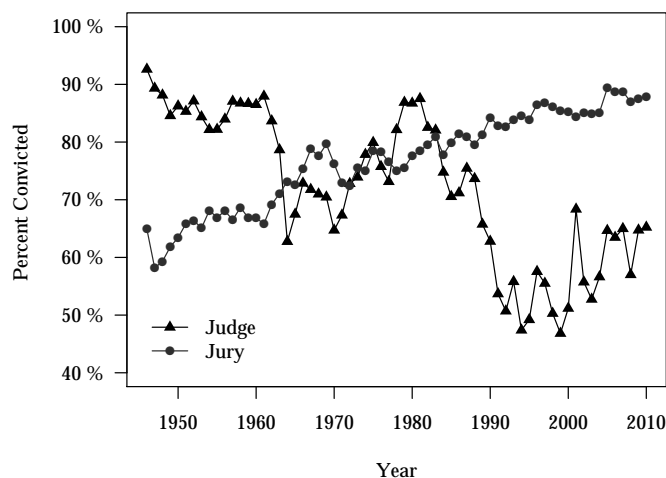
Figure 2.1: Federal Conviction Rates 1946-2010

average rate of 57%. The gap is surprising in light of the conventional wisdom that a defendant is better off going before an unpredictable jury—that a bench trial is a "slow guilty plea"—and it does not suggest that sentencing discretion is leading judges to compromise on verdicts, at least in recent years.

Of course, no fair comparison can be made without accounting for the selection of cases into bench and jury trials. Leipold (2005) finds that case selection only partially explains the gap in conviction rates, but even assessing the remaining gap is difficult because juries may have different information or preferences than judges.[15]

A full assessment of the gap would also involve historical developments. One interesting feature of the data in line with the model is the sharp drop in bench conviction rates alongside the imposition of federal sentencing guidelines in 1987. Were judges unwilling to convict at sentences perceived as excessive? Conviction rates appear to increase after *United States v. Booker,* 543 U.S. 220 (2005) rendered the guidelines advisory, though rates remain lower than before the guidelines were enacted.[16]

---

[15]For example, judges face one cost that jurors do not: the possibility of appellate review. The desire to avoid reversal may influence judges in their sentencing and conviction decisions. Shavell (2007) develops a theoretical model of how judicial decision making can be constrained by the prospect of appellate review, while Fischman and Schanzenbach (2011) find empirical evidence that judges are indeed constrained by the prospect.

[16]Compare Figure 2.1 here with Figure 1 in Fischman and Schanzenbach (2011) for an even

2.5 Conclusion

Economic theory raises concern over the possibility of compromise verdicts. Both expected utility theory and behavioral economics suggest that a fact finder with sentencing discretion will be tempted to adopt a lower burden of proof while imposing a more lenient sentence. Empirical work on the question of whether compromise verdicts occur in practice is unsurprisingly scant. As noted in Section 5, the identification of such verdicts is tricky, and conviction rates may not provide insight.

Although their welfare implications are ambiguous, compromise verdicts continue to be cited as a cost of jury sentencing. This argument seems partial given that bench trials currently afford judges the same opportunity to compromise. Again, expertise may protect judges from the temptation, but further work on the topic would be helpful.[17] A limitation of the model here is that only individual decisions are considered, but group dynamics may inspire a different tendency to compromise among juries and judges (see, e.g., Sunstein 2005). As theory stands, compromise verdicts are not an argument that is uniquely suited to jury sentencing. Critics who cite the risk of compromise must either explain why judges are immune or expand their disapproval to bench trials, which are widespread in many countries and jurisdictions.

---

more suggestive story that judges are more willing to convict when they are free to impose lighter sentences. After *US v. Booker*, both conviction rates and downward departures from the guidelines immediately increase.

[17]In an interesting parallel to this work, Siegel and Strulovici (2015, Working Paper) model and explore the welfare implications of a three-verdict system (see also Daughety and Reinganum 2015, Working Paper). In Scotland, for example, a defendant may be acquitted as either "not guilty" or "not proven," with the latter carrying perhaps a greater social stigma. Granting a fact finder the option of multiple verdicts may offer a coarse alternative to full sentencing discretion. Siegel and Strulovici (2015, Working Paper) find that offering multiple verdicts for convicted rather than acquitted defendants can be welfare-improving, but that a two-verdict system with plea bargaining may in fact be the optimal mechanism.

Acknowledgement

Chapter 3 Theoretical Model Expanded

In economic models of reasonable doubt, a trier-of-fact decides whether the evidence in at trial meets a threshold of reasonable doubt, which is chosen to balance the risks of a wrongful conviction and false acquittal (e.g., Andreoni 1991, Davis 1994, Feess and Wohlschlegel 2009, Friedman and Wickelgren 2006 [couched in Bayesian terms], Miceli 1990, Dharmapala et al. 2014). For example, the famous Blackstone ratio, "Better that ten guilty persons escape than that one innocent suffer," translates to a cost of a wrongful conviction equal to ten times the cost of a false acquittal, say, c. The burden of proof is then $10c/(10c + c) \approx .91$. The trier-of-fact should vote to convict if and only if he is 91% certain of guilt. Although the parsimony of the basic model is attractive, typical formulations omit two elements required for the examination of the conviction paradox and compromise verdict hypotheses.[1]

First, the actual sentence imposed needs to be separated from the target sentence, or the punishment that fits the crime. Many economic models perfectly correlate the actual and target sentence. Furthermore, the cost of a wrongful conviction needs to depend on the actual sentence. A longer sentence raises the cost of a wrongful conviction.

Second, punishing the guilty must offer a direct utility. Without that utility, the reasonable doubt model has an unsatisfying feature. If the cost of a wrongful conviction is increasing in the length of the sentence, then a trier-of-fact with the authority to sentence would, upon conviction, never impose any sentence.

---

[1]The use of the word "omit" is not meant to criticize the models; they are employed for other purposes in which the "omissions" are useful assumptions.

3.0.1 Assumptions

The model is an elaboration of Lundberg (2016). A trier-of-fact in a criminal trial wants to return a correct verdict. That is, he wants to convict when the defendant is guilty and acquit when the defendant is innocent. A wrongful conviction incurs a cost $c_1(l)$, which depends on the sentence severity, or "length" of imprisonment, l. Since a longer sentence is more harmful to an innocent person, $c_1'(l) > 0$. A false acquittal incurs a cost $c_2(t)$, which depends on the target sentence, t, or the punishment that "fits the crime." Since letting a crime go unpunished is worse for more serious crimes, $c_2'(t) > 0$. (The end of this section examines the possible dependence of $c_2(\cdot)$ on l.)

In his decision, the trier-of-fact balances the risks of wrongful conviction and false acquittal, but the cost of an incorrect verdict cannot be the only thing motivating him. Otherwise, if he were free to choose the sentence upon conviction, he would not impose any sentence at all. In other words, a rightful punishment must offer some benefit. The utility from punishment $u(l,t)$, perhaps coming from a desire for retribution or deterrence, is a function of the target sentence t and the actual sentence l, satisfying

$$\text{Assumption 1}: \frac{\partial u(l,t)}{\partial t} \geq 0, \text{ and } \frac{\partial u(l,t)}{\partial l} \begin{cases} > 0 & \text{if } l < t \\ < 0 & \text{if } l > t. \end{cases}$$

A more serious crime always results in at least as much punishment utility as a less serious one. When the sentence is too lenient ($l < t$), a longer sentence increases utility, and when the sentence is too harsh ($l > t$), a shorter sentence increases utility. Normalizing the utility of a correct verdict to zero, $u(l,t)$ represents the additional utility from punishing the guilty.

Lastly, the trier-of-fact summarizes the evidence at trial in a probability p that a defendant is guilty.

3.0.2 Burden of Proof—No Discretion

Without the authority to set the sentence, the trier-of-fact must take l as given, at least in expectation. This scenario represents either a determinate sentencing scheme or a single juror who must accept the sentence of a judge. The expected utility of conviction is $EU_c = p \cdot u(l,t) - (1-p) \cdot c_1(l)$, while that of acquittal is $EU_a = -p \cdot c_2(t)$. The trier-of-fact votes to convict if and only if $EU_c \geq EU_a$, and the inequality defines a threshold of reasonable doubt

$$p(l,t) = \frac{c_1(l)}{c_1(l) + c_2(t) + u(l,t)}. \tag{3.1}$$

For any $p \geq p(l,t)$, guilt is proved beyond reasonable doubt, but if $p < p(l,t)$, the evidence is not convincing enough for a conviction.

The response of the burden of proof to a change in the target sentence is

$$\frac{\partial p(l,t)}{\partial t} = -\frac{c_1(l)[c_2'(t) + \frac{\partial u(l,t)}{\partial t}]}{[c_1(l) + c_2(t) + u(l,t)]^2}. \tag{3.2}$$

Holding the sentence constant, a more serious crime always results in a lower burden of proof. Notice the result holds under any cost of an incorrect verdict and any utility of punishment, as well as any given sentence, even one considered too harsh.

Claim 1: $\dfrac{\partial p(l,t)}{\partial t} < 0.$

The response of the burden of proof to a change in the actual sentence is

$$\frac{\partial p(l,t)}{\partial l} = \frac{c_1'(l)[c_2(t) + u(l,t)] - c_1(l)\frac{\partial u(l,t)}{\partial l}}{[c_1(l) + c_2(t) + u(l,t)]^2}. \tag{3.3}$$

If the sentence is too harsh ($l > t$), then $\partial u(l,t)/\partial l < 0$ by Assumption 1, and $\partial p(l,t)/\partial l > 0$. A longer sentence always increases the potential cost of a wrongful conviction, so when it also decreases the utility of punishing the guilty, the trier-of-fact naturally responds by lowering his burden of proof.

If the sentence is lenient ($l < t$), then $\partial p(l,t)/\partial l$ has an indeterminate sign. A longer sentence still worsens a wrongful conviction but now increases the utility from punishing the guilty. The two effects work in offsetting directions.

Proposition 3: If the sentence is harsh ($l > t$), then $\dfrac{\partial p(l,t)}{\partial l} > 0$. If the sentence is lenient ($l > t$), then $\dfrac{\partial p(l,t)}{\partial l} \lesseqqgtr 0$.

In both *Claim 1* and *Proposition 3*, either the actual or the target sentence is held constant. However, the general version of the conviction paradox would allow for both the actual sentence and the burden of proof to vary simultaneously.

3.0.3 Burden of Proof—Sentencing Discretion

Suppose the trier-of-fact is now free to choose the sentence upon a conviction. One interpretation is a bench trial, in which the judge decides both the verdict and the sentence. Another is that of a juror in the same situation. in one of the six states employing jury sentencing (King and Noble 2004). Lastly, the juror and judge might share identical preferences, with the juror anticipating the sentence to come.

In this more general environment, the intuition of the conviction paradox is thus: A more serious crime demands a higher sentence, but the burden of proof must then account for both the extra cost of a wrongful conviction and the greater desire to punish the serious offense, so the net effect is indeterminate. Although the intuition is accurate, the burden of proof in fact rises in response to the target sentence unless the relative desire to punish is extreme.

If the trier-of-fact can choose the sentence, he anticipates that upon a conviction he will choose the sentence that maximizes his expected utility. The expected utility from conviction, $EU_c = p \cdot u(l,t) - (1-p) \cdot c_1(l)$, is maximized according to the first-

order condition

$$\frac{\partial u(l^*, t)}{\partial l} = \frac{(1-p)}{p} c_1'(l^*),$$

(3.4)

where l^* is the optimal sentence length. Since $c_1'(l) > 0$ $\forall l$, we have that $\partial u(l^*, t)/\partial l > 0$, which implies $l^* < t$ by *Assumption 1*.

Claim 2: $l^* < t$.

Although the punishment does vary to match the severity of the crime, the trier-of-fact chooses a more lenient sentence than t to mitigate the potential cost of a wrongful conviction.

Notice again the importance of having a direct utility from punishment. If the cost of an incorrect verdict were the only thing motivating the trier-of-fact, he would always choose the lower bound on sentence length for any conviction.[2]

For (4) to define the optimal sentence length, the second-order condition must be satisfied:

$$p\frac{\partial^2 u(l, t)}{\partial l^2} - (1-p)c_1''(l) < 0.$$

(3.5)

If deviations from the target sentence become more upsetting the farther they are from the target $(\partial^2 u(l, t)/\partial l^2 < 0)$, a sufficient condition for (5) is weak risk aversion in the cost of a wrongful conviction. Recall that $c_1(l)$ is a loss, so weak risk aversion is equivalent to $c_1''(l) \geq 0$ $\forall l$. Condition (5) is assumed going forward.

With the optimal sentence in mind, the optimal threshold of reasonable doubt, p^*, is simply Equation (1) evaluated at l^*. The question is then, how do l^* and p^* respond to an increase in the severity of the crime? First we require a mild regularity assumption:

---

[2]This claim does not come from the independence of l and $c_2(\cdot)$. It comes from the bifurcated decision, in which the trier-of-fact first returns a verdict and then decides a sentence. Bifurcation is the predominant trial format in the US. Unitary trials, in which the trier-of-fact simultaneously determines the facts and the sentence, are very unusual.

Assumption 2: $\dfrac{\partial^2 u(l^*; t)}{\partial l \partial t} \geq 0.$

Since $l^* < t$, the marginal utility of punishment is presumably weakly increasing in t. When the sentence is already below the target, it is hard to imagine a more serious crime yielding less punishment utility. Under *Assumption 2*, the response of l^* to a change in t is unambiguous. A more serious crime is always met with a longer sentence. Stated without proof,

Proposition 4: $\dfrac{\partial l^*}{\partial t} \geq 0.$

Next, with the longer sentence, the optimal burden of proof may go up or down. On the one hand, a longer sentence increases the cost of a wrongful conviction. On the other, both the cost of a false acquittal and the utility of punishment go up when the crime is more serious.

Proposition 5: Depending on preferences, $\dfrac{\partial p^*}{\partial t} \lesseqqgtr 0.$

However, constructing examples for which the burden of proof goes down requires unrealistic parameters; see Appendix A for details. As an example, consider preferences of the following form:

Assumption 3: $u(l, t) = -a(l - t)^\alpha + bt^\beta + c$, $c_1(l) = dl^\alpha$, and $c_2(t) = et^\gamma$, for some $a, \beta, \gamma > 0$; $b, c, d, e \geq 0$; $\alpha > 1$. Also, $\beta, \gamma \leq \alpha$.

Since *Assumption 1* implicitly requires α to be an even number, the utility of punishment is increasing in the target sentence.[3] That α appears in both $u(\cdot, \cdot)$ and $c_1(\cdot)$

---

[3] Depending on the parameters, $\frac{\partial u(l,t)}{\partial t} \geq 0$ might not hold for small l, but $\frac{\partial u(l^*,t)}{\partial t} \geq 0$, and *Assumption 1* holds in the range of interest.

is merely an analytical convenience. The restriction on $\beta, \gamma \leq \alpha$ is sufficient but not necessary for the following result.

Claim 3: Under *Assumption 3*, both $\dfrac{\partial l^*}{\partial t} > 0$ and $\dfrac{\partial p^*}{\partial t} > 0$.

The intuition behind *Claim 3* is that when the crime is more severe, the fact finder chooses a tougher sentence but then raises his burden of proof to lessen the now higher potential cost of a wrongful conviction.

One way to loosely interpret $\beta, \gamma \leq \alpha$ is by saying the trier-of-fact is not, at the margin, more invested in punishing the guilty than protecting the innocent. However, the result holds *even though* the marginal increase in either the cost of the false acquittal or the utility from punishment (both from the increase in t) may be far greater than the marginal increase in the cost of a wrongful conviction (from the increase in l^*). By letting b and e grow, the former can dominate the latter.

Claim 4: Under *Assumption 3*, $\dfrac{\partial p^*}{\partial t}$ can be greater than zero with $\dfrac{\partial u(l^*, t)}{\partial t}$ and $c_2'(t)$ arbitrarily greater than $c_1'(l^*)$.

Finding preferences for which $\partial p^* / \partial t < 0$ is possible but requires the juror to be far more invested in punishment than protecting the innocent.

While *Proposition 4* showed the burden of proof to have an ambiguous response to an increase in the actual sentence l (for $l < t$), *Claim 3* shows that when the sentence is optimally chosen (recall $l^* < t$), the net response to criminal severity can in fact be signed, positively, for a wide range of preferences.

3.0.4 A Brief Note on the False Acquittal

The results above assume the cost of a false acquittal is independent of the sentence. Yet in practice, a longer sentence can either increase or decrease the cost of a false acquittal. A longer prison sentence might better deter or incapacitate a violent offender, thereby increasing $c_2(\cdot)$. Alternatively, a longer prison sentence might harden an offender, reduce his employment opportunities, and expose him to larger criminal networks, thereby decreasing $c_2(\cdot)$.[4] A natural question is how the results hold up if $c_2(\cdot)$ is allowed to depend on l.

To answer that question, *Claim 1*, *Proposition 4*, and *Claim 3* still hold with minor modifications, but the remaining results become fully ambiguous because the optimal burden of proof must now account for the impact of the sentence on the cost of a false acquittal, which could be positive or negative. However, there is one natural case where everything simplifies nicely—when the cost of a false acquittal is simply the forgone benefit of punishment.

If $c_2(l, t) = u(l, t)$, then all results go through in a simplified form. For example, take *Proposition 3*. The optimal burden of proof is

$$p(l, t) = \frac{c_1(l)}{c_1(l) + 2u(l, t)},$$

and

$$\frac{\partial p(l, t)}{\partial l} > 0 \iff \frac{c_1'(l)}{c_1(l)} > \frac{\dfrac{\partial u(l, t)}{\partial l}}{u(l, t)}.$$

The convenient interpretation is that the burden of proof rises if and only if the growth in the cost of a wrongful conviction is greater than the growth in the utility of punishment.

---

[4]As an example of why this reasoning is more than conjectural, take the reckless driving case cited in Bartels (1981). The judge was so firmly convinced the defendant was guilty that when the jury acquitted, he asked for their reasoning. They said they believed the defendant, a teenager, would be sentenced to prison, which would increase the probability of his future criminality.

3.0.5 Compromise Verdicts

In the preceding model, Lundberg (2016) shows that in response to sentencing discretion, both the optimal sentence and burden of proof go down. The exact arguments are not repeated here, but the intuition is that from *Claim 1*, the optimal sentence is less than the target. Next, the burden of proof goes down because the lighter sentence lowers the potential cost of a wrongful conviction. A sufficient condition for the result to hold is: $c_1(t)/c_1(l^*) > u(t;t)/u(l^*;t)$. If, in going from the smaller sentence l^* to the larger sentence t, the growth in the cost of a wrongful conviction outweighs the growth in the utility from punishing the guilty, then a trier-of-fact with sentencing discretion will engage in what looks to be a compromise verdict. Compared to a determinant sentence scheming, both the sentence and the burden of proof will be lower under a discretionary regime.

Chapter 4   Empirical Results

The two hypotheses of the Chapter above are tested using federal trial data over the last decade. The primary data source is the Federal Justice Statistics Program of the Department of Justice (ICPSR). Information is provided by three federal agencies. The United States Sentencing Commission (USSC) provides sentencing outcomes, and the Administrative Office of the US Courts (AOUSC) provide trial outcomes. Defendant characteristics (age, race, gender, marital status) are taken from the US Marshals' Service (USMS) arrest records because they are not available in AOUSC files. Thus, defendants are linked from the USMS to the AOUSC to the USSC.[1]

Data cover all federal trials for the years 1998-2010, the most recent year available. Federal defendants are tried in one of 94 district courts, each of which randomly assigns cases to judges. Five border districts are removed from the analysis because they are uniquely formatted to handle a high volume of immigration issues.

The data are best described as multi-level. A single observation is indexed by three levels: the trial, the district in which it took place, and the date on which it took place. Even though districts are geographically static over the time period, the data are not structured as a panel because the unit of observation is the single trial.

The secondary data source is Biographical Directory of Federal Judges maintained by the Federal Judicial Center (FJC). The USSC does not (and will not) provide judge identifiers, so judicial characteristics are monthly, district averages of active judges in the FJC data. For example, if a defendant is tried in a district and month containing five judges, two of which are female, the "percent female" variable will equal 0.4 for

---

[1]Observations are linked through a dyadic file that pairs both inter-agency and intra-agency files. There are two possible pathways to link data from the USMS to the USSC. Each initially runs through the Executive Office for the US Attorneys (EOUSA) suspect investigation file. Both pathways are utilized to minimize missing values. For details on the linking algorithm, see Starr (2015).

the observation.

Ethnicity is not recorded by the USMS, so Hispanic defendants are either coded under black or non-black in the race dummy. The variable for criminal history points theoretically has no upper bound, but individual records equal to 98 or 99 are assumed missing since those values are implausibly high. To avoid issues of multiple convictions and concurrent sentences, the analytical sample only includes cases with a single charge. After dropping the remaining observations with missing values, the final sample totals 765 trials. The jury trial results reported in Table A1 are obtained through the same mechanism described above. Jury trials have been more frequent than bench trials over the last decade, and the sample size is larger: 3,447.

4.1 Estimation Strategy

Understanding the test for each hypothesis requires an understanding of the framework in which judges make decisions. Specifically, judges are not just the moderators of the courtroom. A judge in a bench trial decides both the law and the verdict. Bench trials can therefore be used to test if judges apply a higher burden of proof to more serious charges. Additionally, federal sentencing policy has shifted several times over the last decade, each time effectively shifting judicial discretion, and those changes are exploited in the test of compromise verdicts.

4.1.1 Institutional Background

The federal judicial system is comprised of three court tiers. The first is the district or trial court. Federal defendants who contest their guilt are tried in one of 94 US district courts. Next, appeals are handled by one of 11 circuit or appellate courts. Lastly, the Supreme Court of the United States sits atop the system as the highest court in the land. Federal judges are appointed by the President for life, and cases are randomly allocated to judges within each district.

As part of the Comprehensive Crime Control Act of 1984 (Pub.L. 98–473, S. 1762, 98 Stat. 1976, October 12, 1984), Congress founded the United States Sentencing Commission (USSC). Prior to the Act, judges were free to sentence within a wide range allowed by law. The USSC was created in response to the perceived sentencing inequity that similarly culpable offenders would face when going before different judges; its principle job is to develop guidelines for sentencing convicted federal offenders. Those guidelines take the form of a grid or table. One axis contains the severity of the offense, while the other contains the criminal history of the offender (see Figure A1). Upon a conviction, a judge calculates the offender's position in the table and chooses a sentence within the given range. The ranges are fairly narrow, with the lower bound typically being about 75% of the upper bound.

The guidelines became binding in 1987, at which point federal parole was abolished. Judges were able to depart from the guidelines, but any departure could be reversed on appeal. Over time, Congress grew concerned about the frequency of downward departures and took steps to reduce them in part of an omnibus crime bill, the PROTECT Act (Pub.L. 108–21, 117 Stat. 650, S. 151, April 30, 2003).[2] In the years preceding the Act, departures were reviewed under an abuse-of-discretion standard. The PROTECT Act instated a stricter de novo standard of review. It also strengthened reporting requirements and put extra limits on the grounds for departure. Furthermore, then Attorney General John Ashcroft urged prosecutors to oppose downward departures (Freeborn and Hartmann 2010).

Roughly a year and a half later, in a watershed case, the US Supreme Court excised the portion of law making the guidelines mandatory upon judges (*United States v. Booker*, 543 U.S. 220 [2005]). As courts grappled with how much weight to afford the guidelines, now formally advisory, the Court clarified that a sentence within the guidelines

---

[2] "Substantial assistance" departures, in which the government departs downward in exchange for the defendant's cooperation in an investigation, are not relevant here because the sample includes only cases going to trial.

range could be held presumptively reasonable in *Rita v. United States*, 551 U.S. 338 (2007). But shortly afterward, the Court all but eliminated the safe harbor of the guidelines in ruling that sentences outside of them could not be held presumptively *un*-reasonable (*Gall v. United States*, 552 U.S. 38 [2007]). Furthermore, on the same day as *Gall*, the Court declared that judges may depart from the guidelines if they disagree with the congressional policy of treating crack and powder cocaine in a 100-to-1 disparity (*Kimbrough v. United States*, 552 U.S. 85 [2007]). The rulings in *Gall* and *Kimbrough* established a clear precedent for judicial discretion in sentencing.

4.1.2 Severity of Crime

The burden of proof may not be observable, but the decision of whether to convict is. To test whether judges increase their burden of proof with the severity of the crime (or, conversely, drop the burden with the lack of severity), a natural first step is to estimate the relationship between the probability of conviction and the seriousness of the charge. Perhaps the best available measure of the latter is the "filing severity code" of the AOUSC database. The first digit of the code provides a category for the possible sentence. Table 4.1 provides the reference categories.

A negative relationship between the probability of a conviction and the severity of the offense would offer prima facie evidence of a changing burden of proof. A probit regression can test for such a relationship after controlling for other defendant characteristics, but the lack of a good measure for the strength of the case is an obstacle. If the strength of the evidence is randomly distributed across cases, its omission is no great problem.[3] However, certain types of crimes might be more difficult to prove than others, and a serious charge increases the incentives for both the defense and the prosecution to put forward the best case possible (as well as whether to strike a plea deal). If incentives are asymmetric, the burden of proof

---

[3]Recall that omitted variables in a probit regression can bias estimates even when they are uncorrelated with the regressors, but average marginal effects, arguably the object of greatest interest, remain consistent (Wooldridge 2010).

might differ systematically with the charge.

The results below cannot rule out a story whereby changes in conviction probabilities are driven entirely by changes in evidentiary strength, but two factors lessen the possibility. The inclusion of counsel type in the model can partially address any difference in evidence attributable to the quality of the defense lawyer. More importantly, it is unclear why the distribution of evidentiary strength would differ across jury and bench trials. Judges are are intimately familiar with sentencing, while jurors, aside from being instructed to ignore it altogether, are unlikely to be familiar with sentencing in the first place. If judges become less likely to convict when the charge is serious, but jurors do not, then a variation in the burden of proof becomes a more likely story. In fact, separate probit regressions for bench and jury trials yield just that result.

Table 4.1: Severity of Filing Offense

0	None
1	6 mos and under
2	7 mos - 1 yr
3	1 yr 1 day - 2 yrs
4	2 yrs 1day - 3 yrs
5	4 - 5 yrs
6	6 - 10 yrs
7	11 - 15 yrs
8	16 - 20 yrs
9	21 - 25 yrs
10	Over 25 yrs
11	Life
12	Death

Note: The first three categories of the "FSEV1" variable in the AOUSC database have been recoded from "A," "B," and "C" to -3, -2, and -1. The resulting variable is translated by three to begin at zero.

A lesser obstacle is the omission of any judge identifier. The USSC does not iden-

tify judges, so judicial characteristics have to be averaged from biographical information at the district level. Although Schanzenbach (2005) finds judicial characteristics have a limited overall effect on sentencing outcomes, they might still affect conviction decisions differently and are included in all regressions.

4.1.3 Compromise Verdicts

The idea of a compromise verdict requires a joint test. Do judges both convict more often and impose lower sentences in periods of greater judicial discretion? Consider how changes in review standards over the last decade have effectively varied the freedom of judges to choose a sentence outside of the guideline range. For various reasons, judges are averse to having their decisions overturned on appeal. A reverse and remand creates more work for a judge, incurs an obvious disutility, and might affect promotion opportunities if reversals become frequent (Epstein et al. 2013). Thus, when review standards are lax, judges have more discretion in sentencing. Fischman and Schanzenbach (2011) and Freeborn and Hartmann (2010) provide evidence that judges are in fact constrained by the prospect of appellate review. One way to test for compromise verdicts is then to check whether conviction probabilities go up while sentences go down in periods of relatively lax review standards.

Figure 1 shows the monthly rates of conviction and departures from the guidelines for judges in bench trials, along with the timing of the important laws and Supreme Court decisions outlined in Section III.1. These policy changes create the variation in judicial sentencing discretion exploited in the empirical test of compromise verdicts. The first period is one of relative discretion, with lax review standards (*Koon v. US*), while the second is one of low discretion and strict review standards (PROTECT Act). In January 2005, the *Booker* decision created more discretion by rendering the guidelines advisory, and departure rates notably rise. In 2007, the Court ruled that a sentence within the guidelines could be held presumptively reasonable (*Rita v. US*),

but later that year, the Court simultaneously decided *Gall* and *Kimbrough*, stating (1) an out-of-range sentence could not be held presumptively unreasonable, and (2) judges were free to ignore the Congressional mandate of treating crack and powder cocaine in a 100-to-1 disparity. The period between *Rita* and *Gall/Kimbrough* can be construed as a period of less discretion since the guidelines provided a safe harbor for a judge averse to review. The period after *Gall/Kimbrough* is one of unequivocal discretion.

Figure 4.1: Monthly Rates of Conviction and Guideline Departures in Bench Trials
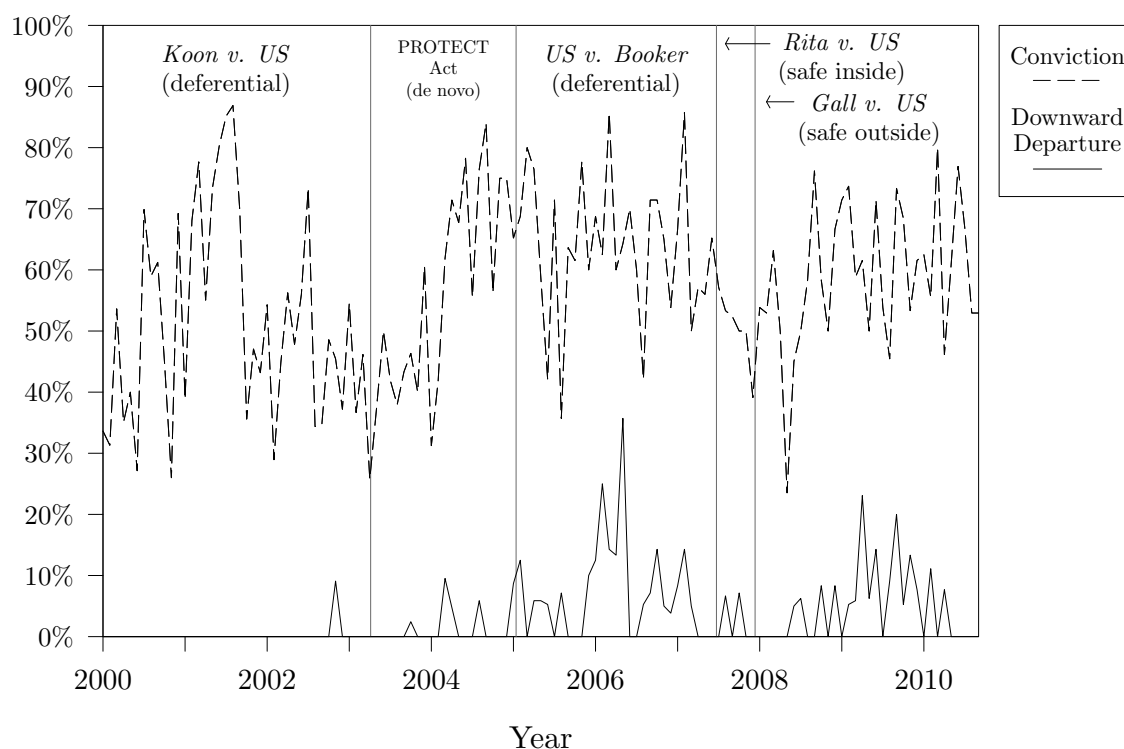
Although Figure 4.1 contains a fair amount of noise, downward departures and conviction rates do bear something of a positive relationship. The question is whether and to what extent a positive relationship exists after accounting for defendant and judge characteristics.

A judge in a bench trial first decides whether to convict and then, if applicable, decides the sentence. The two-part model is therefore an appropriate estimation

strategy (Cragg 1971). The model is an extension of a probit maximum likelihood estimator, which allows for a joint test of variables across both a selection equation, here the conviction decision, and a conditional equation, here the length of the sentence. After including a dummy variable for periods of greater discretion, the joint test of its significance is a check for whether judges increase the probability of conviction and reduce sentence length when given more discretion.

Most criminological research focuses on sentencing outcomes. Studies typically condition on a conviction and apply a two-part model to the decision of whether to incarcerate or not (Bushway et al. 2007). One reason to focus on the incarceration decision is the difficulty of comparing prison and non-prison sentences. In terms of the model, the issue is that a zero from a positive selection is not the same as a zero from a negative selection. A conviction is a sanction per se. This study accepts the difficulty of comparing sentences because some of type of comparison is required to test for compromise verdicts. Furthermore, results are robust to reasonable translations. Coding a conviction with a zero prison time as equivalent to a sentence of anywhere between one day and one month does not qualitatively change the relationship between the variables of interest.

Criminological research also tends to adopt the Heckman correction for sample selection (Bushway et al. 2007, Heckman 1979). The Heckman model technically nests the log-normal hurdle model used here (Wooldridge 2010), but the correlation coefficient between the error terms of the selection and outcome equation is small and insignificant ($\rho = .03$), so results are presented only for the two-part model.

4.2 Results

Table 4.2 presents the summary statistics for the sample of bench trials used in the estimation. "Severity" is a categorical variable for the possible sentence attaching

to the offense; see Table 1 above. "Discretion Regime" is a dummy set to one if the case disposition occurred during a period of relatively greater discretion. It is equal to one during the *Koon*, *Booker*, and *Gall* periods, and zero otherwise; see Figure 1 above. "History Points" is only calculated for convicted offenders. It refers to the subtotal of criminal history points as determined by the sentencing guidelines.

"Public," "Private," and "Appointed" are dummies for the type of defense counsel, which leaves the reference category as pro se defense. Lastly, the judicial variables refer to the percentage of judges with a given characteristic in the district and time of the case disposition.

Table 4.2: Summary Statistics

	Stage One				Stage Two			
	avg.	sd.	min.	max.	avg.	sd.	min.	max.
Sentence Length	-	-	-	-	2.31	2.47	-0.69	6.90
Severity	4.21	3.30	0	12	5.25	3.26	1	12
Discretion Regime	0.90	0.30	0	1	0.87	0.34	0	1
Age	36.04	12.63	14	87	33.92	10.98	18	77
Black	0.33	0.47	0	1	0.41	0.49	0	1
Male	0.82	0.39	0	1	0.80	0.40	0	1
Married	0.23	0.42	0	1	0.14	0.34	0	1
History Points	-	-	-	-	3.65	4.81	0	24
Public	0.28	0.45	0	1	0.41	0.49	0	1
Private	0.28	0.45	0	1	0.18	0.39	0	1
Appointed	0.20	0.40	0	1	0.25	0.43	0	1
%White Judge	0.79	0.28	0	1	0.88	0.12	0	1
%Female Judge	0.11	0.10	0	0.33	0.12	0.11	0	0.50
%Democrat Judge	0.45	0.14	0	0.74	0.44	0.14	0	0.71

Note: n=765 for stage one; n=300 for stage two. The Sentence Length variable is the natural log of the months of prison time imposed.

Table 4.3 presents the results from estimating the two-part model. The first stage is a probit regression, and the second stage is an ordinary least squares regression on the log of the sentence length. For the purposes here, the advantage of the two-part

model is the ability to conduct joint significance tests across both parts of the model.

The sentencing guidelines are calibrated above mandatory minimums, which in some cases limit a judge from departing downward. Only a handful of cases in the sample involved mandatory minimums, and a dummy for the presence of a minimum was perfectly collinear with other regressors and therefore dropped from the anaysis.

4.2.1 Severity of Crime

The severity of the crime significantly reduces the likelihood of conviction. Moving up one category in the severity index reduces the probability of conviction by 1.9 percentage points, and the effect is nearly significant at the .01 level. Being married reduces the conviction probability by 7.8 percentage points, significant at the .05 level. Having a public defender was the only other variable significant at the .10 level. Having a public defender raises the probability of conviction relative to a pro se defense.

4.2.2 Compromise verdicts

In each equation, the sign of the coefficient on the discretion regime dummy matches the prediction. Judges convict more frequently when given more discretion and impose more lenient sentences. However, the latter effect is not statistically significant. Likewise, a chi-square test of joint significance across the two stages fails to reject the null, with a p-value of 0.126.

4.2.3 Robustness

Results are robust to a variety of extensions, such as imputing the criminal history score for missing observations and splitting the severity variable into mutually

exclusive dummies. Please see Appendix B for a discussion of sensitivity analysis.[4]

4.3   Discussion

The results provide limited support of compromise verdicts, but, compatible with economic reasoning, judges do significantly reduce their probability of conviction when the crime is more serious. Again, the strength of the evidence might systematically vary with the severity of the crime. The defense certainly has a powerful incentive to present the best possible case when the charge is serious. On the other hand, so does the prosecution. Boylan (2005) finds career prospects of US Attorneys are partially determined by the length of prison sentences, and serious charges carry higher sentences.

The type of attorney might also serve as a crude proxy for expected quality of the defense. Counsel may be hired privately, appointed by the court, or provided by a public defender organization. Alternatively, the defendant may forego the option of counsel and represent him or herself pro se. Court-appointed attorneys are typically paid hourly while public defenders are salaried. Two studies find the former obtain worse outcomes for their clients than the latter (Iyengar 2007, Roach 2014). In any case, judicial perception of attorney quality is what truly matters for bench trials. In a survey of district court judges, Posner and Yoon (2011) find prosecutors and public defenders to be rated with the highest quality, while court appointed and retained counsel are seen as lower quality on average (though disparities in the two sides of legal representation are generally perceived as larger in civil than criminal litigation). In the two-part model above, the reference category for legal representation is pro se defense. Public defenders are more likely to generate convictions, and appointed

---

[4]The marginal effect on the black-male interaction dummy is not correct because it is calculated from a nonlinear index model (Norton et al. 2004). Appendix B presents the correct values, but the marginal effect is not statistically significant for any observation.

counsel admit higher sentences, but the coefficients on each type of lawyer are positive, so those results might owe more to a general leniency toward pro se defendants than anything else.

Perhaps more telling, it remains unclear why the distribution of evidentiary strength would differ in bench and jury trials. Table A1 shows that severity does not significantly impact the conviction probability for juries. Judges, of course, have experience in trials and sentencing outcomes, while jurors do not. The only task of a juror is to determine the facts. That severity is highly signficant for judges but not juries would seem to support the story of judges setting their burden of proof economically. For a more detailed account of the selection between bench and jury trials, see Appendix C.

If judges do behave like economic agents, why might the prediction of compromise verdicts fail to materialize? For one, the sample size might not allow for the identification of what is probably a small effect, if any. Furthermore, sentencing most frequently occurs on a delay after conviction. The decisions of the verdict and the sentence might simply be too remote in time for a judge to associate them. Lastly, if judges are happy with the guidelines range in most cases, then granting them greater discretion will only influence the outcome in a minority of cases. A USSC survey on district judges finds them to be generally happy with the guidelines ranges for most offenses (USSC 2010, Question 8). Again, the null result could arise from a lack of statistical power.

One important, unresolved question for future research is what drives the decision to choose a bench trial over a jury trial. Although a number of studies explore the decision in a civil trial (e.g., Helland and Tabarrok), the question remains open in the criminal context. Gay et al. (1989) present a game theoretic model in which judges and defendants behave strategically, juries behave naively, all innocent offenders opt for the bench, and guilty offenders mix between the bench and the jury. They also present stylized facts consistent with the equilibrium, specifically that judges will have

lower conviction probabilities than juries. Over the last two decades, federal judges do have lower conviction probabilities than juries (Leipold 2005). This paper offers a different account of the disparity in conviction rates. Judges convict less frequently because they are more familiar with sentencing than juries, and they think more often in terms of economic tradeoffs.

Conventional wisdom says the bench is preferred when the case is either too sensational or technical to present tor a jury. Given the high stakes of trial, one would expect the decision to be strategic, based on expected conviction probabilities or sentences, but Leipold (2005) interviewed twelve federal defense lawyers and found strategic considerations were muddled. For example, eleven of the twelve laywers wrongly believed juries had higher acquittal rates than judges. In a telephone survey, MacCoun and Tyler (1998) found juries to be perceived as more accurate and more representative of minorities than judges, but Rose et al. (2008) found partial evidence to the contrary. Hispanics who took a survey in Spanish preferred judges over juries, and blacks had a weaker preference for juries than whites.

Given the volume of trials going before the court, additional research on the choice of trial format is clearly warranted. Again, see Appendix C for a preliminary investigation.

Table 4.3: Two-Part Model Results – Federal Bench Trials

Dependent variables: Participation=conviction dummy
Continuous=log(sentence length)

	Conviction (Probit)		Sentence (OLS)	
	marginal effect	p-val	$\dfrac{\partial \mathrm{E}(ln(y)\|y > 0, x)}{\partial x_j}$	p-val
Severity	-.019	.012	.327	.000
Discretion Regime	.085	.046	-.187	.548
Age	-.013	.021	.038	.421
Age^2	.000	.013	-.001	.320
Black	-.050	.471	.132	.680
Male	-.024	.588	.319	.348
Black·Male	.049	.506	-.325	.388
Married	-.078	.041	-.113	.730
History Points	-	-	.135	.000
Public	.124	.072	.631	.124
Private	.123	.141	.260	.553
Appointed	.102	.216	.963	.005
%White Judge	-1.44	.049	-1.408	.833
%Female Judge	-.717	.172	-4.330	.285
%Democrat Judge	-.148	.366	2.671	.213

Note: n=765 for stage one; n=300 for stage two; years run from 1998-2010. The 'marginal effect' column reports average marginal effects. In the second stage, the average marginal effect is simply the coefficient in the OLS regression: $\partial \mathrm{E}(ln(y)|y > 0, x)/\partial x_j$ for continuous regressors and $\Delta_j \mathrm{E}(ln(y)|y > 0, x)$ for discrete ones. The regression contains unreported dummies for each US district. Standard errors are clustered by district. Convictions with a zero prison sentence are coded as a sentence of one-half month.

Chapter 5   Deliberation in Committees

The previous chapters viewed the fact finder in a trial as a single decision maker, but the group setting of juries involves communication between members of the group. This chapter develops a model in which members update their belief throughout a stream of evidence before voting on a decision. Individuals may or may not behave strategically.

At what point in time should you form an opinion of something? Conventional wisdom urges restraint. You shouldn't "judge a book by its cover" because hasty conclusions are prone to be less accurate than slow, careful ones. In many settings, individuals receive a stream of information before deciding whether or not to adopt a status quo. A reader commits to finishing a novel before choosing whether to recommend it to a friend. A juror sits through a trial before returning a verdict. A member of the Federal Open Market Committee reviews economic and financial developments before voting on interest rates. At any point in time, the individual is free to form judgment.

For a Bayesian individual, the timing of judgment is not relevant. A quick belief update is only a problem when subsequent, probative information is ignored. Otherwise, a Bayesian will come to the same conclusion whether she awaits all of the information before updating her beliefs or she updates them as each piece of information arrives. In other words, simultaneous and sequential information are identical to her. The equivalence is logically beautiful, and it follows from the basic rules of conditional probability.

That equivalence, however, depends on the ability to assign probabilities within a complete system of events. What if the information arrives in the form of a noisy

signal? Such a format can arise in either the receiver or the sender. On the receiver side, cognitive limitations might rule out a complete probabilistic description of an event system. Instead, the individual might collapse information into a binary signal indicating which of two options is the better. On the sender side, consider an expert who recommends one of two options. For example, a doctor condenses a complicated system of information into a recommendation of whether a patient should take a certain drug. Experts are perhaps more likely to be right than wrong, but they are not always right, so the signal is useful but noisy.

In this setting of binary signals, the timing of belief updates is material, even for a Bayesian. Surprisingly, when time has no opportunity cost, the best policy is not always to wait for all evidence before updating beliefs. With a symmetric loss function and a constant rate of information arrival, the individual should update beliefs early and often—what might appear as jumping to conclusions.

Many economic models analyze optimal stopping times when information acquisition is costly (e.g., Davis and Cairns 2012, Diamond 1971, Lizzeri and Yariv 2013, Wald 1947). This article inverts the question of an optimal stopping time. Given an information process with exogenous termination, at what point(s) in the process should an individual update beliefs? Learning is passive in that the only action is the binary choice at the end of the process. Contrary to the classical Bayesian case, identical information is interpreted differently when presented sequentially vs. simultaneously. The difference lies in the format of the information. Here, as in a large number of economic models, information arrives in a collapsed binary signal (e.g., Austen-Smith and Banks 1996, Feddersen and Pesendorfer 1998, Jackson and Tan 2013, Mukhopadhaya 2003, Persico 2004, to name just a few).

In the standard model, a Bayesian wants to identify the true state of the world, $\mathcal{A}$ or $\mathcal{B}$. She updates beliefs from a binary signal, equaling a or b, which satisfies $P(a|\mathcal{A}) > 1/2$ and $P(b|\mathcal{B}) > 1/2$. The current model relaxes the restriction on the

number of updates to one, allowing the agent a choice of when to convert information into an update. The key result is that an agent who frequently updates beliefs will tend to hold a more extreme posterior than an agent who infrequently updates beliefs. Whether a more extreme posterior is helpful depends on the loss function and the accuracy of the signal.

To build an intuition, consider a juror in a murder trial. The prosecution offers three pieces of evidence against the defendant: an eyewitness, a blunt object, and a confession. The defense rebuts the evidence, and no item is conclusive by itself. The eyewitness places the defendant near the scene of the crime, but it was dark outside and difficult to see. The blunt object contains traces of the victim's blood, but it also contains several sets of partially smudged fingerprints. The confession is recorded on videotape, but it was obtained after ten hours of interrogation, and the defendant was visibly confused.

If the juror receives a binary signal based on the evidence, the probability of receiving a signal for either guilt or innocence must somehow depend on the strength of the evidence supporting each state of the world. The more the evidence supports guilt, the more likely the juror should be to receive a guilty signal. Suppose each piece of evidence 80% supports guilt, and the evidence taken altogether 80% supports guilt. The word "support" is left deliberately vague for now; Section III speaks to the interpretation in detail. For simplicity, let the probability of receiving either a guilty or innocent signal equal the "strength" of the evidence in favor of that state of the world—80% for guilt and 20% for innocence.[1]

The juror has two options. She can wait through the trial to form a single update, as in Case 1 of Figure 5.1, or she can update beliefs as each piece of evidence is presented, as in Case 2. Suppose the juror makes no prejudgment on whether the

---

[1]The purpose of the example is purely informational. Whether every *statutory* element of a crime should be proved beyond reasonable doubt, or whether only the evidence taken as a whole must meet that criterion is a separate question. The Supreme Court has ruled for the former on the foundation of *In re Winship*, 397 U.S. 358 (1970).

defendant committed the crime, so her priors are uninformative.

Case 1. If the juror receives one signal, then she updates beliefs via Bayes' rule,

$$P(\mathcal{G}|s) = \frac{P(s|\mathcal{G})P(\mathcal{G})}{P(s|\mathcal{G})P(\mathcal{G}) + P(s|\mathcal{I})P(\mathcal{I})}, \tag{5.1}$$

where $\mathcal{G}$ and $\mathcal{I}$ denote the events of guilt and innocence, and s is the received signal. She will come to believe in guilt with probability 0.8 and innocence with probability 0.2. With equal priors, her posterior in either case will be 0.8 in favor of the received signal.

Case 2. Suppose the juror updates her beliefs after each piece of evidence is presented. If so, she forms a posterior belief through (1) with not one but three independent signals. According to the binomial formula, with probability 0.512 (0.008) she receives three guilty (innocent) signals, and with probability 0.384 (0.096) she receives two guilty (innocent) signals and one innocent (guilty) signal. Her probability of coming to believe in guilt is no longer 0.8 but 0.896, and she will tend to be more extreme in her beliefs. Her posterior will again be 0.8 if she receives two guilty signals but will be 0.985 if she receives three guilty signals.

Figure 5.1: Simultaneous vs. Sequential Updating
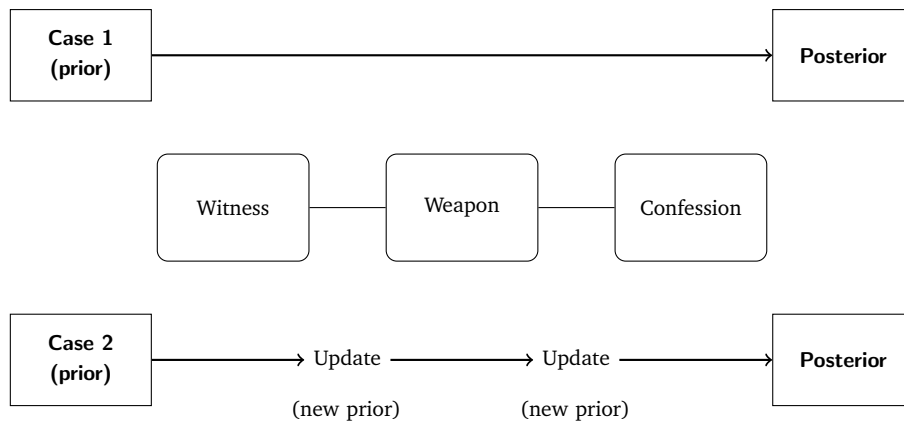
The two scenarios are logically identical. Splitting the overall evidence into its constituent parts changes nothing about the information. Yet the juror can alter her

behavior with the timing of her updates. Surprisingly enough, if her only objective is to determine the true state of the world, she is better off taking the multiple update position, since then she becomes more likely to choose the side with more supporting evidence. (This assumes, of course, a positive correlation between the majority of the evidence and the true state of the world.) For a relatively asymmetric loss function, however, she should await all evidence before forming an update. For example, suppose the "reasonable doubt" standard in a criminal trial is numerically interpreted as a 95% confidence in guilt. Then the juror should not vote to convict because the evidence only 80% supports guilt. With a single update, she never does vote to convict, but with a triple update, she does if she receives three guilty signals, an event with probability 0.512 (recall that her posterior is 0.985 in that case).

The example can easily be extended through additional updates. In the limit, the juror becomes arbitrarily confident in guilt. Intuitively, since a guilty signal is more likely than an innocent signal, it becomes more and more likely that the number of guilty signals exceeds the number of innocent signals by any given threshold, yielding an arbitrarily strong posterior.[2] Because of this potential for distortion, a Bayesian with a relatively asymmetric loss function is well advised to wait for all of the evidence before making any judgment.

To summarize the main results, when the rate of information arrival varies widely, the safe option is to hold off on judgment; otherwise, beliefs might be updated on a part of the information not representative of the whole. When information arrives at a roughly constant rate, an individual becomes increasingly overconfident in a belief as she updates it more frequently. Whether she should be patient or quick to form judgment then depends on the relationship between her loss function and the accuracy of the information signal. Loosely speaking, if the costs of a Type I and Type II error are close—e.g., if her only goal is to determine the true state of the world—then she

---

[2]Three signals are chosen over two in the example because the signal probability sequence is non-monotonic.

does best by updating beliefs quickly. If the costs of error are asymmetric, she does best by waiting for the evidence before updating beliefs.

In many applicable settings, the individual is in fact a member of a committee. Existing work typically assumes that each member looks at the evidence at hand to form a single Bayesian update. Members then vote for an alternative based on their posteriors. Again, the assumption of a single update is restrictive. The main results extend naturally to the group setting, with the caveat that previous strategic voting equilibria are not always robust to additional updates.

The potential for overconfidence has clear implications for the optimal timing of committee deliberation. To continue with the courtroom example, jurors are typically prohibited from discussing a trial during breaks. In a criminal trial, the prohibition makes sense because the loss function is asymmetric. A conviction requires proof beyond reasonable doubt. If jurors deliberate and update beliefs during trial breaks, they risk an overconfident conviction, even in the absence of confirmation bias. In a civil trial, however, the loss function is perfectly symmetric. The winning party is decided by the preponderance of the evidence. If that standard is in fact optimal, the prohibition is counterproductive. Jurors would be more likely to choose the correct side with additional updates.[3]

The next section describes the related literature, while section III presents the basic model. Sections IV and V apply the model to the Bayesian individual and group settings, respectively. Section VI explores the implications of an alternative, moving average model of learning, and Section VII describes the incentives of opponents in a debate facing an audience of either learning model. Section VIII concludes. Proofs are left to the Appendix.

---

[3]Though overconfidence might skew any damage awards. Also, Kaplow (2012) argues that the preponderance standard is misguided.

5.1 Related Literature

This article fits within two lines of economic literature. First, economic models of learning set the background for the incorporation of sequential information into beliefs. Bayesian learning in particular ocassions an acknowledgment of the psychological work on the topic. Second, committee voting inspires a natural extension of the model. This section offers a brief review of the relevant literature.

5.1.1 Models of Learning

Several recent models in economics offer a variation on the classical Bayesian agent. This paper fits best with Ortoleva (2012), Rabin and Schrag (1999), Schwartzstein (2014), and Wilson (2014), in that each examines passive learning under cognitive limitations. Again, a behavioral assumption is not required in the current model. The sender might be the one collapsing the message into a binary signal. Other work examines learning in a more active environment, where an action is taken in multiple stages (e.g., Lehrer and Smorodinsky 2000), as in repeated games (see Sobel 2000 for an early survey). Here the only game is the vote in the committee setting, so no individual has the opportunity to learn from any action aside from votes.

Inattention is a potential source of error in the model, but unlike the work of Sims (2003) or Caplin and Dean (2015), attention is not derived from a direct constraint on cognition. The only explicit cost of inattention is in reducing the quality of the ultimate decision. Furthermore, since the stopping rule is exogenous, the opportunity cost of continuing to review information is assumed zero, unlike classic stopping rule models following Wald (1947).

Lastly, Section VI presents a moving average model inspired by Hastie (1993) and Pennington and Hastie (1992), though Cross (1973) introduces a similar model where the average moves over probabilities instead of states.

5.1.2 Bayesian Cognition

A large body of psychological evidence suggests that people integrate information in a Bayesian fashion, at least at the "computational" level (e.g., Gigerenzer and Hoffrage 1995, Jacobs and Kruschke 2010, Tenenbaum et al. 2011). Applications range from sensory inputs to higher cognition.[4] This paper takes for granted some degree of Bayesian thought and provides a model in which, contrary to the classic Bayesian result, the timing of updates is critical. The finding is not without precedent. The legal literature has previously noted the possibility of a differential effect of sequential vs. simultaneous information on Bayesian beliefs (Schum and Martin 1982), and Hoffman et al. (2011) find experimental evidence in support of the idea.

5.1.3 Committee Voting

The primary strand of related work on committee decision making begins with Austen-Smith and Banks (1996), who, using a version of the basic model below, show that for a binary group decision made by simultaneous vote, it is not always a Nash equilibrium for members to vote in line with their own information. Feddersen and Pesendorfer (1998) conclude that unanimity is an especially poor voting rule in this context, but Coughlan (2000) defends the rule with two extensions of the model. Feddersen and Austen-Smith (2006) show more generally that uncertainty about private, individual preferences is necessary for full information sharing, and if there exists a truth-revealing equilibrium under unanimous rule, there does for any voting rule (but the converse is not true). See Gerling et al. (2005) for an early survey of information aggregation in committee decision making.

Gerardi and Yariv (2007, 2008) approach the topic from the perspective of mechanism design, showing the equivalence of different voting rules when members deliber-

---

[4]However, Bowers and Davids (2012) criticize the models for effectively retrofitting data via free parameters, and Marcus and Davis (2013) echo the sentiment (but see Griffiths et al. 2012 for a rebuttal of the critique).

ate before casting votes. Jackson and Tan (2013) find that when a committee consults experts who can conceal information but cannot lie, the maximization of information aggregation and total utility share no necessary relationship.

While strategic voting and information aggregation are natural foci, other work examines whether specialization is desirable in committee decision making (Ben-Yasher et al. 2012), how restricting communication protocol can improve group decisions (Schulte 2012), when a deliberating group will reach a consensus (DeGroot 1974, Neilson and Winter 2008), and when that conensus is accurate (Golub and Jackson 2010). Persico (2004) and Mukhopadhaya (2003) examine the effect of voting rules and committee size on endogenous information collection. Those factors can be important in the group context but are omitted here to focus on the main result.

5.2 The Basic Model

An individual plans to evaluate a continuous stream of evidence over a finite period with exogenous stopping time. After her evaluation, she decides which of two alternatives to adopt, C or A. The state of the world $s \in \{\mathcal{G}, \mathcal{I}\}$ is unknown, and the individual wants to choose the alternative that corresponds to the true state—C in state $\mathcal{G}$ and A in state $\mathcal{I}$. Following the convention in the committee voting literature, the notation reflects the running example of a criminal trial, where the defendant is either guilty ($\mathcal{G}$) or innocent ($\mathcal{I}$), and the juror decides whether to convict (C) or acquit (A), but the model applies to any choice of whether to adopt a status quo.

The juror has preferences $u(A; \mathcal{I}) = u(C; \mathcal{G}) = 0$, $u(C; \mathcal{I}) = -\bar{p}$, and $u(A; \mathcal{G}) = -(1 - \bar{p})$ for $\bar{p} \in (0, 1)$. The parameter $\bar{p}$ captures the cost of a wrongful conviction relative to the cost of a false acquittal. It also defines a threshold of proof. As an expected utility maximizer, the juror prefers conviction if and only if she believes the defendant to be guilty with probability higher than $\bar{p}$.

The sum of incoming evidence is normalized to unity. At any moment in time a

Figure 5.2: Stream of Evidence
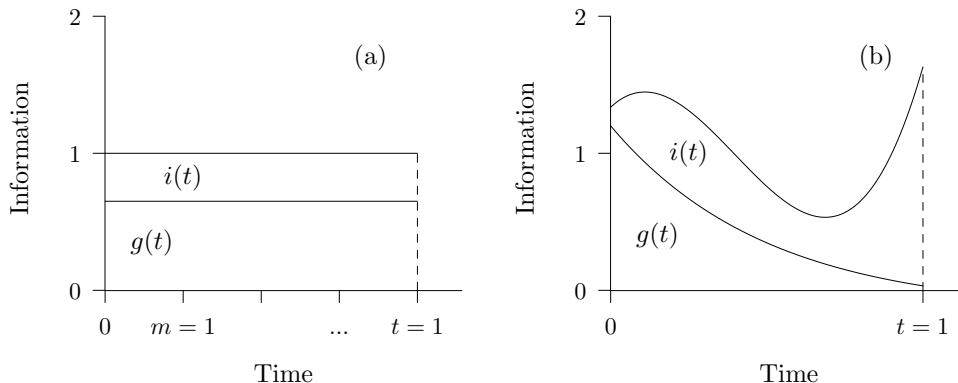
portion $g(t)$ will support state $\mathcal{G}$, and the rest, $i(t)$, will support state $\mathcal{I}$. Any extraneous information is filtered out. As pictured in panel (b) of Figure 5.2, information need not arrive at a constant rate, but the primary focus will be the case where it does, as in panel (a). Aside from being a natural benchmark, a constant rate of information arrival can be motivated as a Nash equilibrium where opposing debaters choose when to present their evidence in order to convince a Bayesian audience of their position; see Section VII.

Although the juror always forms a belief after the final evidence, she can also update beliefs earlier if desired. Let k denote the number of updates, equidistant, over the time period. When $k = 1$, the juror waits until all of the information has arrived before passing judgment. When $k \geq 2$, she updates beliefs at various points in time, at points $m \in \{1, 2, ..., k\}$.

At each point m, the juror receives either a guilty or innocent signal. Denote the overall proportion of guilty information by $\lambda \in (0, 1)$, and define λ_m as the proportion of guilty evidence arriving between the time of the last update (or prior) and point m. Intuitively, the probability of receiving a guilty signal, p_m, at point m should be an increasing function of λ_m,

$$p_m = f(\lambda_m) = f\left(\frac{g_m}{g_m + i_m}\right), \text{ where } g_m = \int_{m-1}^{m} g(t)dt, \ i_m = \int_{m-1}^{m} i(t)dt,$$

and $m = 0$ is understood to mean time $t = 0$. The assumption of an increasing $f(\cdot)$ is a fairly agnostic but commonsense property of a binary signal. The more the evidence favors guilt, the more likely the juror should be to receive a guilty signal. The function also satisfies

$$f(\lambda_m) \in (0, 1), \text{ with } f(1/2) = 1/2.$$

The information process is not referenced in the standard model. Typically, the juror is assumed to receive one signal, either guilty (gs) or innocent (is), satisfying $P(gs|\mathcal{G}) = p > 1/2$ and $P(is|\mathcal{I}) = q > 1/2$, where $p = q$ is often assumed for algebraic convenience. The standard model can be motivated by the current framework. Abstracting from incentives in the creation of information, suppose λ is known to be distributed via $F_{\mathcal{G}}(\lambda)$ when the defendant is guilty and $F_{\mathcal{I}}(\lambda)$ when he is innocent. The former distribution first-order stochastically dominates the latter, making higher values of λ more likely when the true state is guilt. If $k = 1$,

$$p = P(gs|\mathcal{G}) = \int_0^1 f(\lambda)dF_{\mathcal{G}}(\lambda), \text{ and } q = P(is|\mathcal{I}) = \int_0^1 (1 - f(\lambda))dF_{\mathcal{I}}(\lambda).$$

A sufficient and fairly natural condition for $p, q > 1/2$ is $F_{\mathcal{G}}(1/2) < 1/2 < F_{\mathcal{I}}(1/2)$. If information arrives at a constant rate, the decision is again whether to divide one signal into many.

For the signal to be "informative," i.e., positively correlated with the true state of the world, the preponderance of the evidence should favor the true state on average. Let $\bar{\lambda}$ equal the overal ratio of theoretically available evidence. For simplicity, the following assumption is maintained going forward:

Assumption 4. $P(\mathcal{G}|\lambda) = P(\mathcal{G}|\bar{\lambda}) = \bar{\lambda}$.

In words, the conditional probabilities of guilt for the observed evidence and for all theoretical evidence are both equal to the ratio of theoretical evidence favoring guilt, which makes the observed λ interchangeable with $\bar{\lambda}$ for decision purposes. Since λ is generated by a random process, the assumption is clearly stylized, and Section IV.2 discusses its relaxation. One possibility is to let $E(\lambda) = \bar{\lambda}$, where the distribution of λ maintains the decision-value equivalence of λ and $\bar{\lambda}$. In any case, the only use of Assumption 4 is to simplify optimality results.

5.2.1 Interpretation of Evidence

How the evidence "supports" one state of the world or another is open to interpretation. The simplest explanation is that some portion of the inherent quality of a piece of evidence favors a given state. For example, a literal eight out of ten findings in a report on the economy might favor a healthy real sector, while the remaining two indicate weakness, yielding a $\bar{\lambda} = 0.8$ in favor of a strong economy. Under a more satisfactory probabilistic interpretation, if $\bar{\lambda} = 0.8$, then 80% of the time when such information is presented, the real sector of the economy is in fact healthy. Assumption 4 captures the latter interpretation.

If $\bar{\lambda}$ represents a conditional probability of the true state, a constant information stream could mean the juror is viewing a sequence of events, E_i, satisfying the property in the following claim.

Claim 6. For any $x \in (0, 1)$ and $n \in \mathbb{Z}^{++}$, $\exists$ events $\mathcal{G}$ and E_i such that $P(\mathcal{G}|E_i) = x \; \forall i \in \{1, 2, ..., n\}$ and $P(\mathcal{G}|E_1, ..., E_k) = x \; \forall k \in \{2, ..., n\}$.

Events that satisfy Claim 1 may or may not be realistic, but they exemplify the difference between sequential and simultaneous information for a Bayesian receiving signals. While E_2 provides additional information over E_1, it does not provide any

extra value to the ultimate decision. Such events show that while an early update is based on less information than a later update, that information gap is not what is driving the overconfidence in the multiple update position. The key factor is the timing of the update.

The conversion of information into a signal is what generates the difference in sequential vs. simultaneous information. If the juror knew the probability system behind n events, her posterior could be calculated via the odds-form of Bayes' rule by

$$\Phi_1 = \frac{P(\cap E_i|\mathcal{G})}{P(\cap E_i|\mathcal{I})} \frac{P(\mathcal{G})}{P(\mathcal{I})}$$

for the simultaneous update (where Φ_1 stands for the posterior odds in Case 1 from the Introduction), and

$$\Phi_2 = \frac{P(E_n|\mathcal{G}, E_{n-1}, ..., E_1)}{P(E_n|\mathcal{I}, E_{n-1}, ..., E_1)} \frac{P(E_{n-1}|\mathcal{G}, E_{n-2}, ..., E_1)}{P(E_{n-1}|\mathcal{I}, E_{n-2}, ..., E_1)} \cdots \frac{P(E_1|\mathcal{G})}{P(E_1|\mathcal{I})} \frac{P(\mathcal{G})}{P(\mathcal{I})}$$

for the sequential update (Case 2). Since

$$P(\cap E_i|\mathcal{J}) = P(E_n|\mathcal{J}, E_{n-1}, ..., E_1)P(E_{n-1}|\mathcal{J}, E_{n-2}, ..., E_1)...P(E_1|\mathcal{J}), \ \ \mathcal{J} = \mathcal{G}, \mathcal{I},$$

$\Phi_1 = \Phi_2$ generally.

Why should an individual receive a noisy signal when, if she knew the complete probability system behind the events, she could calculate the exact posterior odds? The committee voting literature is mainly quiet on the question, but many factors influence the perception of evidence. Retention will be incomplete because memory is imperfect. Some information will never be processed because individuals are either innattentive or lack the technical ability to comprehend it. All information requires interpretation, which is subject to personal biases. Anecdotally, people tend not to carry probabilities in their mind but use them to draw conclusions about the world.

As a modeling device, then, perhaps individuals are better viewed as receivers of signals than as perfect trackers of information ratios.

For similar reasons, the signal will not always match the side with the preponderance of the evidence. For example, in experimental studies, events people believed were certain to happen only did happen 80% of the time; likewise, events considered impossible happened about 20% of the time (Fischhoff et al. 1977). In simple written and televised communications, Jacoby and Hoyer (1982, 1989) find the average rate of miscomprehension to be within 20-30%.

5.3 Bayesian Updaters

Consider an individual who updates her belief at k points in time.

5.3.1 Known λ

Assume the juror updates beliefs from signals according to Bayes' rule, starting with priors of $p(\mathcal{G}) = p(\mathcal{I})$.[5] Depending on the timing of the evidence, a single additional update, $k = 2$, can have a large effect on the posterior belief. For any overall ratio of evidence, a juror who updates her belief just once throughout the stream of evidence can come to hold any posterior belief depending on the timing of the evidence. For example, even if the total ratio of guilty to innocent evidence is 0.90, by adjusting the timing of the evidence, a juror can arrive at a posterior belief of 0.99 in favor of innocence.

Proposition 6: For linear $f(\cdot)$, take any positive sums of evidence, $g = \int_0^1 g(t)dt$ and $i = \int_0^1 i(t)dt$.[6] For any $k > 1$ and $q \in (0, \infty)$, there exists an ordering of the evidence

---

[5]The assumption of equal priors, or that individuals begin their evaluation with an open mind, is not needed but merely simplifies the results.

[6]Whenever $f(\cdot)$ is linear, events with probability ε-close to zero and one are implicitly ignored. Furthermore, the function should actually be affine, and the abuse of the term linearity is only to

and signal sequence such that for a Bayesian individual the posterior odds in favor of guilt, $\Lambda_g \equiv P(\mathcal{G}|k)/P(\mathcal{I}|k) = q$.

Proposition 6 is dramatic but should not be overstated—the more extreme the posterior beliefs, the less likely are the signals generating them to occur.

Suppose now that information arrival is constant, as in panel (a) of Figure 1, which implies that $p_m = p$ $\forall m$. After receiving n_{gs} guilty signals and n_{is} innocent signals, the juror believes the defendant is guilty with probability

$$P(\mathcal{G}|n_{gs}, n_{is}) = \frac{p^{n_{gs}-n_{is}}}{p^{n_{gs}-n_{is}} + (1-p)^{n_{gs}-n_{is}}}. \tag{5.2}$$

First, for illustration, suppose that $p > \bar{p}$. In this case, if the juror receives at least one more guilty than innocent signal, she will vote to convict. If $k = 1$, she votes to convict if and only if the one signal is guilty. For a general k, the probability that she convicts is

$$\sum_{j=[k/2]+1}^{k} b(j; k, p) = 1 - \mathrm{F}([k/2]; k, p),$$

where $b(j; k, p)$ denotes the probability of j successes in k trials of a binomial distribution with parameter p, $\mathrm{F}(\cdot)$ denotes the cumulative distribution, and $[\cdot]$ is the integer part. Intuition suggests that since a guilty signal is more likely than an innocent signal, then as the number of updates grows, the chance of having at least one more guilty than innocent signal also grows. The juror should end up convicting with near certainty if she updates his beliefs enough times. In fact, this result does hold, and it holds for any threshold of proof $\bar{p}$.

Proposition 7: WLOG suppose $g(t) = g > i(t) = i$ $\forall t \in [0, 1]$. Let $p_c \in (0, 1)$ denote the posterior belief in state $\mathcal{G}$. Then for $\forall \bar{p} \in (0, 1)$, $\lim_{k\to\infty} P(p_c > \bar{p}) = 1$.

---

keep algebra clean. No result depends on it.

If the arrival of information is constant, then even if the total evidence just barely favors one state of the world, an individual who updates frequently enough will come to believe strongly in that state. Proposition 7 is a restatement of the well-known result that a Bayesian agent with enough informative signals will come to hold a correct belief with arbitrary strength, but here the agent *should not* retain such strong beliefs. A patient, objective juror who reserves judgment until the simultaneous presentation of all the facts (by setting $k = 1$) tends to have a weaker posterior. The juror who quickly updates her belief becomes overconfident. The following corollary of Proposition 7 expresses the danger.

Corollary 7.1: Let $g(t) = g > i(t) = i \ \forall t \in [0, 1]$. Then $\lim_{k \to \infty} P(\text{C}) = 1$.

With frequent enough updates, a juror will become so confident in her belief of guilt that she will vote to convict regardless of how strict her threshold of proof may be. If the information marginally favored innocence instead of guilt, the same would be true of acquittal.

The juror votes to convict if and only if her posterior belief in guilt is greater than $\bar{p}$, but, under Assumption 1, she *wants* to convict whenever the ratio of total guilty evidence, $\bar{\lambda}$, is greater than $\bar{p}$. The difference between her action and her desire comes from the uncertainty in the signal. Surprisingly, overconfidence is not always harmful. In fact, it helps the juror when her loss function is relatively symmetric.

Proposition 8: WLOG let $g(t) = g > i(t) = i \ \forall t \in [0, 1]$, and denote by k^* the optimal number of updates for a Bayesian juror. If $\bar{p} > \bar{\lambda}$, then $k^* = 1$. If $\bar{p} < \bar{\lambda}$, then $k^* \to \infty$.

Proposition 8 follows from Corollary 1, Assumption 4, and the utility function of the

juror. If $\bar{p} > \bar{\lambda}$, the juror does not want to vote to convict, but overconfidence might lead her to convict. Therefore, $k^* = 1$.[7] On the other hand, if $\bar{p} < \bar{\lambda}$, the juror wants to convict. For any given k, she may or may not vote to convict depending on the signals she receives, but as long as k is great enough, the probability of conviction is arbitrarily close to one when $\bar{\lambda} > 1/2$. A similar proposition holds for the case where $g(t) = g < i(t) = i \ \forall t \in [0,1]$. As a rule of thumb, a juror with either a stringent or lax threshold of proof should wait for all evidence before updating her belief, but a juror with an intermediate threshold should start forming a judgment right away.

The foregoing is an asymptotic result. If information arrival is approximately constant, and overconfidence develops very slowly, the timing of updates would be unimportant. Furthermore, while the continuity of information is a convenient modeling device, evidence is generally discrete—it makes little sense to subpoena 1.67 of a witness or present 0.932 of a DNA test. In practice, the number of potential updates is bounded from above. How fast does overconfidence occur?

Figure 5.3: Overconfidence – Speed of Convergence

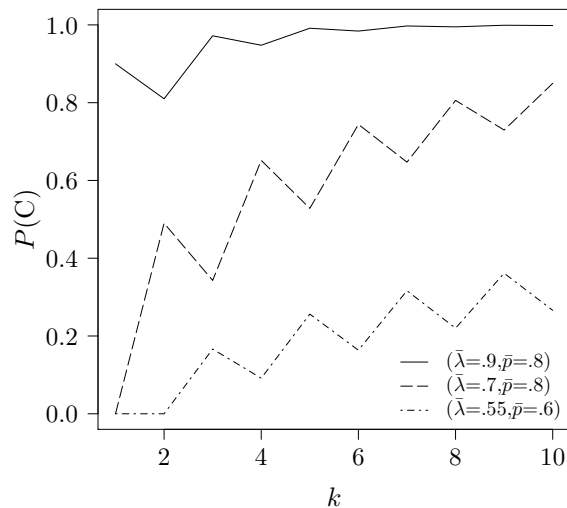Figure 5.3 shows the probability that a Bayesian juror will come to hold a posterior

---

[7]Proposition 3 actually provides for a subset of k^* when the loss function is relatively asymmetric. For example, with $\bar{p}$ much larger than $\bar{\lambda}$, a juror may not be willing to convict with two guilty signals but would with three. Since she does not want to convict when $\bar{p} > \bar{\lambda}$, k^* can be 1 or 2.

belief exceeding her threshold of reasonable doubt for a given number of updates k, assuming $f(\cdot)$ is the identity function. Overconfidence is a serious risk even for a small number of updates. For example, if the total ratio of guilty to innocent information is 0.9, and the juror has a threshold of $\bar{p} = 0.8$, the probability she will hold a belief stronger than her threshold after four updates, and thus vote to convict, is 0.95. Of course, overconfidence is a benefit in this case because the juror is more likely to make the right decision. However, when $\bar{p} = 0.8$ but the total ratio of guilty information is 0.7, the juror who updates one extra time ($k = 2$) votes to convict 49% of the time, when she never would if she had waited for all the information to form her beliefs.

Interestingly, frequent updates yield a kind of analog to confirmation bias without any path-dependent misinterpretation of evidence. Confirmation bias, the tendency to disproportionately find information in support of a currently held belief, is a pervasive phenomenon (Nickerson 1998). Rabin and Schrag (1999) develop a model in which an agent who comes to hold a belief occasionally misinterprets future disconfirmatory signals as evidence in favor of the belief; the agent then becomes overconfident in any position held for a long enough time. In the current model, the juror always holds the correct belief asymptotically, but an interesting question for future work is how likely a juror with confirmation bias is to hold the wrong belief in the long run.

As another important cognitive bias, Möbius et al. (2014) and Hoffman et al. (2011) find evidence of incomplete updating in experiments. Subjects moved their beliefs in the correct direction but with a smaller magnitude than prescribed by Bayes' rule. If people dampen their updates, then overconfidence is less of a risk.

5.3.2 Unknown λ

The results above hold λ constant, but a juror is not likely to know λ ahead of time. Furthermore, the optimal update policy is always one of two extremes. The juror should either update only once or as many times as possible. Would an

intermediate policy ever be appropriate?

For any $\bar{p}$, $\bar{\lambda}$, and priors, $\exists \ n_{\bar{p},\bar{\lambda}} \in \mathbb{Z}$ s.t. the juror believes in guilt enough to convict iff $\Delta_n \equiv n_{gs} - n_{is} > n_{\bar{p},\bar{\lambda}}$. (Note that $n_{\bar{p},\bar{\lambda}}$ bears no sign restriction and the dependence on the priors is suppressed by the notation.) Given k updates and constant information arrival, the juror has expected utility

$$EU_k = -\bar{p} \cdot P(C; \mathcal{I}|k) - (1-\bar{p}) \cdot P(A; \mathcal{G}|k).$$

Using Assumption 4, the expression can be simplified to

$$
\begin{aligned}
EU_k &= -\bar{p}[1 - \mathrm{F}(n_{\bar{p},\bar{\lambda}}; k, f(\lambda))](1-\lambda) - (1-\bar{p})[\mathrm{F}(n_{\bar{p},\bar{\lambda}}; k, f(\lambda))](\lambda) \\
&= -\bar{p}(1-\lambda) + (\bar{p}-\lambda)\mathrm{F}(n_{\bar{p},\bar{\lambda}}; k, f(\lambda)).
\end{aligned}
\tag{5.3}
$$

Equation (3) poses the algebra behind Proposition 8. The first term does not depend on k. When $\bar{p} > \lambda > 1/2$, the second term is maximized by choosing small k, since a small k yields a large $\mathrm{F}(\cdot)$. When $\lambda > 1/2$ and $\lambda > \bar{p}$, the second term is maximized by choosing large k, since a large k yields a small $\mathrm{F}(\cdot)$. Similar reasoning holds when $\lambda < 1/2$.

Suppose the juror knows only that λ is distributed via distribution $H(\lambda)$, which is derived from the mixture of $F_{\mathcal{G}}(\cdot)$ and $F_{\mathcal{I}}(\cdot)$. Dropping the first term in (3), her expected utility as a function of k is then

$$EU_k = \int_0^1 (\bar{p}-\lambda)\mathrm{F}(n_{\bar{p},\bar{\lambda}}; k, f(\lambda))dH(\lambda). \tag{5.4}$$

Although its maximization over k does not yield a clean analytical solution, Expression (4) shows why an intermediate number of updates can be optimal. For example, suppose all priors are uninformative, meaning that $H(\cdot)$ is uniform, and $f(\cdot)$ is the identity function. For $\bar{p} = 0.7$, $k^* = 3$ is the number of updates that balances the risk

and benefit of overconfidence across the possible values of λ when a Type I error is 0.7/0.3=2.33 times worse than a Type II error.

5.4 Group Voting

The literature on committee voting typically adopts the model above, assuming that $k = 1$, that each member is Bayesian, and that each member receives a single, informative signal upon receipt of the final evidence.

The results for the individual extend directly to the group setting if voting is sincere, meaning that each member votes to convict if and only if her own posterior belief of guilt is higher than her threshold of proof. While an individual might have the restraint to wait for all the facts before judging them, a deliberating group is unlikely to be so disciplined. In fact, the very purpose of deliberating throughout a stream of evidence would seem to include making some type of judgment. A simple application of Proposition 2 shows that if the rate of information arrival is constant, then a Bayesian group will become overconfident in its beliefs very quickly.[8] If the group conducts a straw poll at each meeting prior to the last, additional meetings are analytically identical to additional signals. Each time a committee deliberates, each member receives an additional $n - 1$ signals that would otherwise be absent.

Proposition 9: WLOG let $\lambda_m = \lambda > 1/2 \ \forall m \in \{1, 2, ..., k\}$. Consider a committee deliberating at points L $\subseteq \{1, 2, ..., k\}$ with $k \in$ L. If members are Bayesian, and if their standard of proof is relatively high $(\bar{\lambda} < \bar{p})$, a committee that deliberates early (L $\supset \{k\}$) will make worse decisions on average than one that deliberates only after

---

[8]This statement is true of a group subject to what Glaeser and Sunstein (2009) call "credulous Bayesianism," where group members fail to account for their common source of information and, as a result, put too much weight on each others' opinions. A group that accounts for the common information should move to overconfidence less quickly. However, Bénabou (2013) shows how willful blindness and wishful thinking can be individually rational in a group setting when agents have anticipatory preferences.

receiving all evidence (L = $\{k\}$). If members have a low standard of proof ($\bar{\lambda} > \bar{p}$), then a committee will tend to make better decisions when deliberating early.

The optimal number of meetings also depends on the opportunity cost of additional meetings, which is assumed zero here. Notice that Proposition 9 assumes the rate of information arrival is proportionally constant, or that two sides in a debate employ their equilibrium strategies (as shown in the next section). If the task at hand is to evaluate a natural process, and evidence arrives at variable rates, then a committee may still be wise to await the evidence before meeting.

Thus far, voting has been assumed sincere, but in some environments voting may be strategic. Austen-Smith and Banks (1996) note that, depending on $\bar{p}$, sincere voting is not a Nash equilibrium for Bayesian jurors. The result is initially surprising because jurors are assumed to have identical preferences, but the basic intuition is clear. Each juror bases his strategy on the situation where his vote is pivotal, i.e., where the group decision turns on his vote. Suppose a juror would like to convict if all but one of the signals are guilty (or if they are all guilty). Under a unanimous rule, his vote is pivotal only when every other member votes guilty. But if all the other members are voting sincerely, they all must have received guilty signals, and the juror would prefer to vote guilty even if he received an innocent signal. Therefore, sincere voting cannot be an equilibrium. Feddersen and Pesendorfer (1998) extend the result to show that unanimity is a uniquely bad voting rule in the strategic setting.

Just one early update can have a large impact on strategic equilibria. As an example, consider the case where jurors each update their own beliefs independently, and then meet to cast simultaneous votes upon receipt of the final evidence. Jurors vote strategically under a unanimous rule, as in Feddersen and Pesendorfer (1998). With one extra update, there are two symmetric, responsive equilibria.[9] In the first,

---

[9]Following Feddersen and Pesendorfer (1998), an equilibrium is "responsive" if a juror conditions his voting on his signal. A non-responsive equilibrium exists under any voting rule. For example,

jurors with two innocent signals always vote acquit, jurors with one guilty and one innocent signal mix, and jurors with two guilty signals always vote to acquit. In the second, jurors with two innocent signals mix on voting and all other jurors always vote guilty. See Appendix B for details.

Table 5.1: Required Burdens of Proof for Strategic Equilibria

	p=.55	.65	.75	.85	.95
Equilibrium 1	0.9999	1	1	1	1
Equilibrium 2	0.9881	0.9841	0.9708	0.9240	0.7307

The table contains minimum thresholds of proof, $\bar{p}$, for a strategic equilibrium to exist for different true ratios of evidence, p, given $k = 2$. Equilibriums 1 and 2 refer to the cases where jurors with one of each signal mix and jurors with two innocent signals mix, respectively. See Appendix B for details. The jury size is n=12.

Table 5.1 shows the parameters required for strategic equilibrium to be empirically suspect. For equilibrium to exist, the required threshold of proof must be very high. Even taking the less extreme—but certainly less natural—second equilibrium, an overall information ratio of 0.75 requires a minimum $\bar{p} = 0.97$. The strategic equilibrium is evidently not robust to additional updates. For example, even in the beyond-reasonable-doubt criminal trial context, Horowitz (1997) summarizes the empirical estimates of proof beyond reasonable doubt to range from approximately 0.8 to 0.9 (see also Dhami 2008). In many other situations, $\bar{p}$ should be much lower yet.

5.5 Moving Average Updaters

The Bayesian model is the standard in economics for its logical foundation and its normative beauty. Unfortunately, empirical failures of the model are numerous, and individuals frequently follow alternative modes of cognition (cf. Kahneman, Slovic, and Tversky 1982, Hilbert 2012). In the context of sequential learning, Pennington

---

under a unanimous rule, always voting to acquit is an equilibrium strategy since a juror cannot affect the outcome by changing his vote.

and Hastie (1992) find a moving average model to be empirically preferable. In that model, the juror begins with an initial strength of belief $J_0 \in [0, 1]$. The upper and lower thresholds correspond to full confidence in guilt and innocence, respectively. At each update the juror forms a new strength of belief that is a weighted average of the prior belief and the current signal,

$$J_m = wJ_{m-1} + (1 - w)s_m, \text{ for } w \in (0, 1), \ m \in \mathbb{Z}^+.$$

The signal s_m could equal λ_m if the juror perfectly tracks information, but for a fair comparison with the Bayesian model, assume the signal is either guilty or innocent, where a guilty (innocent) signal maps to one (zero) on the strength-of-belief scale. After the final update, the juror compares J_k to a threshold $\bar{J}$. If $J_k \geq \bar{J}$, he prefers to convict; otherwise, he prefers to acquit. Mathematically, the primary difference between the moving average and Bayesian models is that in the former updates are additive while in the latter they are multiplicative.

One advantage of the moving average model is that a result analogous to Proposition 1 does not hold.[10] A single additional update cannot lead to arbitrarily extreme beliefs because the weight on the initial belief constrains the update. Consider that

$$J_2 = J_0 w^2 + w(1 - w)\frac{g_1}{g_1 + i_1} + (1 - w)\frac{g_2}{g_2 + i_2}.$$

Maximizing J_2 with respect to the order of evidence requires numerical methods, but, for example, if $J_0 = .5$, $w = .6$, and $g = .7$, the maximum possible value of $J_2 \approx 0.75$. Another difference between the models is that the order of evidence matters independently of its effect on the signal probabilities, which, depending on the perspective, can either be an advantage or disadvantage. For example, contrary to

---

[10]Another advantage is that beliefs are not absorbing at zero and one, as they are for a Bayesian model. Situations of certainty are ignored here, however.

the Bayesian model, a weak witness followed by a strong one tends to yield different beliefs than a strong one followed by a weak one. The advantage is that such order effects are empirically observed. The disadvantage is that, normatively, the order of testimony should not matter.

Most importantly, as a moving average juror engages in more frequent updates, he does not become overconfident, but instead approaches the correct strength of belief. If he perfectly tracks the information, this result holds exactly. Under the signal model, the result holds in expectation.

Proposition 10: Suppose $\lambda_m = g_m/(g_m + i_m) = \lambda \ \forall m \in \{1, 2, ..., k\}$ and $f(\cdot)$ is the identity function. If a committee member updates with a moving average, then $\lim_{m \to \infty} \mathrm{E}(J_m) = \frac{g}{g+i}$.

In contrast to the overconfident Bayesian juror, if the rate of information arrival is constant ($g(t) = g, \ i(t) = i \ \forall t \in [0, 1]$), the moving average juror becomes more accurate in her beliefs with more frequent updates. The reason Proposition 10 is stated in terms of λ_m instead of $g(t)$ and $i(t)$ is that constant information arrival is not a Nash equilibrium for two debaters in front of a moving average audience. Rather, the unique Nash equilibrium involves setting a similar ratio of evidence over time, as described in Section VII.

Suppose now that a committee is comprised of moving average members. Since a moving average individual becomes more accurate in her beliefs as she updates more frequently, such committees should do better by deliberating early, regardless of their proof threshold. To formalize the result, assume that members incorporate information at each meeting through a simple mean of the group signals. In that case, while Bayesian committees with a high (low) threshold of proof should (not) deliberate early, moving average committees always should.

Proposition 11: WLOG let $\lambda_m = \lambda > 1/2 \; \forall m \in \{1, 2, ..., k\}$. Consider a committee deliberating at points $L \subseteq \{1, 2, ..., k\}$ with $k \in L$, and members incorporate information at each deliberation through the mean of the group signals. If members adopt a moving average process of beliefs, early deliberation ($L \supset \{k\}$) yields better decisions on average (compared to $L = \{k\}$).

5.6 Winning a Debate

Earlier it was claimed that a constant rate of information arrival can be motivated by a situation where two agents each defend the opposite side of an adversarial debate. The time between audience belief updates can be thought of as rounds of the debate. Proposition 7 formalizes the result for a Bayesian audience. Interestingly, neither debater should concede lower odds in one round of a debate in the attempt to win another round. Since the order of the evidence only matters to a Bayesian audience through its effect on the signal probabilities, the best strategy for a debater is to maximize the signal probabilities in favor of her position, which is done by dispensing information at a constant rate.[11]

Proposition 12: Let two debaters each seek to convince a committee of their own position. Suppose $f(\cdot)$ is linear, committee members update beliefs $k > 1$ times following Bayes' rule, and each agent has a fixed quantity of evidence to present, the timing of which may be chosen. Then the unique Nash equilibrium is for each side to present her evidence at a constant rate over time.

For a moving average audience, the equilibrium strategies become slightly more complicated, but the proportion of evidence favoring each side is again constant for

---

[11]Note the difference between the result here and Kamenica and Gentzkow (2011). Here the Bayesian receiver draws an inference from a signal derived from the evidence presented by the sender. In Kamenica and Gentzkow (2011), the sender manipulates the signal directly.

every round.

Proposition 13: Let two debaters each seek to convince a committee of their own position. Suppose $f(\cdot)$ is linear, committee members update beliefs $k > 1$ times following the moving average rule, and each agent has a fixed quantity of evidence to present, the timing of which may be chosen. Then the unique Nash equilibrium is for the two sides to present their evidence at rates

$$g_j = \sqrt{w^{k-j}}\left(\frac{1 - \sqrt{(w)}}{1 - \sqrt{(w)^k}}\right)g, \text{ and } i_j = \sqrt{w^{k-j}}\left(\frac{1 - \sqrt{(w)}}{1 - \sqrt{(w)^k}}\right)i \quad \forall j \in \{1, 2, ..., k\}.$$

The rate of information is not constant in equilibrium because the order of evidence matters to a moving average updater independently of its effect on the signal probability. The weight placed on previous information, w, compared to new information, alters the incentive of the debater to present evidence at one time versus another. Depending on the value of w, the audience could exhibit a primacy or recency effect, where information presented first or last is weighted more heavily.

Although a debate could allow one side to speak for the first half of the alloted time and the other side to speak for the second, in practice each side is typically allowed to rebut the other several times. This alternating structure speaks to the underlying audience model, apparently discounting the Bayesian model since the order of evidence is unimportant to a Bayesian. Instead, the structure supports a moving average model since an alternating platform is fairer to each side in that model.

For a simple illustration, suppose one debater has two signals equal to zero and the other has two signals equal to one. If one side is given the platform for the first half of the debate, that side is either advantaged or disadvantaged conditional on

Table 5.2: Order of Evidence – Moving Average Audience

Signal Order	(1,1,0,0)	(1,0,1,0)	(0,1,0,1)	(0,0,1,1)
J_4	0.22	0.34	0.66	0.78

The table contains the strength of belief a moving audience member will hold after four rounds of debate when $J_0 = 1/2$ and $w = 1/2$.

w. For a particular set of parameters, Table 5.2 shows how the order of evidence alters the strength of belief an audience member will hold after four rounds of debate. Having the floor during the second half of the debate is clearly valuable to either side. Alternating the speaker dampens the advantage. Of course, the moving average model is a consistent but not necessary motivation for the common debate structure. A variety of departures from the Bayesian model, including confirmation bias and other cognitive dangers, might justify the format.

5.7 Conclusion

Whether a Bayesian should reserve judgment or jump to conclusions depends on her loss function and the accuracy of the information signal. While a cynic might observe a positive correlation between individuals who feel most confident in their opinions and those who spend the least timing forming them, those individuals may in fact be following a rational, Bayesian strategy with a symmetric loss function.

Unlike a Bayesian, however, a moving average individual who updates his beliefs more frequently becomes more accurate on average, without the attendant overconfidence. Sound policy therefore depends on the approximate model of learning that people actually adopt. For example, prohibiting jurors from discussing their criminal case during trial breaks is wise for a Bayesian jury but unwise for a moving average jury.

The courtroom environment in particular appears to have a tangled motivation. In a Bayesian juror model, the common jury prohibition on discussing a case during a

criminal trial makes sense since the standard of proof is high, but the same prohibition is undesirable in a civil trial, where the standard of proof is low. Moreover, the alternating rebuttal structure of attorney debate would appear to favor a moving average model.[12] Of course, cognitive factors like confirmation bias certainly play a role too, and resolving their interaction with in the current model has the potential to improve the environment for committee decision making.

One limitation here is that no great complementarities in information over time are allowed. Changing the order of evidence is only sensible if a chunk of information has a meaningful interpretation on its own. Certain orders of information might even combine to produce a synergistic effect, as when a debater delivers a "knock-out punch." Also, in the courtroom, jurors do not always know what exactly they are evaluating until the end of the process (as noted by Hastie 1993). Lastly, committee members are assumed identical, but heterogenous members can bring different information with which to evaluate the evidence at hand. Understanding how external information affects the interpretation of evidence is another important question on the proper timing of belief updates and deliberation.

---

[12]In a courtroom, the structure is at least partly logistical. It is easier for a witness to remain on the stand for cross-examination than to return later for questioning.

References

Andreoni, James. 1991. Reasonable doubt and the optimal magnitude of fines: should the penalty fit the crime? *The RAND Journal of Economics* 22(3): 385–395.

Andreoni, James. 1995. Criminal deterrence in the reduced form: A new perspective on Ehrlich's seminal study. *Economic Inquiry* 33 (3): 476–483.

Andridge, R., and R. Little. 2010. A review of hot deck imputation for survey non-response. *International Statistics Review* 78(1): 40–64.

Austen-Smith, David, and Jeffrey S. Banks. 1996. Information aggregation, rationality, and the Condorcet jury theorem. *The American Political Science Review* 90 (1): 34–45.

Austen-Smith, David, and Timothy J. Feddersen. 2006. Deliberation, preference uncertainty, and voting rules. *The American Political Science Review* 100 (2): 209–217.

Bartels, Robert D. 1981. Punishment and the burden of proof in criminal cases: A modest proposal. *Iowa Law Review* 66: 899–930.

Bellin, Jeffrey. 2010. Is punishment relevant after all? A prescription for informing juries of the consequences of conviction. *Boston University Law Review* 90: 2223–2265.

Bénabou, Roland. 2013. Groupthink: collective delusions in organizations and markets. *Review of Economic Studies* 80: 429–462.

Ben-Yashar, Ruth, Winston T. H. Koh, and Shmuel Nitzan. 2012. Is specialization desirable in committee decision making? *Theory and Decision* 72 (3): 341–357.

Bibas, Stephanos. 2012. *The Machinery of Criminal Justice.* Oxford University Press.

Bowers, Jeffrey S., and Colin J. Davis. 2012. Bayesian just-so stories in psychology and neuroscience. *Psychological Bulletin* 138 (3): 389–414.

Richard, Boylan T. 2005. What do prosecutors maximize? Evidence from the careers of U.S. attorneys. *American Law and Economics Review* 7(2): 379–380.

Bureau of Justice Statistics. 2010. Table 5.22: Criminal defendants disposed of in U.S. district courts. *Sourcebook of Criminal Justice Statistics Online.* Access date: 07-12-2016. doi: http://www.albany.edu/sourcebook/pdf/t5222010.pdf.

Bushway, Shawn, Emily Owens, and Anne Morrison Piehl. 2012. Sentencing guidelines and judicial discretion: Quasi-experimental evidence from human calculation errors. *Journal of Empirical Legal Studies* 9(2): 291–319.

Bushway, Shawn, Brian D. Johnson, and Lee Ann Slocum. 2007. Is the magic still there? The use of the Heckman two-step correction for selection bias in criminology. *Journal of Quantitative Criminology* 23: 151–178.

Caplin, Andrew, and Mark Dean. 2015. Revealed preference, rational inattention, and costly information acquisition. *American Economic Review* 105 (7): 2183–2203.

Coughlan, Peter J. 2000. In defense of unanimous jury verdicts: Mistrials, communication, and strategic voting. *The American Political Science Review* 94 (2): 375–393.

Cragg, John G. 1971. Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica* 39(5): 829–844.

Cross, John G. 1973. A stochastic learning model of economic behavior. *The Quarterly Journal of Economics* 87 (2): 239–266.

Daughety, Andrew, and Jennifer Reignanum. 2015. Informal sanctions on prosecutors and defendants and the disposition of criminal cases. *Working Paper.*

Davis, Graham A., and Robert D. Cairns. 2012. Good timing: the economics of optimal stopping. *Journal of Economic Dynamics & Control* 69 (345): 118–121.

Davis, Michael L. 1994. The value of truth and the optimal standard of proof in legal disputes. *Journal of Law, Economics, & Organization* 10(2): 343–359.

Dharmapala, Dhammika, Nuno Garoupa, and Richard H. McAdams. 2014. Do exclusionary rules convict the innocent? In *Research Handbook on Economic Models of Law*, edited by Matthew Baker and Thomas Miceli. Edward Elgar Publishing.

DeGroot, Morris H. 1974. Reaching a consensus. *Journal of the American Statistical Association* 69 (345): 118–121.

Dhami, Mandeep K. 2008. On measuring quantitative interpretations of reasonable doubt. *Journal of Experimental Psychology: Applied* 36: 255–265.

Diamond, Peter A. 1971. A model of price adjustment. *Journal of Economic Theory* 3: 156–168.

Diamond, Shari S., and Stalans, Loretta J. 1989. The myth of judicial leniency in sentencing. *Behavioral Sciences & the Law* 7 (1): 73–89.

Dzur, Albert W. 2012. *Punishment, participatory democracy, and the jury.* Oxford University Press.

Epstein, Lee, William M. Landes, and Richard A. Posner. 2013. *The behavior of federal judges.* Harvard University Press.

Feddersen, Timothy, and Wolfgang Pesendorfer. 1998. The inferiority of unanimous jury verdicts under strategic voting. *The American Political Science Review* 92 (1): 23–35.

Federal Judicial Center. Biographical Directory of Federal Judges, 1789-present. Access date: 06-27-2016. doi: http://www.fjc.gov/history/home.nsf/page/export.html.

Feess, Eberhard, and Ansgar Wohlschlegel. 2009. Why higher punishment may reduce deterrence. *Economics Letters* 104 (April): 69–71.

Fischman, Joshua B., and Max M. Schanzenbach. 2011. Do standards of review matter? The case of federal criminal sentencing. *The Journal of Legal Studies* 40(2): 405–437.

Fischhoff, Baruch, Slovic, Paul, and Sarah Lichtenstein. 1997. Knowing with certainty: The appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance* 3(4): 552–564.

Freeborn, Beth A., and Monica E. Hartmann. 2010. Judicial discretion and sentencing behavior: Did the Feeney Amendment rein in district judges? *Journal of Empirical Legal Studies* 7(2): 355–378.

Friedman, Ezra, and Abe Wickelgren. 2006. Bayesian juries and the limits to deterrence. *Journal of Law, Economics, & Organization* 22 (1): 70–86.

Gay, Gerald D., Martin F. Grace, Jayant R. Kale, and Thomas H. Noe. 1989. Noisy jury trials and the choice of trial mode in a sequential signalling game: Theory and evidence. *The RAND Journal of Economics* 20 (2): 196–213.

Gerardi, Dino, and Leeat Yariv. 2007. Deliberative voting. *Journal of Economic Theory* 134: 317–338.

Gerardi, Dino, and Leeat Yariv. 2008. Information acquisition in committees. *Games and Economic Behavior* 62: 436–459.

Gerling, Kerstin, Hans Peter Grüner, Alexandra Kiel, and Elisabeth Schulte. 2005. Information acquisition and decision making in committees: A survey. *European Journal of Political Economy* 21 (3): 563–597.

Gigerenzer, Gerd, and Ulrich Hoffrage. 1995. How to improve Bayesian reasoning without instruction: frequency formats. *Psychological Review* 102 (4): 684–704.

Glaeser, Edward L., and Cass R. Sunstein. 2009. Extremism and social learning. *Journal of Legal Analysis* 1 (1): 263–324.

Golub, Benjamin, and Matthew O. Jackson. 2010. Naïve learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics* 2 (1): 112–149.

Griffiths, Thomas L., Chater, Nick, Norris, Dennis, and Alexandre Pouget. 2012. How the Bayesians got their beliefs (and what those beliefs actually are): comment on Bowers and Davis (2012). *Psychological Bulletin* 138(3): 415–422.

Hastie, Reid. 1993. *Inside the Juror.* Cambridge University Press.

Heckman, James. 1979. Sample selection bias as a specification error. *Econometrica* 47(1): 153–161.

Helland, Eric, and Alexander Tabarrok. 2000. Runaway judges? Selection effects and the jury. *Journal of Law, Economics, & Organization* 16 (2): 306–333.

Hilbert, Martin. 2012. Toward a synthesis of cognitive biases: How noisy information processing can bias human decision making. *Psychological Bulletin.* 138 (2): 211–237.

Hoeffding, Wassily. 1963. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association* 58 (301): 13–30.

Hoffman, Morris B. 2003. The case for jury sentencing. *Duke Law Journal* 52: 951–1010.

Hoffman, Ross M., Kagel, John H., and Dan Levin. 2011. Simultaneous versus sequential information processing. *Economics Letters* 112: 16–18.

Horowitz, Irwin A. 1997. Reasonable doubt instructions: Commonsense justice and standard of proof. *Psychology, Public Policy, and Law* 3 (2/3): 285–302.

Iontcheva, Jenia. 2003. Jury sentencing as democratic practice. *Virginia Law Review* 89 (2): 311–383.

Iyengar, Radha. 2007. An analysis of the performance of federal indigent defense counsel. *NBER Working Paper 13187*.

Jackson, Matthew O., and Xu Tan. 2013. Deliberation, disclosure of information, and voting. *Journal of Economic Theory* 148: 2–30.

Jackson, Randall R. 1999. Missouri's jury sentencing law: a relic the legislature should put to rest. *Journal of the Missouri Bar* 55: 14–23.

Jacobs, Robert A., and John K. Kruschke. 2011. Bayesian learning theory applied to human cognition. *Wiley Interdisciplinary Reviews: Cognitive Science* 2: 8–21.

Jacoby, Jacob, and Wayne D. Hoyer. 1982. Viewer miscomprehension of televised communication: Selected findings. *Journal of Marketing* 46: 12–26.

Jacoby, Jacob, and Wayne D. Hoyer. 1989. The comprehension/miscomprehension of print communication: Selected findings. *Journal of Consumer Research* 15(4): 434–443.

Jouras, James P. 1952. On modernizing Missouri's criminal punishment procedure. *The University of Kansas City Law Review* 20: 299–316.

Kalven, Harry Jr., and Hans Zeisel. 1966. *The American Jury*. The University of Chicago Press.

Kamenica, Emir, and Matthew Gentzkow. 2011. Bayesian persuasion. *American Economic Review* 101(6): 2590–2615.

Kaplan, Martin F., and Sharon Kupra. 1986. Severe penalties under the control of others can reduce guilt verdicts. *Law & Psychology Review* 10: 1-18.

Kaplow, Louis. 2011. On the optimal burden of proof. *Journal of Political Economy* 119(6): 1104–1140.

Kaplow, Louis. 2012. Burden of Proof. *The Yale Law Journal* 121: 738-859.

Kahneman, D., Slovic, P., and Amos Tversky (eds.). 1982. *Judgment under uncertainty: Heuristics and biases.* Cambridge University Press.

de Keijser, Jan W., and Peter J. van Koppen. 2007. Paradoxes of proof and punishment: psychological pitfalls in judicial decision making. *The British Psychological Society* 12: 189–205.

Kerr, Charles. 1918. A needed reform in criminal procedure. *Kentucky Law Journal* 6: 107–111.

Kerr, Norbert L. 1978. Severity of prescribed penalty and mock jurors' verdicts. *Journal of Personality and Social Psychology* 36 (12): 1431–1442.

Kessler, Daniel P., and Anne Morrison Piehl. 1998. The role of discretion in the criminal justice system. *Journal of Law, Economics, & Organization* 14 (2): 256–276.

Koch, Chantal, and Dennis J. Devine. 1999. Effects of reasonable doubt definition and inclusion of a lesser charge on jury verdicts. *Law and Human Behavior* 23 (6): 653–674.

King, Nancy J., and Rosevelt L. Noble. 2004. Felony jury sentencing in practice: a three-state study. *Vanderbilt Law Review* 57(3): 885–962.

King, Nancy J., and Rosevelt L. Noble. 2005. Jury sentencing in noncapital cases: comparing severity and variance with judicial sentences in two states. *Journal of Empirical Legal Studies* 2 (2): 331–367.

LaFountain, R. Schauffler, S. Strickland, K. Holt, & K. Lewis, eds. 2016. *Court Statistics Project Data Viewer.* National Center for State Courts. Updated: 03-15-2016. Access date: 07-12-2016. doi: www.courtstatistics.org.

Lanni, Adriaan. 1999. Jury sentencing in noncapital cases: an idea whose time has come (again)? *The Yale Law Journal* 108 (7): 1775–1803.

Lehrer, Ehud, and Rann Smorodinsky. 2000. Relative entropy in sequential decision problems. *Journal of Mathematical Economics* 33: 425–439.

Leib, Ethan J. 2008. A comparison of criminal jury decision rules in democratic countries. *Ohio State Journal of Criminal Law* 5: 629–644.

Leipold, Andrew. 2005. Why are federal judges so acquittal prone? *Washington University Law Review* 83(1): 151–227.

Lillquist, Erik. 2004. The puzzling return of jury sentencing: misgivings about *Apprendi*. *North Carolina Law Review* 82: 621–715.

Lizzeri, Alessandro, and Leeat Yariv. 2013. Sequential deliberation. *Working Paper.*

Lundberg, Alexander. 2016. Sentencing discretion and burdens of proof. *International Review of Law and Economics* 46: 34–42.

MacCoun, Robert J., and Tom R. Tyler. 1988. The basis of citizens' perceptions of the criminal jury: Procedural fairness, accuracy, and efficiency. *Law and Human Behavior* 12(3): 333–352.

Marcus, Gary F., and Ernest Davis. 2013. How robust are probabilistic models of higher-level cognition? *Psychological Science* 24 (12): 2351–2360.

Miceli, Thomas J. 1990. Optimal prosecution of defendants whose guilt is uncertain. *Journal of Law, Economics, & Organization* 6 (1): 189–201.

Möbius, Markus M., Niederle, Muriel, Niehaus, Paul, and Tanya S. Rosenblat. 2014. Managing self-confidence. *Working Paper.*

Mukhopadhaya, Kaushik. 2003. Jury size and the free rider problem. *Journal of Law, Economics, and Organization* 19 (1):24–44.

Mustard, David. 2001. Racial, ethnic, and gender disparities in sentencing: Evidence from the U.S. federal courts. *Journal of Law & Economics* 44(1): 285–314.

Neilson, William S., and Harold Winter. 2008. Votes based on protracted deliberations. *Journal of Economic Behavior & Organization* 67 (1): 308–321.

Nickerson, Raymond S. 1998. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology* 2 (2): 175–220.

Norton, Edward C., Hua Wang, and Chunrong Ai. 2004. Computing interaction effects and standard errors in logit and probit models. *The Stata Journal* 4(2): 154–167.

Ortoleva, Pietro. 2012. Modeling the change of paradigm: Non-Bayesian reactions to unexpected news." *American Economic Review* 102 (6): 2410–2436.

Persico, Nicola. 2004. Committee design with endogenous information. *Review of Economic Studies* 71 (1): 165–191.

Posner, Richard A., and Albert H. Yoon. 2011. What judges think of the quality of legal representation. *Stanford Law Review* 63: 317–350

Rabin, Matthew, and Joel L. Schrag. 1999. First impressions matter: A model of confirmatory bias. *The Quarterly Journal of Economics* 114 (1): 37–82.

Rehavi, Marit M., and Sonja B. Starr. 2014. Racial disparity in federal criminal sentences. *Journal of Political Economy* 122(6): 1320–1354.

Roberts, Julian V., and Jan W. de Keijser. 2014. Democratising punishment: sentencing, community views and values. *Punishment & Society* 16 (4): 474–498.

Robinson, Paul H., Geoffrey P. Goodwin, and Michael Reisig. 2010. The disutility of injustice. *New York University Law Review* 85 (6): 1940–2033.

Rose, Mary R., Ellison, Christopher, and Shari Seidman Diamond. 2008. Preferences for juries over judges across racial and ethnic groups. *Social Science Quarterly*

89(2): 372–391.

Roach, Michael. 2014. Indigent defense counsel, attorney quality, and defendant outcomes. *American Law and Economics Review* 16(2): 577–619.

Schanzenbach, Max. 2005. Racial and sex disparities in prison sentences: The effect of district-level judicial demographics. *The Journal of Legal Studies* 34(1): 57–92.

Schulhofer, Stephen J. 1988. Criminal justice discretion as a regulatory system. *The Journal of Legal Studies* 17 (1): 43–82.

Schwartzstein, Joshua. 2014. Selective attention and learning. *Journal of the European Economic Association* 12 (6): 1423–1452.

Schulte, Elisabeth. 2012. Communication in committees: Who should listen? *Public Choice* 150 (1-2): 97–117.

Schum, David A., and Anne W. Martin. 1982. Formal and empirical research on cascaded inference in jurisprudence. *Law & Society Review* 17 (1): 105–152.

Shavell, Steven. 2007. Optimal discretion in the application of rules. *American Law and Economics Review* 9 (1): 175–194.

Siegel, Ron, and Brulo Strulovici. 2015. On the design of criminal trials: The benefit of three-verdict systems. *Working Paper*.

Simon, Rita James, and Linda Mahan. 1971. Quantifying burdens of proof: A view from the bench, the jury, and the classroom. *Law & Society Review* 5(3): 319–330.

Sims, Christopher A. 2003. Implications of rational inattention. *Journal of Monetary Economics* 50 (3): 665–690.

Sobel, Joel. 2000. Economists' models of learning. *Journal of Economic Theory* 94 (2): 241–261.

Snyder, Edward A. 1990. The effect of higher criminal penalties on antitrust enforcement. *The Journal of Law & Economics* 33(2): 439–462.

Starr, Sonja. 2015. Estimating gender disparities in federal criminal cases. *American Law and Economics Review* 17(1): 127–159.

Stoffelmayr, Elisabeth, and Shari Diamond. 2000. The conflict between precision and flexibility in explaining "beyond a reasonable doubt." *Psychology, Public Policy, and Law* 6 (3): 769–787.

Tenenbaum, Joshua B., Kemp, Charles, Griffiths, Thomas L., and Noah D. Goodman. 2011. How to grow a mind: statistics, structure, and abstraction. *Science* 331: 1279–1285.

United States Department of Justice. Bureau of Justice Statistics. Federal Justice Statistics Program: Arrests and Bookings for Federal Offenses, 1995-2010 [United States]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2011-03-08.

United States Department of Justice. Bureau of Justice Statistics. Federal Justice Statistics Program: Defendants Sentenced Under the Sentencing Reform Act, 2000-2010 [United States]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2011-03-08.

United States Department of Justice. Bureau of Justice Statistics. Federal Justice Statistics Program: Defendants in Federal Criminal Cases in District Court – Terminated, 2000-2010 [United States]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2011-03-08.

United States Sentencing Commission. 2006. *Final Report of the Impact of United States v. Booker on Federal Sentencing.* Washington, D.C.: United States Sentencing Commission.

United States Sentencing Commission. 2010. *Results of survey of United States district judges January 2010 through March 2010.* Washington, D.C.: United States Sentencing Commission.

United States Sentencing Commission. 2012. *Report on the Continuing Impact of United States v. Booker on Federal Sentencing.* Washington, D.C.: United States Sentencing Commission.

Viscusi, Kip W. 2001. Jurors, judges, and the mistreatment of risk by the courts. *The Journal of Legal Studies* 30(1): 107–142.

Wald, Abraham. 1947. Foundations of a general theory of sequential decision functions. *Econometrica* 15 (4): 279–313.

Webster, Charles W. 1999. Jury sentencing—grab-bag justice. *Southwestern Law Journal* 14: 221–231.

Weninger, Robert A. 1994. Jury sentencing in noncapital cases: a case study of El Paso County, Texas. *Washington University Journal of Urban & Contemporary Law* 45: 3–40.

Wooldridge, Jeffrey M. 2010. Econometric analysis of cross section and panel data. MIT Press.

Wilson, Andrea. 2014. Bounded memory and biases in information processing. *Econometrica* 82 (6): 2257–2294.

Zeelenberg, M. 1999. Anticipated regret, expected feedback and behavioral decision making. *Journal of Behavioral Decision Making* 12 (2): 93-106.

Appendix A   Proofs for Chapter 2

We want to show that

$$p(\ell^*) = \frac{c_1(\ell^*)}{c_1(\ell^*) + c_2 + u(\ell^*; t)} < p(t) = \frac{c_1(t)}{c_1(t) + c_2 + u(t; t)}$$

if and only if

$$\frac{c_2 + u(t; t)}{c_2 + u(\ell^*; t)} < \frac{c_1(t)}{c_1(\ell^*)}.$$

Define $a = c_1(\ell^*), b = c_1(t), c = u(\ell^*; t), d = u(t; t)$, and $e = c_2$. Then $p(\ell^*) = a/(a + e + c)$, and $p(t) = b/(b + e + d)$. We have

$$p(\ell^*) < p(t) \Leftrightarrow$$

$$\frac{a}{a + e + c} < \frac{b}{b + e + d} \Leftrightarrow$$

$$ae + ad < be + bc \Leftrightarrow$$

$$\frac{e + d}{e + c} < \frac{b}{a}.$$

Lastly, a sufficient condition for the final inequality is that $\frac{b}{a} > \frac{d}{c}$: Given that $b > a$,

$$\frac{b}{a} > \frac{d}{c} \implies bc > ad \implies bc + be > ad + ae \implies \frac{b}{a} > \frac{e + d}{e + c}.$$

Appendix B   Proofs for Chapter 3

Proof of Proposition 5. Let l^* and p^* define the optimal sentence and burden of proof for a trier-of-fact with discretion, given a target sentence t. From (1) and (4),

$$p^* - \frac{c_1(l^*)}{c_1(l^*) + c_2(t) + u(l^*, t)} = 0, \quad \text{and} \tag{A1}$$

$$\frac{\partial u(l^*, t)}{\partial l} - \frac{(1-p)}{p} c_1'(l^*) = 0. \tag{A2}$$

By the implicit function theorem,

$$
\begin{bmatrix}
\frac{c_1(l^*)[c_1'(l^*) + \frac{\partial u(l^*,t)}{\partial l}] - c_1'(l^*)[c_1(l^*) + c_2(t) + u(l^*,t)]}{[c_1(ll^*) + c_2(t) + u(l^*,t)]^2} & 1 \\[2em]
\frac{\partial^2 u(l^*,t)}{\partial l^2} - \frac{(1-p)}{p} c_1''(l^*) & 0
\end{bmatrix}
\begin{bmatrix}
\frac{\partial l^*}{\partial t} \\[2em]
\frac{\partial p^*}{\partial t}
\end{bmatrix}
=
\begin{bmatrix}
-\frac{c_1(l^*)[\frac{\partial u(l^*,t)}{\partial t} + c_2'(t)]}{[c_1(l^*) + c_2(t) + u(l^*,t)]^2} \\[2em]
-\frac{\partial^2 u(l^*,t)}{\partial l \partial t}
\end{bmatrix}
$$

in a neighborhood of (l^*, p^*) provided the Jacobian matrix on the left, denoted by J, has a non-zero determinant. But by the second-order condition (5), $|J| < 0$. Solving the system by Cramer's Rule yields

$$\frac{\partial l^*}{\partial t} = \frac{1}{|J|}\left(-\frac{\partial^2 u(l^*, t)}{\partial l \partial t}\right), \quad \text{and}$$

$$\frac{\partial p^*}{\partial t} = \frac{1}{|J|}\left(\left[\frac{c_1'(l^*)[c_1(l^*) + c_2(t) + u(l^*, t)] - c_1(l^*)[c_1'(l^*) + \frac{\partial u(l^*,t)}{\partial l}]}{[c_1(l^*) + c_2(t) + u(l^*, t)]^2}\right] \cdot \left[\frac{\partial^2 u(l^*, t)}{\partial l \partial t}\right] + \left[\frac{\partial^2 u(l^*, t)}{\partial l^2} - \frac{(1-p)}{p} c_1''(l^*)\right] \cdot \left[\frac{c_1(l^*)[\frac{\partial u(l^*,t)}{\partial t} + c_2'(t)]}{[c_1(l^*) + c_2(t) + u(l^*, t)]^2}\right]\right).$$

Since $|J| < 0$, then, by *Assumption 2*, $\partial l^*/\partial t > 0$. However, $\partial p^*/\partial t$ may be positive or negative depending on preferences. *Assumptions 1* and *2* are not sufficient for the result. □

Proof of Claim 3. Although $\frac{\partial l^*}{\partial t} > 0$ by *Proposition 2*, the algebra may be helpful. Under *Assumption 3*, A1 and A2 yield

$$l^* = t\left(\frac{1}{1 - [\frac{-d(1-p)}{ap}]^{\frac{1}{\alpha-1}}}\right),$$

and, defining $z \equiv [\frac{-d(1-p)}{ap}]^{\frac{1}{\alpha-1}}$,

$$\frac{\partial l^*}{\partial t} = \left(\frac{1}{1 - [\frac{-d(1-p)}{ap}]^{\frac{1}{\alpha-1}}}\right) > 0.$$

(*Assumption 1* requires that α be even, so $\alpha - 1$ is odd, and $z < 0$, whence the sign.) Next, the optimal burden of proof is

$$p^* = \frac{d(\frac{1}{1-z})^\alpha t^\alpha}{d(\frac{1}{1-z})^\alpha t^\alpha + et^\gamma - a(\frac{z}{1-z})^\alpha t^\alpha + bt^\beta + c}.$$

Denoting the denominator of p^* by ξ,

$$\frac{\partial p^*}{\partial t} = \left(\frac{1}{\xi^2}\right)\left(\xi \alpha d(\frac{1}{1-z})^\alpha t^{\alpha-1} - d(\frac{1}{1-z})^\alpha t^\alpha \left[\alpha d(\frac{1}{1-z})^\alpha t^{\alpha-1} + e\gamma t^{\gamma-1} - \right.\right.$$
$$\left.\left. a\alpha(\frac{z}{1-z})^\alpha t^{\alpha-1} + b\beta t^{\beta-1}\right]\right).$$

Dropping the $(1/\xi^2)$ term, the expression simplifies to

$$d(\frac{1}{1-z})^\alpha[(\alpha - \gamma)et^{\alpha+\gamma-1} + (\alpha - \beta)bt^{\alpha+\beta-1} + \alpha ct^{\alpha-1}],$$

which is greater than zero by assumption. Therefore, under *Assumption 3*, both

$\partial l^*/\partial t > 0$ and $\partial p^*/\partial t > 0$. $\square$

Proof of Claim 4. Under *Assumption 3*,

$$c_1'(l^*) = d\alpha \left(\frac{1}{1 - [\frac{-d(1-p)}{ap}]} \right) t^{\alpha-1}, \quad c_2'(t) = e\gamma t^{\gamma-1}, \quad \text{and}$$

$$\frac{\partial u(l^*, t)}{\partial t} = -a\alpha \left(\frac{[\frac{-d(1-p)}{ap}]^{\frac{1}{\alpha-1}}}{1 - [\frac{-d(1-p)}{ap}]^{\frac{1}{\alpha-1}}} \right)^{\alpha-1} t^{\alpha-1} + b\beta t^{\beta-1}.$$

By allowing $b, e \to \infty$, $\frac{\partial u(l^*, t)}{\partial t} > c_1'(l^*)$ and $c_2'(t) > c_1'(l^*)$. $\square$

Appendix C Proofs for Chapter 4

C.1 Propositions

Proof of Claim 6. The proof is by construction. Take first the case of $k = n$. We seek unique events $\mathcal{G}$ and E_i such that $P(\mathcal{G}|E_i) = x$ $\forall i \in \{1, 2, ..., n\}$ and $P(\mathcal{G}|E_1, ..., E_n) = x$ for $x \in (0, 1)$. We focus on the more likely signal of $x \in (.5, 1)$.[1] It is enough to show the existence of polygons, defined by different sets of coordinates, such that the ratio of the area of the intersection of polygons $\mathcal{G}$ and E_i to the area of polygon E_i is x $\forall i$, and the ratio of the intersection of all of the polygons to the intersection of all but $\mathcal{G}$ is also x (where both an event and its polygon are denoted by the same variable).

To that end, consider first the case of $n = 2$ in Figure A1. The polygon $\mathcal{G}$ is a square with area normalized to unity. The polygons E_1 and E_2 are squares translated to the left and right by $1 - x$, respectively. Let A_Y denote the area of polygon Y. Then $A_{\mathcal{G} \cap E_i}/A_{E_i} = x$ for $i = 1, 2$.

Figure A1: Proof of *Claim 4*
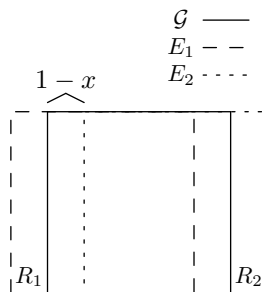
As pictured, the regions R_1 and R_2 do not overlap, but we are free to move them

---

[1] After the demonstration for $x \in (.5, 1)$, the case for $x \in (0, .5]$ should be clear.

anywhere outside of $\mathcal{G}$ without undoing the uniqueness of E_1 and E_2. The question is whether, by moving R_1 and R_2, we can satisfy $A_{\mathcal{G} \cap E_1 \cap E_2} / A_{E_1 \cap E_2} = x$. The smallest such ratio is obtained by a complete overlap, in which case

$$\frac{A_{\mathcal{G} \cap E_1 \cap E_2}}{A_{E_1 \cap E_2}} = \frac{(2x-1)}{(2x-1)+(1-x)}.$$

Therefore, if

$$\frac{(2x-1)}{(2x-1)+(1-x)} = \frac{2x-1}{x} \leq x \iff -x^2 + 2x - 1 \leq 0,$$

there exists an overlap of R_1 and R_2 such that $A_{\mathcal{G} \cap E_1 \cap E_2} / A_{E_1 \cap E_2} = x$, and the claim is proven for $n = 2$. But since $-x^2 + 2x - 1 \leq 0 \ \forall x$, such an overlap exists. Next, note that

$$\frac{A_{E_1 \cap E_2 ... \cap E_n \cap \mathcal{G}}}{A_{E_1 \cap E_2 ... \cap E_n}} = \frac{A_{(E_1 \cap E_2 ... \cap E_n) \cap \mathcal{G}}}{A_{(E_1 \cap E_2 ... \cap E_n) \cap \mathcal{G}} + A_{(E_1 \cap E_2 ... \cap E_n) \backslash \mathcal{G}}},$$

which is increasing in $A_{(E_1 \cap E_2 ... \cap E_n) \cap \mathcal{G}}$. For the case of $n > 2$, since n is finite, $\exists$ unique E_i $i \in \{3, ..., n\}$, each with area x inside $\mathcal{G}$, and each leaving a region of area $R_i = R_1 = R_2$ outside of $\mathcal{G}$. Thus, by choosing to completely overlap the R_i, $A_{(E_1 \cap E_2 ... \cap E_n) \backslash \mathcal{G}} = A_{(E_1 \cap E_2) \backslash \mathcal{G}}$, and

$$\frac{A_{E_1 \cap E_2 ... \cap E_n \cap \mathcal{G}}}{A_{E_1 \cap E_2 ... \cap E_n}} = \frac{A_{(E_1 \cap E_2 ... \cap E_n) \cap \mathcal{G}}}{A_{(E_1 \cap E_2 ... \cap E_n) \cap \mathcal{G}} + A_{(E_1 \cap E_2 ... \cap E_n) \backslash \mathcal{G}}} = \frac{A_{(E_1 \cap E_2 ... \cap E_n) \cap \mathcal{G}}}{A_{(E_1 \cap E_2 ... \cap E_n) \cap \mathcal{G}} + A_{(E_1 \cap E_2) \backslash \mathcal{G}}}$$

$$\leq \frac{A_{E_1 \cap E_2 \cap \mathcal{G}}}{A_{E_1 \cap E_2 \cap \mathcal{G} \cap} + A_{(E_1 \cap E_2) \backslash \mathcal{G}}} \leq x$$

by the result for $n = 2$ above.

For $k \in \{2, ..., n-1\}$, assume $n \geq 3$ and take once more the events in Figure A1. Define E_i ($i \in \{3, ..., n\}$) by shifting to the right a rectangle whose left edge is the left edge of $\mathcal{G}$, whose top and bottom edges follow the top and bottom of edges of $\mathcal{G}$, and

Figure A2: Overlapping Events
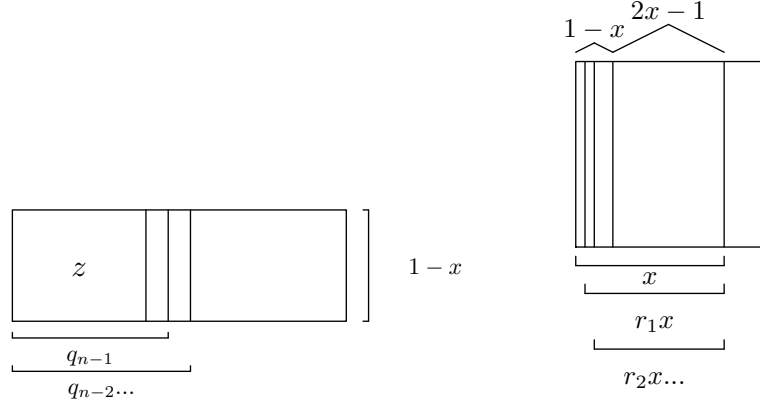
whose right edge is the right edge of E_1. The rightward shifts are different amounts less than $1 - x$, to be made precise momentarily. Thus, $A_{\mathcal{G} \cap E_i}/A_{E_i} = x \ \forall i$, the area $A_{E_1 \cap E_2 ... \cap E_n \cap \mathcal{G}}$ is still $2x - 1$, and for each E_i we are again left with a region R_i of area $1 - x$ outside of $\mathcal{G}$. Define z as the area satisfying

$$\frac{(2x - 1)}{(2x - 1) + z} = x$$

$$\Rightarrow z = 3 - 2x - 1/x.$$

In words, z is the overlap of all remaining R_i rectangles that makes $P(\mathcal{G}|E_1, ..., E_n) = x$. An area of $(1 - x) - z$ is left over to satisfy $P(\mathcal{G}|E_1, ..., E_k) = x$ for each $k \in \{2, ..., n - 1\}$. Define

$$q_{n-i} = z + \frac{i(1 - x - z)}{2n} \quad i \in \{2, ..., n - 1\}$$

as the successive areas of overlap for rectangles of the remaining $n - 2$ intersections of events, and define r_{n-i} to satisfy

$$\frac{r_{n-i}x}{r_{n-i}x + q_{n-i}} = x.$$

Since $\frac{A_{E_1 \cap E_2 ... \cap E_k \cap \mathcal{G}}}{A_{E_1 \cap E_2 ... \cap E_k}} = x$, then $P(\mathcal{G}|E_1, ..., E_k) = x$. The r_{n-i} make precise the afore-mentioned shifts of the rectangle inside $\mathcal{G}$, where they are expressed as the proportion of x from the original right edge of the rectangle at which to begin the new left edge. To confirm that each $(2x - 1) < r_{n-i} < 1 \; \forall i \in \{2, ..., n-1\}$, note that

$$r_{n-i} = \frac{2nz + i(1 - x - z)}{2n(1 - x)} < \frac{2nz + 2n(1 - x - z)}{2n(1 - x)} = 1,$$

and

$$(2x - 1) < \frac{2nz + i(1 - x - z)}{2n(1 - x)} \Leftrightarrow$$

$$2n(1 - x)(2x - 1) < 2nz + i(1 - x - z)$$

A sufficient condition for the latter expression is

$$2n(1 - x)(2x - 1) < 2nz + (1 - x - z) \Leftrightarrow$$

$$(2x - 1) < \frac{1 - z}{1 - x} + \frac{1}{2n}.$$

Again sufficient,

$$(2x - 1) < \frac{1 - z}{1 - x} \Leftrightarrow$$

$$2x - 2x^2 - 1 + x < 1 - 3 + 2x + 1/x \Leftrightarrow$$

$$0 < 2x^3 - x^2 - x + 1,$$

but the inequality holds for $x \in (0, 1)$. $\square$

Proof of Proposition 5. Let $f(\cdot)$ be linear, $k > 1$, $q \in (0, \infty)$, $g = \int_0^1 g(t)d(t)$, and

$i = \int_0^1 i(t)d(t)$. Since $g = \sum_{m=1}^k g_m$ and $i = \sum_{m=1}^k i_m$, if $k = 2$, we have

$$\Lambda_g = \frac{P(\mathcal{G}|k)}{P(\mathcal{I}|k)} = \frac{g_1(g - g_1)}{i_1(i - i_1)}.$$

Then for $c \in (0, 1)$, $\lim_{g_1 \to 0, i_1 \to c} \Lambda_g = 0$, and $\lim_{g_1 \to c, i_1 \to 0} \Lambda_g = \infty$. If $k > 2$, we can define $g^* = \sum_{m=3}^k g_m$ and $i^* = \sum_{m=3}^k i_m$. Then

$$\Lambda_g = \frac{g_1(g^* - g_1)}{i_1(i^* - i_1)} \prod_{m=3}^k \frac{g_m}{i_m},$$

and similar limits apply holding g^* and i^* constant. $\square$

Proof of Proposition 6. Define $\alpha = n_{gs} - n_{is}$, and note that by (2), $\lim_{\alpha \to \infty} p_c \equiv P(\mathcal{G}|n_{gs}, n_{is}) = 1$. Then $\forall \bar{p} \; \exists \alpha_c$ s.t. $p_c > \bar{p}$. The probability of receiving at least α_c more guilty than innocent signals after k updates is

$$P_k(\alpha \geq \alpha_c) = \begin{cases} 0 & \text{if } \alpha_c > k \\ \\ \sum_{j=\lceil \frac{k+\alpha_c+1}{2} \rceil}^k b(j; k, p) & \text{if } \alpha_c \leq k. \end{cases}$$

For large enough k, $\alpha_c \leq k$, and

$$P_k(\alpha \geq \alpha_c) = \sum_{j=\lceil \frac{k+\alpha_c+1}{2} \rceil}^k b(j; k, p) = 1 - \mathrm{F}(\lceil \tfrac{k+\alpha_c-1}{2} \rceil; k, p).$$

In addition, for large enough k, $[\frac{k+\alpha_c-1}{2}] < kp$. Then by Hoeffding's inequality,

$$\mathrm{F}([\tfrac{k+\alpha_c-1}{2}]; k, p) \leq \exp\left(\frac{-(kp - [\frac{k+\alpha_c-1}{2}])^2}{2kp}\right)$$

$$\Rightarrow \lim_{k\to\infty} \mathrm{F}([\tfrac{k+\alpha_c-1}{2}]; k, p) \leq \lim_{k\to\infty} \exp\left(\frac{-(kp - [\frac{k+\alpha_c-1}{2}])^2}{2kp}\right) = 0$$

$$\Rightarrow \lim_{k\to\infty} P_k(\alpha \geq \alpha_c) = \lim_{k\to\infty} [1 - \mathrm{F}([\tfrac{k+\alpha_c-1}{2}]; k, p)] = 1. \ \square$$

Proof of Proposition 8. In text. See Section IV.

Proof of Proposition 9. The result follows directly from the sincerity of voting and Proposition 2.

Proof of Proposition 10. Let $f(\cdot)$ be the identity function and $\lambda_m = g_m/(g_m + i_m) = \lambda \ \forall m \in \{1, 2, ..., k\}$. Then, in the non-signal case,

$$J_m = J_0 w^m + (1-w)\left(\frac{g}{g+i}\right) \sum_{j=1}^{m} w^{m-j} = J_0 w^m + (1-w)\left(\frac{g}{g+i}\right) \sum_{j=0}^{m-1} w^j.$$

By the sum of a geometric series, we have $\lim_{m\to\infty} J_m = \frac{g}{g+i}$. In the signal case, where a guilty signal maps to a 1 on the strength of belief scale and an innocent signal maps to 0, we have $\mathrm{E}(\lambda_m) = \frac{g}{g+i}$, implying $\lim_{m\to\infty} \mathrm{E}(J_m) = \frac{g}{g+i}$. $\square$

Proof of Proposition 11. By Proposition 4, $\lim_{k\to\infty} \mathrm{E}(J_k) = g/(g+i)$, the true ratio of evidence, so accuracy can be defined in terms of the variance, $\mathrm{V}(J_{k,\mathrm{L}})$, where $J_{k,\mathrm{L}}$ denotes the final strength of belief for a member whose committee deliberates at points L. For a committee of size n, let p_{im} denote the probability that member i receives a guilty signal at time m. By assumption, $p_{im} = p \ \forall i \in \{1, 2, ..., n\}, \ m \in \{1, 2, ..., k\}$.

Denoting the variance of the signal by $\sigma^2 = p(1-p)$,

$$J_{k,\mathrm{L}} = J_0 w^k + (1-w)\Big[\sum_{j\in\mathrm{L}} w^{k-j}\Big(\frac{1}{n}\Big)\sum_{i=1}^{n} p_{ij} + \sum_{j\notin\mathrm{L}} w^{k-j} p_{ij}\Big]$$

$$\Rightarrow\ \mathrm{V}(J_{k,\mathrm{L}}) = (1-w)^2\Big[\sum_{j\in\mathrm{L}}(w^{k-j})^2\Big(\frac{\sigma^2}{n}\Big) + \sum_{j\notin\mathrm{L}}(w^{k-j})^2\sigma^2\Big]$$

$$< (1-w)^2\Big[\sum_{j\in\{\mathrm{L}/k\}}(w^{k-j})^2\sigma^2 + \sum_{j\notin\{\mathrm{L}/k\}}(w^{k-j})^2\sigma^2 + \Big(\frac{\sigma^2}{n}\Big)\Big]$$

$$= (1-w)^2\Big[\sum_{j=1}^{k-1}(w^{k-j})^2\sigma^2 + \Big(\frac{\sigma^2}{n}\Big)\Big] = \mathrm{V}(J_{k,\{k\}}).\ \square$$

Proof of Proposition 12. Let $\int_0^1 g(t)d(t) = g$ and $\int_0^1 i(t)d(t) = i$. Suppose that $f(\cdot)$ is linear, that committee members update beliefs $k > 1$ times following Bayes' rule, and that the defense presents i_m of the total innocent evidence at belief update m (where $\sum_{m=1}^{k} i_m = i$). With every member identical, the prosecutor maximizes the odds in favor of guilt that any one of them will come to hold:

$$\max_{\{g_m\}_{m=1}^{k}}\ \Lambda_g = \frac{P(\mathcal{G}|k)}{P(\mathcal{I}|k)} = \prod_{m=1}^{k}\Big[2pt\Big]\frac{g_m}{g_m + i_m}\frac{i_m}{g_m + i_m} = \prod_{m=1}^{k}\frac{g_m}{i_m}$$

$$\text{s.t.}\quad \sum_{m=1}^{k} g_m = g,\ \text{with}\ g \ge g_m \ge 0.$$

Dropping the constant $\prod_{m=1}^{k} 1/i_m$, the first order conditions for the Lagrangian function are

$$\sum_{m=1}^{k} g_m^* = g,\ \text{and}\ \forall j\ \prod_{m\neq j}(g_m^*) = \lambda^*.$$

Therefore, $g_\ell^* = g_j^* = g/k\ \forall\ \ell, j$, and the restrictions on all g_m are satisfied. Lastly, the bordered Hessian is $\mathrm{H} = ee' - \mathrm{I}_k$, where e is a vector of ones and I_k is the identity matrix of dimension k. Since for $n \ge 2$ the leading principal minors have determinant $|\mathrm{H}_n| = (n-1)(-1)^{n-1}$, the solution is indeed a maximum. (To see that

$|H_n| = (n-1)(-1)^{n-1}$, recall the invariance of the determinant to the addition of one row to another. Subtract the first row of H_n from each of the other rows, and then add each of the others to the first. The result is a diagonal matrix whose diagonal elements multiply to $(n-1)(-1)^{n-1}$.) Similarly, the best response of the defense to any strategy of the prosecution is to set $i_m = i/k \ \forall \ m$, and the unique Nash equilibrium is characterized by constant information rates. $\square$

Proof of Proposition 13. Let $\int_0^1 g(t)d(t) = g$ and $\int_0^1 i(t)d(t) = i$. Suppose that $f(\cdot)$ is linear, that committee members update beliefs $k > 1$ times following Bayes' rule, and that the defense presents i_m of the total innocent evidence at belief update m (where $\sum_{m=1}^k i_m = i$). With every member identical, the prosecutor maximizes the posterior strength of belief in favor of guilt that any one of them will come to hold:

$$\max_{\{g_m\}_{m=1}^k} \quad J_k = J_0 w^k + (1-w) \sum_{j=1}^k w^{k-j} \left(\frac{g_m}{g_m + i_m} \right)$$

$$\text{s.t.} \quad \sum_{m=1}^k g_m = g, \ \text{with} \ g \geq g_m \geq 0.$$

Since the prosecutor is maximizing a concave function subject to linear inequality constraints, the constraint qualification holds, and the Kuhn-Tucker conditions are sufficient for a maximum. The first order conditions for the Lagrangian function are

$$\sum_{m=1}^k g_m^* = g, \ \text{and} \ \forall m \ \frac{(g_m^* + i_m)^2}{w^{k-m} g_m^*} = \lambda^*. \tag{A1}$$

The defense faces an analogous maximization problem with Lagrangian first order conditions

$$\sum_{m=1}^k i_m^* = i, \ \text{and} \ \forall m \ \frac{(i_m^* + g_m)^2}{w^{k-m} i_m^*} = \lambda^*. \tag{A2}$$

The unique solution to (A1) and (A2) is obtained by setting

$$g_j = \sqrt{w^{k-j}}\left(\frac{1 - \sqrt{(w)}}{1 - \sqrt{(w)^k}}\right)g, \text{ and } i_j = \sqrt{w^{k-j}}\left(\frac{1 - \sqrt{(w)}}{1 - \sqrt{(w)^k}}\right)i \quad \forall j \in \{1, 2, ..., k\}. \;\square$$

C.2   Nash Equilibria in Strategic Game

The following characterizes the symmetric, responsive Nash equilibria of the multiple update extension of Feddersen and Pesendorfer (1998), hereafter FP. Jurors each update their own beliefs independently k times, and upon receipt of the final evidence, they meet to cast simultaneous votes under a unanimous rule. Assume $P(\mathcal{G}) = P(\mathcal{I}) = 1/2$, and that

$$\beta(k(n-1), kn) > \bar{p}, \tag{B1}$$

where $\beta(k, n)$ is the probability the defendant is guilty when observing k guilty signals out of n total. Condition B1 is analogous to condition 1 in FP. It requires that a juror who receives all innocent signals would still prefer to convict if all other jurors received all guilty signals. In that case, informative voting is not an equilibrium. Define

$$\gamma_G = \sum_{n_{gs}=0}^{k-1} b(n_{gs}; k, p)[p\sigma_{n_{gs}}(g) + (1-p)\sigma_{n_{gs}}(i)],$$

$$\gamma_I = \sum_{n_{gs}=0}^{k-1} b(n_{gs}; k, 1-p)[(1-p)\sigma_{n_{gs}}(g) + p\sigma_{n_{gs}}(i)].$$

Then γ_G (γ_I) is the probability that a juror votes to convict if the defendant is guilty (innocent), where n_{gs} is the number of guilty signals observed by the juror in the first $k-1$ updates, and $\sigma_{n_{gs}}(g)$ ($\sigma_{n_{gs}}(i)$) is the probability the juror votes to convict upon receiving a guilty (innocent) signal at update k, conditional on having observed n_{gs} guilty signals in the first $k-1$ updates. The strategy profile $(\sigma_{n_{gs}}(g), \sigma_{n_{gs}}(i))$

$\forall \, n_{gs} \in \{0, 1, ..., k\}$ constitutes the equilibrium.

For a juror to mix strategies, he must be indifferent between convicting and acquitting, or

$$\frac{b(n_{gs}; , k, p)(\gamma_G)^{n-1}}{b(n_{gs}; , k, p)(\gamma_G)^{n-1} + b(n_{gs}; , k, 1-p)(\gamma_I)^{n-1}} = \bar{p}. \tag{B2}$$

If (B2) holds for some n_{gs}^*, then for any $n_{gs} > n_{gs}^*$, the juror would strictly prefer to convict, and for any $n_{gs} < n_{gs}^*$, he would strictly prefer to acquit. Therefore, he only mixes on $\sigma_{n_{gs}^*}(\cdot)$, but by reasoning identical to Footnote 10 in FP, there are no mixed strategies where $\sigma_{n_{gs}^*}(g) < 1$. We have

$$\sigma_{n_{gs}}(g) \begin{cases} = 1 & \text{if } n_{gs} \geq n_{gs}^* \\ = 0 & \text{if } n_{gs} < n_{gs}^*, \end{cases} \quad \text{and} \, \sigma_{n_{gs}}(i) \begin{cases} = 1 & \text{if } n_{gs} > n_{gs}^* \\ = \sigma & \text{if } n_{gs} = n_{gs}^* \\ = 0 & \text{if } n_{gs} < n_{gs}^*, \end{cases}$$

where σ needs to be solved for. By plugging each $\sigma_{n_{gs}}(j)$, $j \in \{g, i\}$, into the expressions for γ_G and γ_I and solving for σ, the equilibrium is fully characterized. This is where the numbers in Table 1 are derived. Lastly, it must be checked that $\sigma < 1$. If jurors update beliefs k times, then k equilibria exist, one for each $n_{gs}^* \in \{0, 1, ..., k-1\}$.

Appendix D   Additional Empirical Results

D.1   Jury Trials

Table D.1: Probit - Federal Jury Trials

	Conviction	
	marginal effect	p-val
Severity	-.003	.160
Discretion Regime	-.035	.027
Age	.008	.030
Age^2	.000	.102
Black	.037	.294
Male	.043	.077
Black·Male	-.039	.323
Married	-.029	.069
Public	-.103	.022
Private	-.097	.030
Appointed	-.065	.149
%White Judge	.421	.071
%Female Judge	-.113	.401
%Democrat Judge	-.155	.109

Note: n=3,447; years include 1998-2010. The 'marginal effect' column reports average marginal effects. The regression contains unreported dummies for US district. Standard errors are clustered by district.

D.2   Judicial Sentencing Guidelines

	Criminal History Category (Criminal History Points)					
Offense Level	**I** **(0 or 1)**	**II** **(2 or 3)**	**III** **(4, 5, 6)**	**IV** **(7, 8, 9)**	**V** **(10, 11, 12)**	**VI** **(13 or more)**
1	0-6	0-6	0-6	0-6	0-6	0-6
2	0-6	0-6	0-6	0-6	0-6	1-7
3	0-6	0-6	0-6	0-6	2-8	3-9
4	0-6	0-6	0-6	2-8	4-10	6-12
5	0-6	0-6	1-7	4-10	6-12	9-15
6	0-6	1-7	2-8	6-12	9-15	12-18
7	0-6	2-8	4-10	8-14	12-18	15-21
8	0-6	4-10	6-12	10-16	15-21	18-24
9	4-10	6-12	8-14	12-18	18-24	21-27
10	6-12	8-14	10-16	15-21	21-27	24-30
11	8-14	10-16	12-18	18-24	24-30	27-33
12	10-16	12-18	15-21	21-27	27-33	30-37
13	12-18	15-21	18-24	24-30	30-37	33-41
14	15-21	18-24	21-27	27-33	33-41	37-46
15	18-24	21-27	24-30	30-37	37-46	41-51
16	21-27	24-30	27-33	33-41	41-51	46-57
17	24-30	27-33	30-37	37-46	46-57	51-63
18	27-33	30-37	33-41	41-51	51-63	57-71
19	30-37	33-41	37-46	46-57	57-71	63-78
20	33-41	37-46	41-51	51-63	63-78	70-87
21	37-46	41-51	46-57	57-71	70-87	77-96
22	41-51	46-57	51-63	63-78	77-96	84-105
23	46-57	51-63	57-71	70-87	84-105	92-115
24	51-63	57-71	63-78	77-96	92-115	100-125
25	57-71	63-78	70-87	84-105	100-125	110-137
26	63-78	70-87	78-97	92-115	110-137	120-150
27	70-87	78-97	87-108	100-125	120-150	130-162
28	78-97	87-108	97-121	110-137	130-162	140-175
29	87-108	97-121	108-135	121-151	140-175	151-188
30	97-121	108-135	121-151	135-168	151-188	168-210
31	108-135	121-151	135-168	151-188	168-210	188-235
32	121-151	135-168	151-188	168-210	188-235	210-262
33	135-168	151-188	168-210	188-235	210-262	235-293
34	151-188	168-210	188-235	210-262	235-293	262-327
35	168-210	188-235	210-262	235-293	262-327	292-365
36	188-235	210-262	235-293	262-327	292-365	324-405
37	210-262	235-293	262-327	292-365	324-405	360-life
38	235-293	262-327	292-365	324-405	360-life	360-life
39	262-327	292-365	324-405	360-life	360-life	360-life
40	292-365	324-405	360-life	360-life	360-life	360-life
41	324-405	360-life	360-life	360-life	360-life	360-life
42	360-life	360-life	360-life	360-life	360-life	360-life
43	life	life	life	life	life	life

Zones: Zone A (levels 1–8 region), Zone B, Zone C, Zone D

Source: http://www.ussc.gov/sites/default/files/pdf/guidelines-manual/2014/2014sentencing_table.pdf

D.3   Sensitivity Analysis

The dependent variable is logged in the outcome (sentence) equation for overall fit, but results are qualitatively similar in the level estimation. In the only notable difference, the coefficients on the married and male dummies become significant in the level estimation. The preferred specification drops cases with multiple charges, avoiding both concurrent sentences and the possible correlation among verdicts for a single offender, but no qualitative changes arise from naively treating cases with multiple charges as multiple cases with a single charge each. The preferred specification also recodes any conviction with zero prison time as a sentence of one-half month. Using other reasonable translations can affect the p-value for the coefficient on the discretion regime dummy, but in no case does it become statistically significant. It is worth noting that translations of the zero sentence only affect the outcome equation, not the selection equation.

Sentencing guidelines do undergo minor yearly revisions. The preferred specification omits dummies for guideline year because they are highly collinear with the discretion regime dummy. Their inclusion leaves the signs of the coefficients unchanged but the estimates become imprecise. Including both year and year-district interaction dummies yields similar results.

One problematic feature of the data is the calculation of the criminal history score. The courts only calculate the variable for convicted defendants, leaving criminal history unobservable for acquitted defendants. Following Rehavi and Starr (2014), a hot deck imputation method is a viable strategy for dealing with missing values (Andridge and Little 2010). Defendants with a missing criminal history score are matched to defendants with complete information, based on age, race, gender, marital status, and offense severity. Each missing criminal history score is imputed from a random observation in the pool of sample donors. Table A2 in the Appendix presents

the results for the median average marginal effect of severity in the selection equation, taken from 99 hot deck imputations. All main results are qualitatively unaltered.

Most importantly, the reported marginal effects treat severity, a categorical variable, as continuous. The sign and significance of the severity variable are maintained under log and exponentional transformations, excepting that severity becomes marginally insigficant when exponentiated. The main argument for treating severity as continuous is parsimony. The overall story is much the same when the categories are broken into dummy variables. For example, see Tables B1 and B2, which present marginal effects for a severity base level category equal to 1 or 7. In general, conviction becomes less likely when moving from less to more severe categories, and more likely vice versa.

The marginal effect of the black-male interaction term in Table 2 is not correct. Norton et al. (2004) provide a method for calculating correct values. Figure B1 presents the correct interaction effects, and Figure B2 presents their corresponding z-statistics. The correct values deviate little from the incorrect values, and for no observation is is the z-statistic significant.
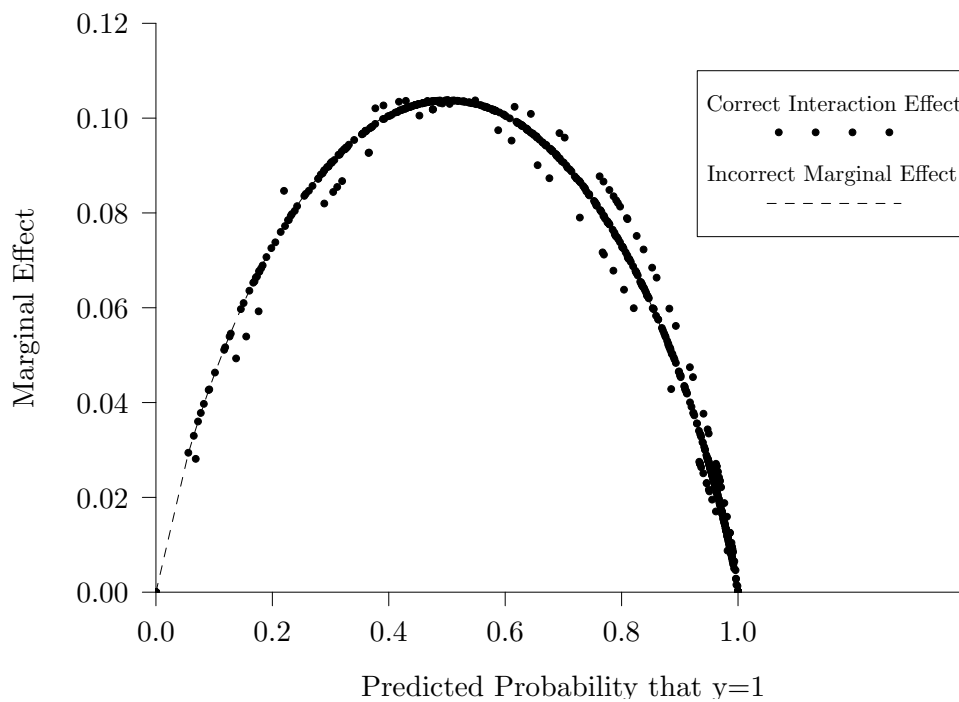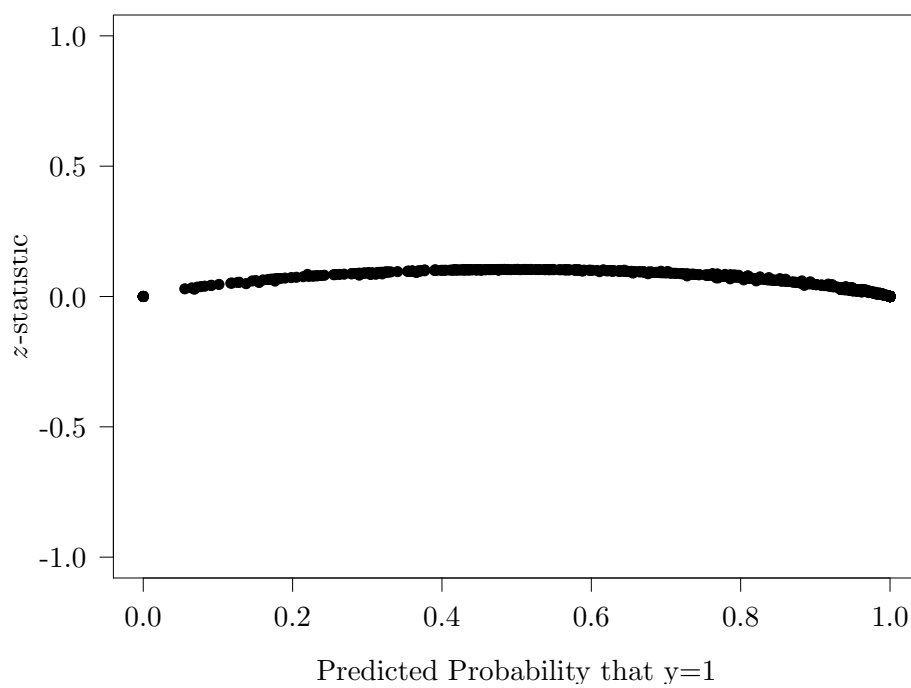
Figure A1: Black-Male Interaction Effects

Figure A2: z-statistics of Interaction Effects

Table D.2: Two-Part Model Results – Hot Deck Imputation of Criminal History

	Conviction (Probit)		Sentence (OLS)	
	marginal effect	p-val	$\dfrac{\partial\mathrm{E}(ln(y)|y>0,x)}{\partial x_j}$	p-val
Severity	-.019	.015	.328	.000
Discretion Regime	.087	.037	-.115	.698
Age	-.012	.031	-.025	.362
Age^2	.000	.021	-.000	.357
Black	-.055	.420	.035	.904
Male	-.017	.709	.175	.417
Black·Male	.060	.421	-.028	.926
Married	-.076	.043	-.171	.275
History Points	-.004	.079	.057	.005
Public	.131	.054	1.519	.000
Private	.129	.123	.980	.010
Appointed	.110	.180	1.410	.000
%White Judge	-1.46	.045	-12.659	.019
%Female Judge	-.770	.153	-4.513	.115
%Democrat Judge	-.142	.368	2.581	.228

Dependent variables: Participation=conviction dummy
Continuous=log(sentence length)

Note: n=765 for stage one; n=576 for stage two; years run from 1998-2010. This table recreates Table 3 using a hot deck imputation procedure for criminal history points. Missing values are filled by a random draw of nearest neighbors as matched on race, gender, age, and offense severity. The values reported correspond to the median marginal effect for the severity variable, based on 99 imputations.

Table D.3: Discrete Marginal Effects - Severity Base 1

	Conviction	
	marginal effect	p-val
0	.034	.213
2	-.161	.001
3	-.103	.147
4	-.380	.012
5	-.366	.000
6	-.481	.000
7	-.430	.000
8	-.252	.001
9	-.910	.000
10	-.129	.118
11	-.341	.000
12	-.542	.000

Note: This table contains average marginal effects from the primary specification in Table 2, with the categorical severity variable broken into dummies. The reference category is "1."

Table D.4: Discrete Marginal Effects - Severity Base 7

	Conviction	
	marginal effect	p-val
0	.464	.000
1	.430	.000
2	.268	.002
3	.325	.000
4	.049	.766
5	.063	.528
6	-.051	.649
8	.178	.050
9	-.481	.000
10	.301	.000
11	.089	.347
12	-.113	.520

Note: This table contains average marginal effects from the primary specification in Table 2, with the categorical severity variable broken into dummies. The reference category is "7."

D.4 Bench and Jury Selection

Any conclusion of whether judges apply a different standard of proof than juries relies on a similar distribution of evidentiary strength over crimes for both bench and jury trials. Figure C1 shows that more serious offenders tend to opt for the jury, but different severity distributions do not imply different evidentiary strengths among offenses of a given severity.

Figure A3: Dependent Variable Distribution: Bench and Jury Trials

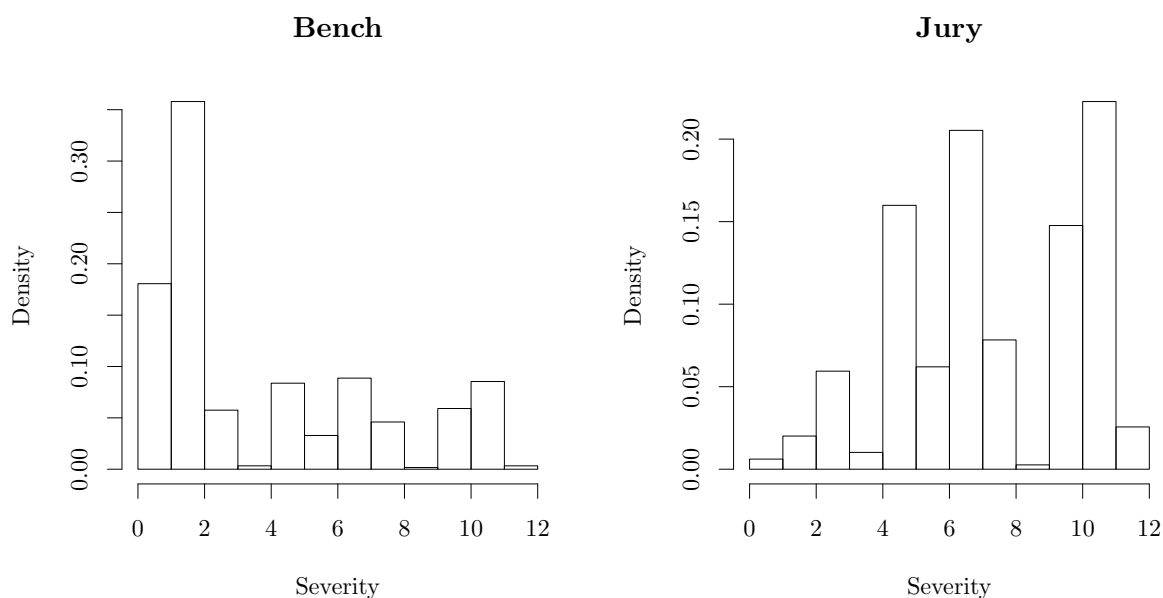**Bench**   **Jury**

Table C1 provides more information on the types of offenders who opt for the bench over the jury. The table contains the average marginal effects of a probit estimation of a bench trial dummy on defendant and judge characteristics. Most charges are pled guilty before trial, which raises the possibility of selection bias. However, a probit model with sample selection fails to reject the null hypothesis that the decision to go to trial and the decision to choose a bench trial are independent equations ($p=0.276$).

Confirming the picture in Figure C1, more severe crimes are significantly less

Table D.5: Probit - Bench Trial

	Conviction	
	marginal effect	p-val
Severity	-.026	.000
Discretion Regime	.047	.014
Age	-.004	.097
Age^2	.000	.253
Black	-.053	.049
Male	-.071	.001
Black·Male	.031	.287
Married	.009	.440
Public	-.046	.247
Private	-.039	.368
Appointed	-.040	.377
%White Judge	-.519	.071
%Female Judge	-.078	.675
%Democrat Judge	.101	.319

Note: n=4,212; years include 1998-2010. The 'marginal effect' column reports average marginal effects. The regression contains unreported dummies for US district. Standard errors are clustered by district. The marginal effect of the black male interaction dummy is not correct (see Norton et al. 2004).

likely to go before the bench. Most other variables are not predictive, but males and blacks are less likely to opt for a bench trial. The question of whether severe crimes with lower evidence opt for the jury cannot be answered by the estimation, but no obvious story emerges for why blacks and males might bring weaker cases to the jury. Furthermore, recall that neither the black nor male dummy was a significant predictor of a conviction decision in a bench trial.

The main results examine separately the conviction tendencies of juries and judges. Table C2 presents the results of a linear probability model pooling both populations together, with an interaction term for a bench trial dummy with the severity of the

crime. The linear probability model avoids the pitfalls of calculating interaction effects in a nonlinear model when one of the variables is ordinal. The results support the main conclusion in treating the trial populations separately. Judges are less likely to convict than juries for every level of severity, but especially so for the highest levels.

Table D.6: Interaction Effects in a Linear Probability Model

Severity	Bench=0		Bench=1	
	coef.	p-val	coef.	p-val
1	-.002	.980	-.142	.130
2	-.133	.015	-.337	.000
3	-.046	.184	-.298	.000
4	-.179	.016	-.437	.043
5	-.085	.006	-.489	.000
6	-.137	.001	-.497	.000
7	-.088	.006	-.454	.000
8	-.028	.439	-.321	.002
9	-.186	.254	-1.037	.000
10	-.007	.838	-.285	.002
11	-.001	.982	-.520	.000
12	-	(omit)	-.735	.003

Note: n=4,212; years run from 1998-2010. The table contains interaction effects from a linear probability model of jury and bench trials. The categorical severity variable is interacted with a bench trial dummy. The omitted category of severity is 0. Standard errors are robust. Defendant and judge characteristics are included but not reported.