

THE ROLE OF PROBABILITY OF DETECTION IN JUDGMENTS OF PUNISHMENT

Jonathan Baron¹, Ilana Ritov

ABSTRACT

In nine experiments—one a questionnaire given to Israeli judges, the rest on the World Wide Web—we examined the effect of probability of detection of an offense on judgments of punishment. When cases differing in probability were separated, we found almost no evidence for attention to probability (as found previously by others). When cases were presented next to each other, however, a substantial minority of subjects took probability into account. Attention to probability was increased in one study by a probe manipulation concerning deterrent effects. We found inconsistent effects of identifying the perpetrator, or of asking subjects to consider policies versus individuals. Some subjects thought that it was unfair to consider probability, but more subjects thought that probability was relevant because of the need for deterrence. We suggest that the failure to consider probability is to some extent an example of the “isolation effect,” in which people do not think much about secondary effects, rather than entirely a result of ideological commitment to a “just deserts” view of punishment.

“To enable the value of the punishment to outweigh that of the profit of the offense, it must be increased, in point of magnitude, in proportion as it falls short in point of certainty.” (Bentham, 1948/1843, Ch. 14, section XVIII).

1. INTRODUCTION

Traditional views of punishment are based on *lex talionis*, the idea of proportional retaliation in kind, which was used extensively (but not to the exclusion of other principles) in the Code of Hammurabi and then in the Old Testament. Presumably, the intuitive justification of punishment is that it balances the scales: If A harms B and then B harms A to the same extent (preferably in the same way), then the two are even. Hence, the idea of “getting even.” In the eighteenth century, Beccaria (1963/1764) introduced

1 This work was supported by a grant from the U.S.-Israel Bi-national Science Foundation. We thank Louis Kaplow, Stephen Shavell, several members of their class at Harvard Law School, Peter Huang and David Hoffman and their class at Temple Law School, Isaac Ehrlich, and Jim Andreoni for comments and specific advice about data. Email: baron@psych.upenn.edu.

the idea that punishment is justified by deterrence; Bentham (1948/1843) extended this idea significantly, stating many of the principles that are now considered part of the economic theory of law (Becker, 1968; Shavell, 2004). In particular, as illustrated in the previous quotation, Bentham recognized that the magnitude of punishment could be increased to compensate for lower probability of punishment, thus maintaining its ability to deter wrongdoing by reducing the overall expected benefit of crime.

2 Of course, punishment has other functions aside from deterrence (Shavell, 2004). It can sometimes allow restitution to victims, if the offender can provide that. When offenders are banished or imprisoned, they are prevented from further crime, an effect that seems to be a major benefit of the criminal justice system in most countries today. In principle, punishments can rehabilitate offenders. Finally, and more controversially, punishment can satisfy the desire for retaliation or retribution. Even if we take a utilitarian view and put aside the claim (advocated by Kant and others) that punishment is inherently required regardless of its future benefits, we must, it is argued, take into account the satisfaction of retribution as a consequence that counts in a utilitarian calculus.

3 Robinson and Darley (for example, 2000) have argued that such satisfaction counts as a major benefit of punishment, largely because the deterrent effects of punishment are in fact minor. The studies they examine, however, are usually concerned with the effects of relatively small modifications in the law. (Even the difference between capital punishment and life in prison can seem relatively small.) Yet, other studies show effects of similar small variation, assessed in other ways. (For a recent example, see Bar-Ilan and Sacerdote, 2004, which shows the effects of both magnitude and probability of punishment. See also Kessler and Levitt, 1999, for a review and additional data.)

4 We do not need to resolve this issue here because we are interested in a general understanding of deterrence. For that purpose, it suffices to observe that the threat of penalties seems to matter: Unsolicited and undesired commercial telephone calls have dropped almost to zero since they were banned in the United States, while unsolicited commercial email, which has not been banned, still thrives. Surely most people are aware of such effects as these and believe that threats of punishment can be effective.

Attention to Probability

5 Our particular interest is in the view of citizens concerning probability of detection. Bentham's idea that magnitude should increase when probability

decreases is not part of the traditional view, so people holding that view might tend to ignore probability. We consider two sorts of reasons for neglect of probability. One is that people explicitly reject the utilitarian view. They might feel that the punishment should match the crime and that nothing should matter except the seriousness of the crime itself. The other reason is that people, while accepting the role of probability in principle, neglect it when it is not called to their attention.

Sunstein, Schkade, and Kahneman (2000) found that manipulation of probability of detection in hypothetical cases about lawsuits had no effect on prospective jurors' judgments of punitive damage awards in lawsuits. Different versions of each case, differing in probability, were presented to different subjects: a between-subject design. Such a design would encourage neglect based on inattention.

In a second study, Sunstein et al. found that most (but not all) of their subjects in a sample of law students, asked directly about the appropriateness of taking probability into account as a matter of policy, objected to this policy on the grounds of unfairness.

Carlsmith, Darley, and Robinson (2002) manipulated probability of detection in a completely between-subject study, in which each subject read a single scenario. Probability had only a very small effect on judgments of punishment, which was found only in subjects who endorsed the idea of deterrence in other questions about their overall views of punishment.

Overview

As noted, people might reject the role of probability and even the role of deterrence. But, even those who think about deterrence and would be open to the idea that probability is relevant can fail to consider probability when it is not a salient property of the situation. Accordingly, in our experiments, we examine two kinds of effects. In one, we manipulate probability across different cases, presented at different times. In the other, we manipulate probability transparently, on the same page, in view at the same time. The transparent test asks whether people think probability is relevant when it is called to their attention.

We also examine the effect of identifying the defendant or the victim. Schelling (1968) pointed out that saving individual lives often has more intuitive appeal than saving more statistical lives. Small and Loewenstein (2003, 2005) found, following Schelling, that people are more generous toward an identifiable victim than toward a statistical victim. A similar effect is revealed

for punitiveness: People are more punitive toward determinate wrong-doers than toward equivalent, but indeterminate, wrongdoers, even when determining the wrongdoer conveys no meaningful information about him. Small and Loewenstein's research addressed the distinction between identifiable and non-identifiable victims. The non-identifiable victim in their studies was not just unidentified, but also as yet undetermined. Their findings clearly show that people are more willing to contribute to a determined victim than to an undetermined one. Small and Loewenstein attribute the identifiability effect they found to stronger affective reactions toward the identifiable target. Kogut and Ritov (2005a,b) found that these reactions intensify even more if the target of the judgment was not just identifiable, but actually identified. This effect, though, was limited to individuals. Identification of the members of groups had little effect.

11 We also examine the role of emotion. Rottenstreich and Hsee (2001; also Hsee and Rottenstreich, 2004) have argued that emotional responses make people less sensitive to quantitative attributes. Possibly, probability is such an attribute.

12 Finally, when we find that people do not take probability into account, we ask whether they would do so if it were called to their attention. This approach is inspired by work on the "isolation effect" (reviewed by McCaffery and Baron, 2006). The general conclusion of this work is that people tend to ignore the secondary or indirect effects of options when evaluating those options. McCaffery and Baron (2006) showed that asking subjects to think of these effects of various public policies could change evaluations of those policies. For example, subjects were less supportive of a tax on businesses when they were asked who would ultimately pay the tax. In general, incentive effects tend to be seen as secondary, and the effect of probability on deterrence would fall in this category.

13 The question of whether people take probability into account is part of a larger issue about the conflict between the intuitive and economic/utilitarian approaches to law. In several other cases, people apply "moral heuristics" (Sunstein, 2005) to questions about legal issues. For example, Baron and Ritov (1993) found that subjects from a variety of sources (including judges and law students) often failed to consider the effects of tort penalties on activity level. In one hypothetical case, a birth-control pill caused sterility in a woman, even though it had been (reasonably extensively) tested and had never caused sterility before. In one version of this case, subjects were told that, if the company lost a lawsuit against the woman, it would

try harder to make a safer pill, even though the current pill was already the safest on the market. In a comparison version presented on the same page, the subjects were told that a loss would lead the company to stop producing the pill, leading to less safe alternatives. Most (but not all) subjects in most groups not only assigned the same tort penalties in the two cases, but also, when asked directly, saw no argument about why anyone would do otherwise. (A group of economics graduate students, however, did make the distinction on the basis of different incentive effects.) In general, then, subjects understood tort penalties in terms of balancing for harm done, rather than in terms of incentives for greater care when care was efficient. When the idea of deterrence was explained to subjects, some of them said that they had not thought of it before. Of these, some thought that it was a good approach and others thought it was unfair. (The numbers varied from study to study.) Baron and Ritov argued that these results pointed to a deficiency in education; given the importance of deterrence in legal theory, subjects should at least have been more familiar with the idea, even if they ended up rejecting it.

In another example, Wilkinson-Ryan and Baron (in press), gave subjects hypothetical scenarios about breach of contract. For example, a person selling a house—who has already moved out—is told that her house will be worth \$10,000 more if she has it painted. She signs a contract with a painter, agreeing to pay him \$6,000. He gets a better job offer (or, in another condition, loses his equipment so that the cost of honoring the contract is now too high) and breaches the contract. Subjects were asked about the penalty for breach. The economic theory of law argues that “expectation damages” are sufficient, for example, \$4,000 (Shavell, 2004). This payment leaves the house seller in the same position she would have been in if the contract were honored. (Recall that her interest is purely financial; she is already out of the house and will not enjoy the new paint job.) And, if the painter expected any higher penalty, he would want prior compensation for bearing the risk of it. Only a few subjects spontaneously chose \$4,000 as the appropriate penalty (although more did so when the painter’s costs increased than when he got a better deal). Many subjects’ comments said that a contract was a promise and that it was morally wrong to break a promise, so the penalty should be higher, sometimes much higher, than \$4,000. Again, we see that a “moral heuristic,” the idea of promise keeping, intrudes.

The last two cases described contrast in an interesting way. The tort example is based on real examples in which beneficial products were removed

from the market, not just birth-control products, but also vaccines (Baron and Ritov, 1993). The contract case, however, does not have such real-world parallels that we know of; apparently, expectation damages are fairly standard in practice (Shavell, 2004). Thus, the law itself is sometimes consistent with moral heuristics and sometimes, it seems, with economic theory.

16 The case of interest here is another potential conflict between a moral heuristic—the idea that the punishment should match the offense—and an implication of utilitarian/economic theory, namely that the punishment should be higher when the probability of detection is lower. In the Appendix, we argue that punishment in the real world is at least somewhat sensitive to probability. We ask here whether human judgments agree. All our studies show the same general pattern as the last two studies described: Many subjects, but by no means all of them, rely on a moral heuristic of matching. Some do so very explicitly, holding that anything else is unfair. In the present case, however, many subjects seem to have labile and malleable intuitions.

17 In most of the rest of this paper, we present several experiments. For the benefit of readers who are not used to reading such experiments, we suggest that many of the details can be skipped. We include them in the text rather than an appendix to make them easier to find when they are needed.

2. EXPERIMENT 1: JUDGES

18 The first study involved determining the sentence of a defendant. We used a 2 (identified/nonidentified defendant) \times 2 (probability of detection) design, with the first factor manipulated between subjects and the second within-subject. Additionally, order of presentation of the two cases was varied between subjects. Subjects were Israeli judges. This experiment is the only one that used judges, and its results are thus relevant to the question of whether judges could somehow compensate for the biases of juries in their sentencing decisions.

Method

19 A sample of 34 judges attending a seminar organized by Israel's ministry of justice responded to a questionnaire. They were presented with the following task (in Hebrew), with the identified condition presented in parenthesis:

Imagine that a defendant [Meir Cohen, age 24, from Beit-Shemesh] is brought before you. The defendant was caught by a police patrol, while attempting to break into a car. The defendant [Meir Cohen] does not have

a criminal record. This is his first known offense. You are aware of the fact that at this time, thanks to the availability of resources and manpower, the police succeeds in catching the thief in 1 out of 50 car-theft cases.

Do you think that the defendant [Meir Cohen] should be convicted?

Yes / No

If yes, what is the sentence you would give? _____ (between 0 to 36 months)

Now imagine another case. Another defendant [Haim Yona, age 22, from Or-Akiva] is brought before you. The defendant was caught by a police patrol, while attempting to break into a car. The defendant [Haim Yona] does not have a criminal record. This is his first known offense. However, you are aware of the fact that at this time, due to budget cuts and shortage in manpower, the police succeeds in catching the thief in 1 out of every 5,000 car-theft cases.

Do you think that the defendant [Haim Yona] should be convicted?

Yes / No

If yes, what is the sentence you would give? _____ (between 0 to 36 months)

Results

Order of presentation did not have an effect on the responses, and data are collapsed across the two orders. Although the questionnaire asked for a response in terms of months in prison, some of the judges set some other punishment (such as conditional imprisonment). We summarize the results in terms of same versus different sentences (within subject) to avoid recoding the responses. Table 1 presents the number of respondents in each cell. As can be seen, 24 of the 34 judges assigned the same sentence to the two defendants. Apparently, they did not consider probability of detection relevant to determination of the sentence. Only 10 judges assigned a more severe sentence when probability of catching the car thief was lower. The sensitivity to probability was not affected by the availability of identifying information regarding the defendant.

20

Table 1. Same/Different Sentence (Rows) by Identified/Unidentified Defendant (Columns)

	Unidentified	Identified	Total
Same sentence	14	10	24
Different sentence	6	4	10
Total	20	14	34

- 21 Although it seems unlikely from these results that we could depend on individual judges to take probability into account in sentencing, the legal system might have other ways of taking probability into account even if juries do not. For example, regional norms for sentences, or even legislative guidelines, might be sensitive to probabilities. (The Appendix provides some relevant evidence.)

3. EXPERIMENT 2

- 22 The remaining experiments were done on the World Wide Web using a questionnaire with several pages. Each page concerned a different offense. In the first of these Web studies, we asked about emotion—specifically, anger—and we asked about probability of getting caught. We could look at the correlation of these probability judgments with judgments of punishment across the 17 offenses. We also manipulated probability of detection within each page by asking the subject to suppose that probability was high (90%) or low (10%).

Method

- 23 The 82 subjects ranged in age from 22 to 70 (median 43); 29% were male. Aside from this imbalance in sex, subjects in this study and those that follow were typical of the U.S. population in income and education (as assessed in other studies with the same population), although some were not U.S. residents. Subjects were paid for each study (through PayPal). To be paid, they had to provide full identification associated with a single email address. Most subjects found our studies by searching for ways to earn money on the Web. A panel of about 1,400 subjects was divided into three groups so that similar studies could be sent to different groups. Potential subjects were notified by email when a study was ready, and the study was removed when about 75 responses had been received, thus aiming for 80 in total.
- 24 Each trial consisted of a web page. The experiment was controlled by a JavaScript program that checked responses and record times (without informing the subjects that times were recorded). The introduction to this study read:

This is about the seriousness of computer abuse of various kinds. Some of these behaviors are now legal, and it is possible that they will remain legal even if some people object to them.

Some of the questions concern the probability of getting caught. In order to be “caught,” the person responsible must be identified so that punishment is possible.

We also ask about punishment. It is possible to punish behaviors with both fines or prison. You might think that a behavior should get a fine but not prison, prison but not a fine, both prison and a fine, or neither.

The 17 cases were presented in an order randomized for each subject. A typical page was 25

Breaking into a bank's database and stealing records, in order to steal money by pretending to be someone else (identity theft).

How angry does this behavior make you?

Not at all A little Moderately Very angry Furious

This question and the next concern penalties: fines and prison. Suppose the perpetrator had \$2,000,000, half of which came from illegal activity.

What fine should he pay? (Pick the closest to what you think.)

\$0 \$5,000 \$50,000 \$500,000 \$1,000,000 \$1,500,000 \$2,000,000

In addition to the fine, what prison term should he serve?

None 1 year 5 years 10 years 25 years Life

Many people do things like this. What proportion of them do you think get caught eventually?

0% 5% 10% 20% 40% 80% 100%

Suppose that, through technical means, it became very easy to catch people who did this, with 90% getting caught eventually. What fine should the perpetrator pay?

\$0 \$5,000 \$50,000 \$500,000 \$1,000,000 \$1,500,000 \$2,000,000

In addition to the fine, what prison term should he serve?

None 1 year 5 years 10 years 25 years Life

Suppose that, through technical means, it became very easy to avoid getting caught, with 10% getting caught eventually. What fine should the perpetrator pay if he is caught?

\$0 \$5,000 \$50,000 \$500,000 \$1,000,000 \$1,500,000 \$2,000,000

In addition to the fine, what prison term should he serve?

None 1 year 5 years 10 years 25 years Life

Please write any comments on this page here (up to 255 characters):

26 The cases were

- Maintaining a political web site for the purpose of attacking a particular ethnic group with derogatory and false statements.
- Tricking an 11-year-old girl into posing nude for pornographic pictures on a web site (where viewers must pay), without the girl's consent.
- Tricking a 21-year-old female student into posing nude for pornographic pictures on a web site (where viewers must pay), without the student's consent.
- Operating a profit-making web site to sell pornographic pictures of children.
- Operating a profit-making web site to sell pornographic pictures of women.
- Operating a profit-making web site for customers to bet on sports events.
- Breaking into an office computer and having it send spam about prescription drugs from its user's email address to 1,000,000 other addresses.
- Sending spam about prescription drugs from thousands of forged email addresses to 1,000,000 other addresses.
- Creating a virus that wipes out files on people's computers, spreads to other computers, and does nothing else.
- Attacking a company's web site so that customers cannot use it, out of anger at the company.
- Breaking into a hospital's database and examining patient records, out of curiosity.
- Breaking into a bank's database and stealing records, to steal money by pretending to be someone else (identity theft).
- Breaking into a home computer, in a home with two young children, and programming it to present pornographic images at random times.
- Breaking into a home computer, in a home without children, and programming it to present pornographic images at random times.
- Downloading music without paying, violating its copyright.
- Making music available free for download, in violation of its copyright.
- Setting up a false web site that looks like a bank's site and then sending spam to entice the bank's customers into revealing their PINs.

Results

27 The two questions about punishment behaved similarly, so we simply treated them as equally spaced and added them together to form an overall index. The mean judgment for fines was 4.9 (s.d., 1.02 across cases, 1.20

across subjects) on the 1–7 scale, and the mean for prison was 3.1 on the 1–6 scale (s.d., 0.95 across cases, 0.87 across subjects). The mean of their sum, our measure of punishment, was 8.0 (s.d., 1.96 across cases, 1.80 across subjects).

Subjects' punishment judgments differed in their sensitivity to probability in the within-page manipulation of probability. Thirty-five subjects (43%) assigned more punishment, on the average, to the low-probability version; 13 (16%) assigned more to the high-probability version; and 34 (41%) assigned equal punishment. Thus, a substantial minority was sensitive to probability in the normatively correct direction. The mean difference was only 0.14 out of a maximum possible difference of 11, although this difference was significant across subjects ($t_{81} = 3.66, p = 0.0004$).

The mean probability assigned to getting caught was 0.15. Across cases, the correlation between mean punishment and mean probability was *positive*, $r = 0.88$ ($p = 0.0000$ with 15 df), opposite to the normative prediction. However, it is likely that more serious offenders *are* more often caught. (See the Appendix.) That said, this positive correlation between probability and punishment is higher in the subjects who did not respond normatively to probability within pages (for example, higher penalties for 10% caught than for 90%). We computed the correlation between probability and punishment for each subject. This within-subject correlation itself was correlated negatively across subjects with the within-subject probability effect, the difference between the 90% and 10% examples, $r = -0.34$ ($p = 0.0020$, 79 df; one subject gave the same probability to all cases). Thus, it appears that subjects who are willing to take probability into account within a page are also sensitive to probability across pages; this sensitivity reduces what would otherwise be a high positive correlation between probability and punishment.

Anger was highly correlated with punishment, with a mean within-subject correlation of 0.78. This correlation, however, was not correlated with the within-subject probability effect ($r = -0.01$). Nor was anger itself correlated with the within-subject probability effect; the across-subject mean of the within-subject correlation across cases between anger and the probability effect was 0.03. Thus, we have no evidence for the view that anger reduces attention to probability.

In sum, this experiment shows that a substantial minority of subjects attend to probability when it is transparently manipulated, and these subjects also seem to attend to it when it varies from case to case.

4. EXPERIMENT 3: FALSE IDENTITY

In Experiment 2, we failed to find a negative correlation across cases between judged probability of getting caught and penalties. We speculated that the positive correlation we found was a result of properties of the cases—namely, a natural correlation between seriousness and probability of punishment. In Experiment 3, we attempted to manipulate the probability of getting caught while holding the crime constant. Specifically, we manipulated whether the wrongdoer tries to hide her identity.

We also asked additional questions about victim responsibility, fairness, and deterrence. The question about fairness helps to determine whether subjects are rejecting the role of probability because they think it is unfair.

Method

The 91 subjects ranged in age from 22 to 90 (median 40); 32% were male.

The 10 cases were modified from some of those used in Experiment 2. Each was presented in two forms, with and without hidden identity, for a total of 20 pages, in random order. Here is an example:

Breaking into office computers and having them send spam about prescription drugs from their users' email addresses to 1,000,000 other addresses.

The perpetrator has not hidden his identity.

[He has used a false identity to do this, so he will be hard to find.]

How angry does this behavior make you?

Not at all A little Moderately Very angry Furious

Suppose the perpetrator has just enough assets to pay the maximum \$1,000,000 fine. What should the punishment be? (Pick the closest to what you think.)

- None • \$3,000 fine • \$10,000 • \$30,000 • \$100,000 • \$250,000
- \$500,000 • \$1,000,000
- \$1,000,000 plus 1 year in prison
- \$1,000,000 plus 3 years in prison
- \$1,000,000 plus 5 years in prison
- More than \$1,000,000 or more than 5 years

Many people do things like this, in this way. What proportion of them do you think get caught eventually? (Pick the closest.)

1% 5% 10% 20% 40% 80% 100%

To what extent are victims responsible for this sort of crime?

Not at all Somewhat Largely Completely

In setting a penalty for this case, how do you consider the chance of getting caught?

- It is not relevant.
- It is unfair if the penalty is much higher for those who happen to be caught.
- If the chance of getting caught is low, then the punishment must be greater to deter potential perpetrators.
- Both fairness and the need to deter are relevant.

The “unfairness” item was poorly worded because it could apply regardless of probability of getting caught. We report the results here, but subsequent experiments report results from better versions. 36

Results

The positive correlation between probability of getting caught and fine (considered as a linear scale) was found, although it was not significant across the 10 cases collapsing over hidden identity, $r = 0.49$. The correlation with anger was high and significant ($r = 0.97$, across the 10 cases). 37

As hoped, the manipulation of hidden identity also affected fines: The mean fine on our 1–12 linear scale was 8.90 with hidden identity and 8.74 without it ($t_{90} = 2.41$, $p = 0.0179$, across subjects). Hidden identity also affected probability of getting caught: 19.5% versus 26.5% ($t_{90} = 5.28$, $p = 0.0000$). 38

We asked whether the effect of hidden identity on probability mediated its effect on fines. In addition to the two correlations just reported, it would be helpful to show that lower probability increased fines directly even when hidden identity is held constant. We could not ask this within subjects because this correlation was positive overall, which is the wrong direction for showing mediation but is clearly caused by other factors. However, subjects differed in both effects of hidden identity—on fine and on probability—and these two effects are correlated across subjects: $r = 0.18$ ($p = 0.0445$ one tailed). Thus, subjects who showed a larger effect of hidden identity on probability also showed a larger effect on fines. This supports a mediation hypothesis, but less directly than usually done. 39

The last question asked about the relevance or irrelevance of the chance of getting caught. Over all cases, 36% of the responses were that probability is irrelevant, 29% that both unfairness and deterrence are relevant, 32% that deterrence is relevant but not unfairness, and only 3% that unfairness 40

is relevant but not deterrence. We conclude tentatively that most subjects do not think it is unfair to take probability into account, but we shall address this further in other experiments.

5. EXPERIMENT 4: PROBES

41 This experiment followed the approach of McCaffery and Baron (2006), finding that a probe to encourage subjects to consider an issue that they might not have considered is effective. We note here that two previous experiments, not otherwise reported, failed to find an effect of probing. These previous experiments used Internet offenses, like Experiments 2 and 3. The present experiment uses tax offenses, in hopes that the idea of computing the expected value of an offense would be more obvious than the idea of computing some sort of expected utility for committing Internet offenses. We did not, however, carry out a formal test of the difference between types of situations.

Method

42 The 85 subjects ranged in age from 22 to 73 (median 41); 22% were male. Two additional subjects' data were eliminated on the basis of very fast times for completing the questions (outliers in the distribution). The 12 crimes used were

- A carpenter asks to be paid in cash and does not report half of his income.
- A waitress in a small tavern does not report half of her tips.
- A waitress in a fancy restaurant does not report half of her tips.
- A retired person with a pension does questionnaires for pay on the Internet and does not report her annual income of \$1,000.
- A writer takes a \$2,000 deduction for a home office that does not actually exist.
- An executive deducts all of her \$10,000 of travel expenses as business expenses, even though \$8,000 was actually vacation.
- An employee of a large company, paid \$100,000/year, fails to file a tax return.
- A psychiatrist in private practice, earning \$100,000/year, fails to file a tax return.
- A U.S. citizen fails to report income from foreign investments that are not otherwise reported to the U.S. government.
- A wealthy executive contributes to a fake charity that actually funnels money back to his family.

- A professional gambler does not report \$5,000 of winnings from playing backgammon as income.
- An accountant embezzles \$100,000 from his company and does not report this on his tax return.

The 12 crimes were presented twice in the same random order, chosen for each subject. On either the first or second presentation of each crime, a probe item was included. The questions, with the probe item in brackets, were (with the carpenter item as an example)

43

A carpenter asks to be paid in cash and does not report half of his income.

How angry does this behavior make you?

Not at all A little Moderately Very angry Furious

If the offender is caught, what should the fine be? (Pick the closest to what you think.)

- No penalty; pay the amount owed only.
- The penalty should equal the amount owed (total payment, twice the amount).
- Twice the amount owed (total payment, 3× amount).
- Four times the amount owed (total payment 5× amount).
- Eight times the amount owed (total payment 9× amount).
- More than eight times the amount owed.

Many people do things like this. What proportion of them do you think gets caught each year? (Pick the closest.)

1% 5% 10% 20% 40% 80% 100%

[PROBES: If the chance of getting caught goes down and the fine stays the same, what happens to the average cost of this violation?

- It would go down, possibly to the point where the violation is profitable on the average.
- It would go down, but never so much that the violation is profitable on the average.
- No change.
- It increases.

If the average cost of this violation goes down, what will happen to the number of people who try it?

Fewer will try. No effect. More will try.]

Suppose that, through technical means, it became very easy to catch people who did this, with 90% getting caught each year. What should the fine be?
[Same choices as previously.]

Suppose that it became very difficult to catch people who did this, with only 1% getting caught each year. What should the fine be?
[Same choices as previously.]

Please write any comments on this page here (up to 255 characters):

Results

44 As found in Experiment 2, a minority of responses assigned higher fines for a lower probability of detection in the within-page comparison. The probe, however, increased the proportion of these responses. When the probe came in the first half of the cases, the proportion with higher fines was 15.7% in the first half and declined to 10.9% in the second half when the probe was absent. When the probe came in the second half, the proportion increased from 18.0% in the first half to 22.5% in the second half. (Evidently, the subjects assigned to the two orders were not well matched, although the group difference in overall proportion was not significant.) Overall, the effect of the probe was significant, as determined from its effect on the mean difference (0.02 steps on the response scale, treating fine as a linear scale; $t_{84} = 2.47$, $p = 0.0157$).

45 This probe effect was larger for subjects whose responses to the two probe questions were more accurate (with accuracy defined as the negative sum of the steps away from the answers taken as correct; $r = 0.27$, $p = 0.0139$, 83 df). Overall, 30% of the answers to the first probe question were fully correct (the average cost could go down to the point where the violation is profitable; however, 46% indicated that the cost would go down, as indicated by the first two possible answers) and 67% of the responses to the second question were correct (more will try).

46 Also as found in Experiment 2, probability of detection correlated positively with fine (considered as a linear scale; mean within-subject correlation $r = 0.24$) and anger also correlated with fine (mean $r = 0.77$).²

2 We also examined, as in Experiment 2, the correlation across subjects between the within-subject correlation between probability and fine and the within-page effect of probability; this was positive ($r = 0.16$), opposite to that found in Experiment 2, but not significant. Note that the correlation between probability of getting caught and punishment was much higher in Experiment 1, possibly because the offenses were more varied.

As in Experiment 1, anger did not reduce the effect of probability. In fact, anger was positively correlated with the within-page effect of probability, but this seemed to result from a floor effect on fines when anger was low.

47

6. EXPERIMENT 5: ADVISORS

The understanding of probability shown in Experiment 4 might have resulted from the perception that the taxpayer was not infringing but rather as simply trying to maximize. To separate infringement from personal gain, the present experiment compared the position of the taxpayer with that of a tax advisor. The advisor would not benefit directly from the taxpayer's infringement but would be held responsible for it.

48

Method

The 82 subjects ranged in age from 22 to 79 (median 42); 35% were male. The items were those used in Experiment 4, with some minor editing. The probe cases were replaced with advisor cases. The following is an example, with the possible answers omitted when they were the same as in Experiment 4:

49

An executive paid \$400,000/year deducts \$50,000 in contributions to a fake charitable trust that actually funnels money back to his family.

This was the result of a letter from a tax advisor. The letter is with the taxpayer's records and will be discovered if the records are audited. The tax advisor has been paid a flat fee for advice and thus does not profit directly from the reduction in taxes. The taxpayer, if caught, must pay what is owed, but only the advisor is held responsible. Any penalty must be paid by the advisor.

How angry does the tax advisor's behavior make you?

[Same answers as in Experiment 4.]

If the tax advisor is caught, what should the fine be? (Pick the closest to what you think.)

- No penalty for the tax advisor.
- The penalty should equal the amount owed by the taxpayer.
- Twice the amount owed.
- Four times the amount owed.
- Eight times the amount owed.
- More than eight times the amount owed.

Many tax advisors do things like this. What proportion of them do you think get caught each year? (Pick the closest.)

[Same choices as previously.]

Suppose that, through technical means, it became very easy to catch tax advisors who did this, with 90% getting caught each year. What should the fine be?

[Same choices as previously.]

Suppose that it became very difficult to catch tax advisors who did this, with only 1% getting caught each year. What should the fine be?

[Same choices as previously.]

Results

50 The mean proportion of cases in which the responses assigned higher fines for a lower probability of detection in the within-page comparison was 26.1% in the advisor condition and 25.5% in the control condition—no difference. (There was also an order effect in which the proportion declined from the first half to the second half. This is not relevant to our concerns here because of the counterbalancing.) Again, a minority of subjects made the distinction, but the proportion was, if anything, slightly higher than in Experiment 4.

51 Probability of detection again correlated positively with fine (considered as a linear scale; mean within-subject correlation $r = 0.29$) and anger also correlated with fine (mean $r = 0.75$). The correlation across subjects between the within-subject correlation between probability and fine and the within-page effect of probability was negative ($r = -0.11$) but not significant. Again, anger was positively correlated with the within-page effect of probability, but this correlation again appeared to result from a floor effect.

52 The main result, then, is that subjects who take probability into account seem to view it as punishment for an infraction, not as part of a simple economic balancing. The advisor does not benefit from the infraction, so the punishment does not simply undo the benefit.

7. EXPERIMENT 6: INDIVIDUAL VERSUS POLICY

53 Experiment 1 found no effect of identifying the offender. The present experiment examines the effect of identification using a more extreme manipulation, one that actually confounds identification with the number of offenders. Kogut and Ritov (2005a,b) found that their identification effect was actually an interaction between identification and uniqueness of the

needy individual. Thus, we might expect a stronger effect with this confounding. In particular, we might expect that the focus on an identified individual would elicit a framework of just deserts, focusing on the act rather than the probability of detecting it. The group presentation, on the other hand, would elicit economic thinking, thus leading subjects to attend more to probability of detection. As Carlsmith et al. (2002, p. 293) put it, "Although [people] may support deterrence at some societal level, when it comes to actually sentencing specific offenders who have committed specific offenses, they are cued toward the intuitive tenets of just deserts."

Identification was further confounded with future versus past. The identified individual had already committed the offense, but the question about the group concerned a policy to be implemented, hence affecting the future. Again, the past perspective would, we expect, promote a concern with just deserts. Policies are for the future.

We also attempted to manipulate the probability of detection explicitly, as in Experiment 3. Instead of hiding the identity of the offender, we changed the situation so as to make the offense inherently detectable or difficult to detect. For example, it is easier for a waitress to avoid detection for not paying taxes on tips when customers pay in cash than when they pay with credit cards.

Method

The 83 subjects ranged in age from 23 to 74 (median 45); 24% were male.

The items were in five groups, with four cases in each group. The following shows the cases, but only the first group is complete. The four cases varied in unique individual versus policy (alternating, here) and in detectability (the first two are more detectable and the last two less detectable). For groups 2–5, the less detectable version is shown in brackets and the policy version is not shown.

1. A waitress in state X did not report half of her tips. The amount not reported was \$25,000. Most customers paid with credit cards.

Waiters and waitresses in state X do not report half of their tips. The amount not reported averages \$25,000. Most customers in X pay with credit cards.

A waitress in state Y did not report half of her tips. The amount not reported was \$25,000. Most customers paid with cash.

Waiters and waitresses in state Y do not report half of their tips. The amount not reported averages \$25,000. Most customers in Y pay with cash.

2. A retired person with a pension did questionnaires for pay on the Internet and did not report his income of \$2,000. The researchers have her Social Security number and reported payments to the government. [They have only her email address and did not report payments to the government.]
3. An executive deducted all her travel expenses as business expenses, even though 50% were actually vacation. The executive had to provide detailed receipts to the company's business office. [She did not need to provide receipts.] The travel expenses were \$10,000. Her income was \$200,000.
4. A psychiatrist who works for a large health maintenance organization [in private practice] failed to file a tax return. He earned \$100,000.
5. A U.S. citizen failed to report income from U.S. investments that were reported to the U.S. government by others [from foreign investments that were not otherwise reported to the U.S. government]. His income was \$100,000, and the amount of unreported income was \$10,000.

58 The individual and policy items were further distinguished by an introductory statement on each page: "This item is about how future offenses should be penalized" or "about an offense already committed." The questions were the same as the no-probe conditions in Experiment 5, except that the two within-subject probability levels (1% and 90%) were counterbalanced across subjects. (The order had no significant effect and is ignored henceforth.) The 20 cases were presented in a random order chosen separately for each subject.

Results

59 Overall, our hypothesis was roundly disconfirmed. We predicted, if anything, an interaction between policy (versus individual) and probability, such that fines would be higher in the policy condition when probability was low but the effect of probability would be small or nonexistent in the individual condition. Instead, we found a significant interaction in the opposite direction ($t_{82} = 2.29$, $p = 0.0244$). In the policy condition, the mean fine (on the linear scale) was higher when probability was high than when it was low (1.38 versus 1.30; $t_{82} = 2.69$, $p = 0.0087$). In the individual condition, the fine was higher when probability was low (1.36 versus 1.33; $t_{82} = 0.70$, n.s.). However, the equivalent effect for the within-subject test of probability was not significant

($t_{82} = 0.88$, n.s., the effect of individual/policy on the difference between the two within-subject probabilities). The within-page high/low probability effect was 0.139 for policy and 0.157 for individual (both $p < 0.0001$ but not significantly different).

The between-page manipulation of probability had its intended effect on the judged probability of being caught ($t_9 = -5.53$, $p = 0.0004$, across the 10 cases), but it had no effect on fines. If anything, the effect was slightly in the wrong direction ($t_9 = -1.23$, $p = 0.2509$). (Nor did it correlate across subjects with the within-page probability effect: $r = 0.05$.) This is a clear demonstration that, in a between-page manipulation, subjects seem not to consider probability in punishment decisions. This result, however, seems to conflict with the effect of false identity in Experiment 3.

8. EXPERIMENT 7

This experiment was another attempt to examine the effect of identifying the perpetrator.

Method

The eight cases were taken from Experiment 3, with a few edits for clarity. The within-page questions specified that offenders knew the probabilities. (This was not explicitly stated in previous studies. Subjects might have assumed that the offenders did not know.) Each case appeared in an identified and a non-identified version.

The 82 subjects ranged in age from 23 to 68 (median 42); 79% were female. A typical item was (with the non-identified version in brackets)

June broke into a college admission system and changed her status from rejected to admitted.

[Breaking into a college admission system and changing one's status from rejected to admitted.]

How angry does this behavior make you?

Not at all A little Moderately Very angry Furious

Suppose the perpetrator has just enough assets to pay a \$1,000,000 fine.

What should the punishment be? (Pick the closest to what you think.)

- None • \$3,000 fine • \$10,000 • \$30,000 • \$100,000 • \$250,000
- \$500,000 • \$1,000,000

- \$1,000,000 plus 1 year in prison
- \$1,000,000 plus 3 years in prison
- \$1,000,000 plus 5 years in prison
- More than \$1,000,000 or more than 5 years

Many people do things like this. What proportion of them do you think get caught eventually? (Pick the closest.)

1% 5% 10% 20% 40% 80% 100%

When people think about doing this, and think about what might happen, how would they answer the last question?

1% 5% 10% 20% 40% 80% 100%

Suppose that, through technical means, it became very easy to catch people who did this, with 90% getting caught each year, and those who think about doing this sort of thing know that 90% get caught. What should the penalty be?

[Same choices as previously.]

Suppose that it is very difficult to catch people who do this, with only 1% getting caught each year, and those who think about doing this sort of thing know that only 1% get caught. What should the penalty be?

[Same choices as previously.]

Results

64 Identifying the offender led to a greater effect of probability in the within-page comparison: Means for identified perpetrators were 9.08 and 8.99 (on our response scale) for probabilities of 1% and 90%, respectively, and 9.24 and 9.27 for unidentified perpetrators. Evidently, probability of detection had the intended effect only for identified perpetrators. The interaction was significant ($t_{681} = 4.07, p = 0.0001$). Again, this result is in the opposite direction of our hypothesis for Experiment 4. Evidently, identifying the perpetrator does not inhibit people from considering probability, and asking about policy in the abstract does not promote the use of probability.

65 Identification had no significant effect on anger, judged probability of getting caught, or penalty when no probability was provided. It did not affect the within-subject correlation between judged probability of getting caught and fined, which averaged 0.05 across all 16 cases (not significant in this study). The question about what others think was usually answered the same as the question about probability.

66 One subject's comment justified an effect of probability opposite to the one we have been seeking: "If you know 1% of people get caught each

year, then you do it because you think chances are you can get away with it and have the ‘everybody’s doing it’ attitude. But if you know 90% of people get caught each year, then you do it because you don’t care and think you are too smart to be caught. Different reasoning and intent for breaking the law can mean different punishment.” This subject consistently gave more punishment for 90% than 1%, but less so when the offender was identified. A few others showed this pattern, but it did not account for the overall effect of identification, which held throughout the range. In the next study, we added an additional question to get at the state of mind of the perpetrator. Subjects might have been thinking more generally about *Menes rea* as a criterion for penalties.

9. EXPERIMENT 8

Experiment 8 used the same cases as Experiment 7 but changed the order of the questions, asking all questions after each probability and dispensing with the questions that leave probability undetermined. As we note, these changes apparently eliminated the effect of identification (although we cannot be certain that this was the critical difference). More importantly, we added questions about fairness, deterrence, and state of mind. We also extended the possible penalties because some subjects complained about not being able to go high enough for some crimes in Experiment 7.

Method

The 88 subjects ranged in age from 23 to 63 (median 42); 65% were female. A typical item was (with the non-identified version in brackets)

Robin set up a false web site that looked like a bank’s site, and then sent spam to entice the bank’s customers into revealing their PINs.

Suppose that it is very difficult to catch people who do this, with only 1% getting caught each year, and those who think about doing this sort of thing know that only 1% get caught. What should the penalty be?

- None • \$3,000 fine • \$10,000 • \$30,000 • \$100,000 • \$250,000
- \$500,000 • \$1,000,000
- \$1,000,000 plus 1 year in prison
- \$1,000,000 plus 3 years in prison
- \$1,000,000 plus 5 years in prison

- \$1,000,000 plus 10 years in prison
- \$1,000,000 plus 20 years (or more) in prison

How would you describe Robin's state of mind, doing this under these conditions?

- Not morally reprehensible
- Somewhat reprehensible with respect to intention and motive
- Moderately reprehensible
- Extremely reprehensible
- Inhuman

How angry does Robin's behavior make you?

[usual response options]

Reminder: Robin set up a false web site that looked like a bank's site, and then sent spam to entice the bank's customers into revealing their PINs.

Suppose that, through technical means, it became very easy to catch people who did this, with 90% getting caught each year, and those who think about doing this sort of thing know that 90% get caught. What should the penalty be?

[Same choices as previously.]

How would you describe Robin's state of mind, doing this under these conditions?

[Same choices as previously.]

How angry does Robin's behavior make you?

[usual response options]

In setting a penalty for this case, how do you consider the chance of getting caught?

- It is not relevant.
 - It is unfair if the penalty depends on the chance of getting caught.
 - If the chance is low, the penalty must be greater, to deter offenders.
 - The state of mind of the offender depends on the chance of getting caught, so the penalty should differ.
- Both unfairness (B) and the need to deter (C) are relevant.
 - Both unfairness (B) and state of mind (D) are relevant.
 - Both need to deter (C) and state of mind (D) are relevant.
 - All three issues (B, C, and D) are relevant.

Results

Identification did not affect the within-page penalty difference between the two probabilities ($t_{87} = 0.27$). It did, however, affect the *sum* of the fines, with identified perpetrators getting lower penalties ($t_{87} = -3.14$, $p = 0.0023$). Similarly, subjects were less angry with identified offenders ($t_{87} = -3.26$, $p = 0.0016$), but identification had no effect on state-of-mind judgments.³ Thus, subjects were attending to identification, but it did not have the same effect as in the last experiment.

Of greater interest here were the within-page effects of probability of getting caught. As found previously, probability affected penalties ($t_{87} = 5.34$, $p = 0.0000$), with means of 9.69 for 1% probability and 9.40 for 90%, on the 1–12 scale of responses above, where the “none” is 1 and “\$1,000,000 plus 1 year in prison” is 9. But probability had no significant effect on judgments of anger or state of mind.

The last question yielded essentially three measures, with the following proportions: endorsement of deterrence as a reason for taking probability into account, 37%; endorsement of unfairness as a reason against it, 27%; and endorsement of state of mind as relevant (without its direction specified), 22%. Across subjects, endorsement of deterrence was correlated with the effect of probability on penalties ($r = 0.36$, $p = 0.0007$), but no such correlations were present for endorsement of unfairness or of the role of mental state. Thus, belief in unfairness does not seem to account for individual differences in the effect of probability on penalties.

10. EXPERIMENT 9

This experiment made one more attempt to manipulate the perception of external factors that would affect probability of detection, in particular, changes in the rate of offenses over time. When the rate of offenses is increasing, people might see more need for severe penalties, especially when the rate of convictions do not increase as well. We compared past increases to expected future increases. The general idea of this manipulation came from Tetlock et al. (2006).

We also tried other questions about the role of fairness and deterrence.

3 The effect on penalties is opposite to that found by Small and Loewenstein (2005). Identification could have two opposing effects; sometimes one dominates, and sometimes the other dominates. In our experiment, identification could lead to sympathy for the offender.

Method

74 The 79 subjects ranged in age from 23 to 66 (median 45); 75% were female. (One subject was omitted because of apparent random responding.)

75 The crimes were the same as in the last experiment, using the non-identified form only. The following are examples of the two types of items (past and future). The possible answers were the same as in the last experiment, except for the last question. Notice that all four possibilities were used: low to high probability of detection; high to low; remaining low; and remaining high. The order of questions being compared was counterbalanced.

Future Condition

Breaking into a bank's computer system and stealing money from the bank's customers by electronic transfer.

How angry does this behavior make you?

Case A: Now, this is very easy to detect, with 90% of the offenders getting caught each year. But it is expected that next year, because of a change in the way it is done, it will become very difficult to detect (10% caught each year), so the amount of it will increase. What should the penalty be?

Case B: Now, this is very difficult to detect, with 10% of the offenders getting caught each year. But it is expected that next year, because of better detection methods, it will become very easy to detect (90% caught each year), so the amount of it will decrease. What should the penalty be?

Case C: This is now very difficult to detect, with 10% of the offenders getting caught each year. In the future, it will still be difficult. What should the penalty be?

Case D: This is now very easy to detect, with 90% of the offenders getting caught each year. In the future, it will still be easy. What should the penalty be?

In setting a penalty for this case, how do you consider the chance of getting caught?

- It is unfair if the penalty is higher when the chance is lower.
- If the chance is lower, the penalty must be greater, to deter offenders.
- Both unfairness and the need to deter are relevant (and must be balanced against each other).
- Neither unfairness nor deterrence is relevant for these reasons.

Past Condition

Operating a profit-making web site to sell pornographic pictures of children.

How angry does this behavior make you?

Case A: In the past, this was very easy to detect, with 90% of the offenders getting caught each year. Now, because of a change in the way it is done, it has become very difficult to detect (10% caught each year), so the amount of it has increased. What should the penalty be?

Case B: In the past, this was very difficult to detect, with 10% of the offenders getting caught each year. Now, because of better detection methods, it has become very easy to detect (90% caught each year), so the amount of it has decreased. What should the penalty be?

Case C: This has always been very difficult to detect, with 10% of the offenders getting caught each year. It is still difficult. What should the penalty be?

Case D: This has always been very easy to detect, with 90% of the offenders getting caught each year. It is still easy. What should the penalty be?

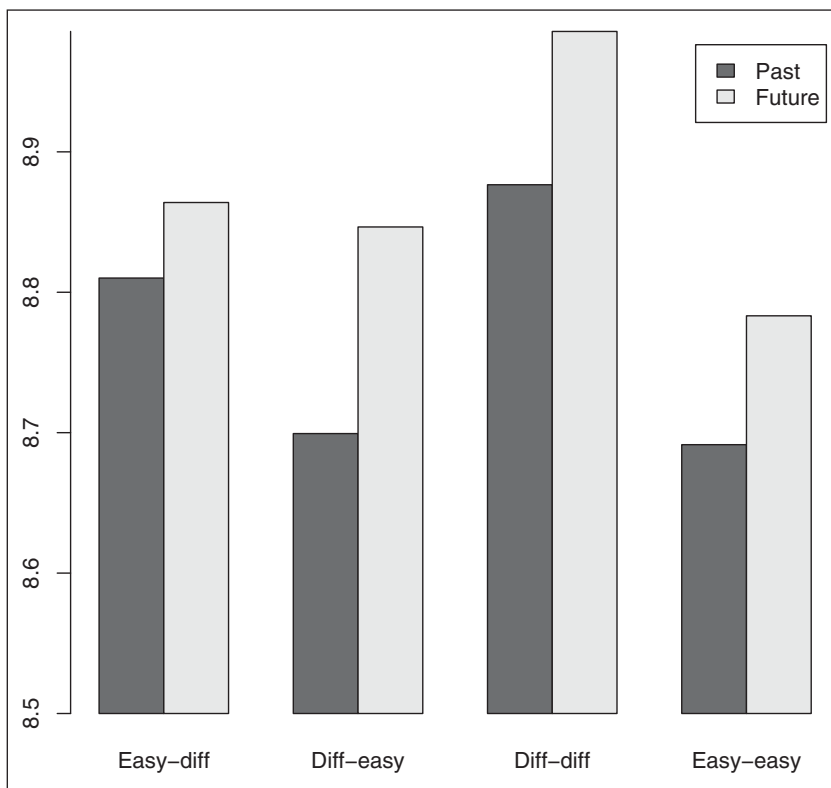
[The same question about the chance of getting caught was used.]

Results

Mean penalties for the four conditions, on the 1–12 scale, are shown in Figure 1. Subjects assigned higher penalties when detection was difficult than when it was easy ($t_{78} = 5.00, p = 0.0000$). Subjects appeared to assign higher penalties for future than for past, but they were so variable that this difference was not significant. However, they did assign reliably higher penalties when conditions were getting worse (Easy-diff) than when they were getting better (Diff-easy), regardless of past/future ($t_{78} = 2.17, p = 0.0331$). This result supports the general idea that worsening conditions can lead to higher penalties when these lead to lower probabilities of detection.

The last question yielded endorsements of unfairness (from making penalties depend on the probability of detection) and of deterrence as relevant to penalties. The mean percent endorsements were 41% for deterrence and 35% for unfairness. Again, deterrence was more often endorsed than was unfairness. And, again, individual differences in endorsement of deterrence were correlated with the tendency to assign higher penalties when detection was difficult ($r = 0.20, p = 0.0790$).

Figure 1. Mean penalties on the 1-12 scale for Experiment 9 as a function of whether the events are in the past or future and whether the change is from easy-to-detect to difficult-to-detect, the opposite, or whether the difficulty of detection is constantly easy or difficult.



two tailed), and the analogous correlation was smaller and not significant for unfairness ($r = -0.12$).⁴ Again, it appears that perceived unfairness does not account for individual differences in the willingness to use probability.

11. GENERAL DISCUSSION

78 The following is a summary of our findings organized by topic:

- Do people consider probability when it varies across cases?

Most experiments: The effect is opposite, but probably spurious because of a positive correlation between severity and probability of detection.

4 Subjects who endorsed deterrence more also endorsed unfairness more ($r = 0.76$).

- Do people consider probability when it is manipulated across matched cases?
Experiment 3: False identity does increase penalties.
Experiment 6: But situational effects do not.
- Do people consider probability when it is transparent (within page)?
All experiments: Affects penalties in general. (Not tested in Experiment 3.)
Experiment 8: Affects penalties but not anger or state of mind judgments.
 Thus, these do not mediate effects on penalties.
Experiment 9: Decreases in detectability increase penalties.
- Do people think that probability should affect penalties?
Experiments 3 and 8: Yes, many people do.
- Do people think that considering probability is unfair?
Experiments 3 and 8: Yes, but less than endorsement of deterrence, and this does not account for individual differences in use of probability.
- Do people think state of mind is relevant?
Experiment 8: Even fewer.
- Can people be induced to think that probability is relevant by asking them about its effect on expected utility?
Experiment 4: Yes.
- Does identification of the perpetrator affect within-page attention to probability of getting caught?
Experiment 6: Individual versus policy has the opposite effect. Probability has more effect in the policy condition.
Experiment 7: Identification increased the effect of probability.
Experiment 8: Failed to replicate Experiment 7; no effect.
- Does emotion reduce the sensitivity to probability?
All experiments: No.

In general, subjects do not seem very sensitive to the probability of detection between cases. Only in Experiment 3 did we find such an effect, and it was induced by the offender's own attempt to hide his identity. Possibly, this attempt made the offense seem more intentional, and it made the offender seem more cowardly, less willing to accept the consequences of the offense.

A substantial minority of subjects in most studies assigned more punishment when probability of detection was lower, within two adjacent cases on a page (within page). In general, even these subjects did not attend to probability across cases, although there was some evidence for this in Experiment 2 (because the *positive* correlation between probability and

punishment was lower in subjects who showed a greater within-page effect, a relationship that was not replicated in other experiments).

81 We found very little support for the possibility that judgments of unfairness lie behind the failure to take probability into account, when it happens. Subjects who think unfairness is relevant are not necessarily those who do not take probability into account (Experiment 8). More likely, this failure is simply a failure to think of the relevance of deterrence.

82 Our results seem to conflict with those of Sunstein et al. (2000), who found that law students tended to think that adjustment of punishment to consider probability of detection was unfair. We note that our cases and sample are different. The difference is more likely due to the cases: as Sunstein et al. point out, their students had been taught the economic theory of optimal deterrence. On the other hand, it could be that they were also taught more deontological theories and the atmosphere tended to favor those over utilitarianism.

83 Although subjects tended not to think of probability as unfair, a few comments explicitly addressed it, for example: “I don’t think the penalty should change based on whether other people are caught or not. Each punishment should be based on the crime committed (and any prior offenses the person has committed).” On the other hand, a few comments also supported a general deterrence principle, for example, “This happens way too much and the punishment should be steep to help it stop.” A few other comments questioned the effectiveness of deterrence, for example, “I don’t think that this type of criminal is influenced by penalties”; “I really wonder if the threat of punitive damages deters many crimes at all.”

84 Probes designed to encourage subjects to think about deterrence did have a small effect, increasing the use of probability in the within-page presentation. At least some of the neglect of probability is the result of simple failure to think of its role, which is an example of the “isolation effect” (McCaffery and Baron, 2006). Consistent with this interpretation is the finding that most subjects—unlike the law students studied by Sunstein et al. (2000)—did not think that different punishments based on probability were inherently unfair.

12. CONCLUSION

85 A recent news story⁵ concerned a 26-year-old man who was facing up to 60 years in prison for infecting over 250,000 computers with “malware” that

5 http://digg.com/security/Was_Botmaster_s_60_year_prison_sentence_too_harsh, accessed Nov. 13, 2007.

enabled him to steal personal information and money. Comments on the story indicated a wide divergence of opinion about the appropriate sentence. Some felt that 60 years was not harsh enough unless it included an element of pain. Others said that it was unfair for the penalty to be greater than that for rape and murder. No comments mentioned the possibility that very few people who commit this sort of crime are ever brought to justice. However, those who wanted a harsh sentence did mention the ubiquity of this sort of crime. They felt that too little was being done about it, and many expressed emotional outrage. It might be that the perception that “crime is rampant” leads to harsher sentences for crimes that are in fact underdeterred. Thus, the system might conform to Bentham’s principles of deterrence, even if very few people explicitly take probability into account in making judgments about penalties. If this is happening, our results also suggest that citizens would not object very strongly to the idea that penalties are negatively correlated with probability of detection, if they knew about it.

More generally, we can ask how our results fit into the study of “behavioral law and economics.” Some writers on law and economics take the view that economic theory is a “positive” theory of law, which is what I would call a “descriptive” theory because it is an empirical theory that attempts to describe the data. Other writers take economic theory to be normative—specifying a standard to which the law should conform—but not necessarily descriptive. The latter view is suggested by findings of cognitive biases in judgments and decisions (Baron, 2008) and by findings of economic ignorance (Caplan, 2007). If citizens’ judgments are inconsistent with economic theory, how could the law come to follow the theory?

Part of the answer is suggested by the first paragraph of this section. It could be that some sort of “invisible hand” brings the law into line. People might not think that deterrence is relevant, but they have other intuitions that coincide with the economic theory of deterrence, so—in their roles as voters, jurors, judges, legislators, and regulators—they end up imposing penalties that coincide, at least roughly, with the economic theory.

Against this possibility is the general fact that many policies, laws, and regulations seem to be clearly irrational from the point of view of economic theory. The most egregious examples can involve risk regulation, where we are faced with very expensive regulations that do little good alongside of relatively inexpensive and beneficial regulations that do not exist (for example, Breyer, 1993; Sunstein, 2002). Other examples involve tort laws that have perverse effects, for example, lawsuits leading to the withdrawal of beneficial vaccines that have harmful side

effects (Baron and Ritov, 1993). Many other examples of apparently irrational policies have been described (Baron, 1998; Baron et al., 2006; Bazerman et al., 2001). Although many of these irrationalities coincide with known cognitive biases, it is also possible that some of them arise for other reasons, for instance, political pressure from interest groups.⁶

89 What about probability and punishment? Although some evidence indicates that both probability and magnitude have deterrent effects (for example, Bar-Ilan and Sacerdote, 2004), we know of no systematic study of the determinants of penalties across various kinds of offenses. Of interest here is whether probability and severity of punishment are negatively correlated, holding other factors constant. In the Appendix, we argue that they are. There also seems to be a positive correlation across crimes. The seriousness of the offense can increase both probability of detection and magnitude of punishment, thus creating a positive correlation that is spurious with respect to the current question.

90 We have ascribed some of our results to moral heuristics. Where do these heuristics come from? One possibility is that they are somehow “built in” by biological evolution. It is difficult to test this possibility, although many psychologists seem to assume it as an axiom. Another possibility, which we are exploring in other works, is that, strangely enough, these heuristics come from the law itself. Just as morality influences law, so does the law influence moral judgment. Because law and morality have different functional roles (Shavell, 2004), they take different forms. For example, the law is necessarily limited in its application. The police cannot go snooping around looking for harmful omissions. They must focus on harmful actions. Thus, moral intuitions often favor harmful omissions over less harmful actions (Baron and Ritov, 2004; Ritov and Baron, 1990). Similarly, the limits of the law lead to a concept of duty (and the corresponding concept of supererogation), which need not be present in moral theory. Utilitarianism in its simplest form, for example, does not distinguish acts and omissions and has no obvious concept of duty (just better and worse).

91 The intuition at issue here is the idea of balancing. It can come from the legal convenience—not always present but often present—of linking

6 For example, one general mechanism is that the effort expended to influence policy is a convex function of the effect of a policy on individual utility. As a result, policies with large effects on a few and small effects on many will be dominated by the interests of the few, even if these are outweighed by the small effects on many. This mechanism requires no cognitive biases, unless we consider the basic convex function, and the dominance of political participation by self-interest, to be cognitive biases in themselves.

compensation and deterrent functions by requiring injurers to compensate victims. This principle, combined with the principle that “making the victim whole” is often the appropriate compensation (especially when damages are pecuniary) leads to the simple idea of making the penalty fit the injury. Note that the Biblical passage in which “an eye for an eye” is most prominently mentioned (Exodus 21: 22–25) concerns the compensation that the injurer must pay for causing a miscarriage. Thus, a legal convenience, a simplification, becomes elevated to the status of a moral rule. We are not sure of this interpretation, of course, but it is one possibility.

What are the implications of our results for policy? Note, first, that the law itself might somehow take probability into account (see Appendix) even if individuals do not. But, taking our results at face value, we would expect that offenses with low probability of detection would be underdeterred and/or those with high probability would be overdeterred. If this is true, then two different kinds of prescriptive solutions suggest themselves. One is to attempt to increase the attention given to probability. For example, judges could instruct jurors to consider the need to increase penalties when the chance of conviction is so low that potential offenders might be inclined to ignore it.

The second implication concerns public policy toward resources devoted to detection and conviction of offenders. As resources increase, the need for long jail sentences (to compensate for reduced probability) is reduced. For serious crimes, then, it makes sense to work around people’s tendency to neglect probability by increasing the resources devoted to apprehension and conviction. For the most serious crimes, it might even be impossible to increase deterrence enough to compensate for a low probability: Even death might not be enough if the probability of detection is low. The Appendix suggests that this has happened: probability of detection is higher for more serious crimes.

For the least serious offenses, however, the law is in a dilemma. If it puts few resources into apprehension and tries to compensate by increasing penalties, the result will seem grossly unfair. Imagine a year in jail for littering. If it puts more resources into apprehension, the cost can exceed the benefit of deterring the offense at all. The inevitable result is that the least serious offenses are underdeterred, if we try to detect and punish them at all. The obvious solution is to give up. The law is understood to have its limits (Shavell, 2004), but our results suggest that it is even more limited than standard economic analysis implies. Minor offenses are best controlled through social norms, not the law. When social norms fail to control them, we just have to put up with them. At least we can understand why.

92

93

94

REFERENCES

- Baayen, R. H., D. J. Davidson, and D. M. Bates, "Mixed-effect modeling with crossed random effects for subjects and items." *Journal of Memory and Language*, 59 (2008): 390–412.
- Bar-Ilan, A., and B. Sacerdote, "The response of criminals and noncriminals to fines." *Journal of Law and Economics*, 47 (2004): 1–17.
- Baron, J. *Judgment misguided: Intuition and error in public decision making*. New York: Oxford University Press, 1998.
- Baron, J. *Thinking and deciding*, 4th edition. New York: Cambridge University Press, 2008.
- Baron, J., M. H. Bazerman, and K. Shonk, "Enlarging the societal pie through wise legislation: A psychological perspective." *Perspectives on Psychological Science*, 1 (2006): 123–132.
- Baron, J., and I. Ritov, "Intuitions about penalties and compensation in the context of tort law." *Journal of Risk and Uncertainty*, 7 (1993): 17–33.
- Baron, J. and I. Ritov, "Omission bias, individual differences, and normality." *Organizational Behavior and Human Decision Processes*, 94 (2004): 74–85.
- Bazerman, M. H., J. Baron, and K. Shonk, *You can't enlarge the pie: The psychology of ineffective government*. New York: Basic Books, 2001.
- Beccaria, C. (1963, original work published 1764), *On crimes and punishments*. Trans. Henry Paolucci. Englewood Cliffs, New Jersey: Prentice Hall, 1963.
- Becker, G. S. "Crime and punishment: An economic approach." *Journal of Political Economy*, 76 (1968): 169–217.
- Bentham, J. *An introduction to the principles of morals and legislation*. Oxford: Blackwell Publisher, 1948. (Original work published 1843)
- Breyer, S. *Breaking the vicious circle: Toward effective risk regulation*. Cambridge, Massachusetts: Harvard University Press, 1993.
- Caplan, B. *The myth of the rational voter: Why democracies choose bad policies*. Princeton, New Jersey: Princeton University Press, 2007.
- Carlsmith, K. M., J. M. Darley, and P. H. Robinson, "Why do we punish? Deterrence and just deserts as motives for punishment." *Journal of Personality and Social Psychology*, 83 (2002): 284–299.
- Ehrlich, I., "Participation in illegitimate activities: A theoretical and empirical investigation." *Journal of Political Economy*, 81 (1973): 521–565.

- Hsee, C. K., and Y. Rottenstreich, "Music, pandas, and muggers: On the affective psychology of value." *Journal of Experimental Psychology: General*, 133 (2004): 23–30.
- Kessler, D. and S. D. Levitt, "Using sentence enhancements to distinguish between deterrence and incapacitation." *Journal of Law and Economics*, 42 (1999): 343–363.
- Kogut, T., and I. Ritov, "The 'identified victim' effect: An identified group, or just a single individual?" *Journal of Behavioral Decision Making*, 18 (2005): 157–167.
- Kogut, T., and I. Ritov, "The singularity effect of identified victims in separate and joint evaluations." *Organizational behavior and human decision processes*, 97 (2005): 106–116.
- McCaffery, E. J., and J. Baron "Isolation effects and the neglect of indirect effects of fiscal policies." *Journal of Behavioral Decision Making*, 19 (2006): 1–14.
- Miller, T. R., M. A. Cohen, and B. Wiersema, *Victim costs and consequences: A new look*. National Institute of Justice Research Report. Washington, D.C.: U.S. Department of Justice, 1996.
- Ritov, I., and J. Baron, "Reluctance to vaccinate: omission bias and ambiguity." *Journal of Behavioral Decision Making*, 3 (1990): 263–277.
- Robinson, P. H. and J. Darley, "The utility of desert." *Northwestern University Law Review*, 91 (1996–7): 453–499.
- Robinson, P. H. and J. Darley, "Testing lay intuitions of justice: How and why?" *Hofstra Law Review*, 28 (2000): 611–634.
- Rottenstreich, Y. and C. K. Hsee, "Money, kisses, and electric shocks: on the affective psychology of risk." *Psychological Science*, 12 (2001): 185–190.
- Shavell, S. M. *Foundations of economic analysis of law*. Cambridge, Massachusetts: Belknap Press of Harvard University Press, 2004.
- Small, D. A. and G. Loewenstein, "Helping the victim or helping a victim: Altruism and Identifiability." *Journal of Risk and Uncertainty*, 26 (2003): 5–16.
- Small, D. A., and G. Loewenstein, "The devil you know: The effects of identifiability on punishment." *Journal of Behavioral Decision Making*, 18 (2005): 311–318.
- Sunstein, C. R. "Moral heuristics (with commentary)." *Behavioral and Brain Sciences*, 28 (2005): 531–573.
- Sunstein, C. R., D. Schkade, and D. Kahneman, "Do people want optimal deterrence?" *Journal of Legal Studies*, 29 (2000): 237–253.

- Sunstein, C. R. *Risk and reason: Safety, law, and the environment*. New York: Cambridge University Press, 2002.
- Tetlock, P. E., P. S. Visser, R. Singh, M. Polifronid, A. Scott, S. B. Elson, P. Mazzocod, and P. Rescober, "People as intuitive prosecutors: The impact of social-control goals on attributions of responsibility." *Journal of Experimental Social Psychology*, 43 (2006): 195–209.
- Vandaele, W. "Participation in illegitimate activities: Ehrlich revisited, 1960." [Computer file]. ICPSR08677-v1. Ann Arbor, Massachusetts: Inter-university Consortium for Political and Social Research, 1987.
- Wilkinson-Ryan, T., and J. Baron, (in press). "Moral judgment and moral heuristics in breach of contract." *Journal of Empirical Legal Studies*.

APPENDIX: THE CORRELATION OF PROBABILITIES AND SENTENCES

Here we address two issues. First, is probability of conviction of different offenses correlated positively with the penalties for them? We have suggested that the intuitive positive correlation that we sometimes find is based on reality. What is the reality? 95

Second, holding constant the offense, is the sentence higher when probability of detection is lower? That is, does the real world follow Bentham's advice? 96

On the first point, the correlation between probability and penalties across offenses, Table 1 of Robinson and Darley (1996–7), presents data for seven crimes (roughly: murder, rape, robbery, assault, burglary, larceny, and auto theft—a classification used by the U.S. Federal Bureau of Investigation). The log of the duration of the average state sentence (available for all seven crimes) is correlated positively with the log probability of conviction (specifically, the ratio of number of prison sentences to the number of crimes committed) across the seven crimes: $r = 0.96$ ($p = 0.0004$, 5 df). Although this sample of crimes is small, the correlation makes it plausible that subjects believe that, in general, serious offenses are more likely to be caught and punished. Such a belief could be based on knowledge of the real world. 97

As pointed out by Louis Kaplow, the same tiny data set can be used to indirectly answer the second question. Although we cannot hold constant the offense, we can roughly control for the harmfulness of the offense. We can look at the ratio of the harshness of the typical sentence (in months) to the (monetized) harmfulness of the crime. Using data from Miller, Cohen, and Wiersma (1996), Prof. Kaplow computed this ratio for us. The log of this ratio was correlated negatively with the log probability of conviction, across the seven crimes: $r = -0.86$ ($p = 0.0130$, 5 df). In other words, controlling for harmfulness, a lower probability of conviction is associated with a harsher sentence. This result suggests that Bentham's principle affects actual sentences, but, of course, this result is very tentative because of the nature of the measures. 98

A larger data set that can be used to answer the same questions is that used first by Ehrlich (1973) and made available by Vandaele (1987). Ehrlich collected data on reports of the same seven major crimes, conviction probabilities, and average sentences (plus many other variables) for each of 47 states of the United States, for at least the year 1960, the year we use (because it is complete in the data set). Thus, we can examine variations across 99

states as well as across crimes. For our analysis, we used the log of the mean sentences and the log of probabilities of conviction.⁷

100 In the most useful regression model, we used a mixed-effects model (Baayen, Davidson, and Bates, 2008), in which state and crime were crossed random effects. The dependent variable was the log sentence (for each of the $47 \times 7 = 329$ data). The fixed-effect predictors were the mean log probability for the state, the mean log probability for the crime, and the log probability of the given crime for the given state corrected for both the crime and the state (that is, the deviation from the log probability expected on the basis of the crime and the state). (Other analyses produced results consistent with this one but seemed less enlightening.) Because all the variables are logged, the coefficients correspond to unit-independent elasticities, the effect of a given proportional change in one variable on the proportional change in another.

101 The estimated coefficients were 0.540 for the effect of crime probabilities, -0.215 for the effect of state probabilities, and -0.033 for the effect of the corrected state/crime probability. Reliability was assessed using MCMC sampling (Baayen et al., 2008), which yielded p-levels of 0.0001, 0.0002, and 0.0340, for the three coefficients, respectively. In sum, crimes with harsher sentences are (again) associated with higher probabilities of conviction, but, holding the crime constant, states that have higher probabilities of conviction tend to have less harsh sentences; this negative relationship between probability and sentence also applies when we hold both the state and crime constant.

102 The best explanation of all these results—which are completely consistent—is that more serious crimes lead to more enforcement effort, which, in turn, leads to higher conviction rates. However, the justice system otherwise does use harsher sentences to compensate for lower conviction probabilities and thus to maintain a situation closer to optimal deterrence.

7 The probabilities were actually ratios based on fallible statistics. The log produced more symmetric distributions than other transforms such as log odds. For both variables, the logs were distributed roughly normally except for somewhat fat tails on both sides.