

Saliency Theory of Judicial Decisions

Pedro Bordalo, Nicola Gennaioli, Andrei Shleifer*

November 2013

Abstract

We present a model of judicial decision making in which the judge overweights the salient facts of the case. The context of the judicial decision, which is comparative by nature, shapes which aspects of the case stand out and draw the judge's attention. By focusing judicial attention on such salient aspects of the case, legally irrelevant information can affect judicial decisions. Our model accounts for a range of recent experimental evidence bearing on the psychology of judicial decisions, including anchoring effects in the setting of damages, decoy effects in choice of legal remedies, and framing effects in the decision to litigate. The model also offers a new approach to positive analysis of damage awards in torts.

*Royal Holloway, University of London, pedro.bordalo@rhul.ac.uk; Università Bocconi, IGIER, nicola.gennaioli@unibocconi.it; Harvard University, shleifer@fas.harvard.edu. We are grateful to Richard Posner for extremely helpful comments.

1 Introduction

According to legal formalism, judges deciding cases impartially verify facts and apply the pertinent law to these facts, with little room left for discretion. After all, the theory goes, judges are selected and trained to perform such formal analysis, they need to explain their opinions in writing, which disciplines thought, they are constrained by the adversarial system from missing relevant facts or rules, and they are kept on their toes by the risk of appellate review. In stark contrast to this perspective, legal realism holds that both trial and appellate judges enjoy significant discretion in deciding cases. Exercise of such discretion turns not just on the merits of the case but also on personal factors such as judges' political views, their beliefs about the litigants, or the lawyers' persuasive tactics.

Much recent research on legal realism has focused on political influences on judicial behavior, especially in the discussions of appellate courts and judicial law-making (Posner 2008, Epstein, Landes, and Posner 2013). Yet the original work on legal realism emphasized judicial psychology, particularly as it influences the decisions of trial judges (Radin 1925, Frank 1973). The role of psychology in judicial decisions is again receiving some attention, especially in experimental research (Guthrie, Rachlinski and Wistrich 2001, Kelman, Rottenstreich and Tversky 1996, Viscusi 1999, Rachlinski, Wistrich and Guthrie 2013). This research is based on the classic work of Kahneman and Tversky (Tversky and Kahneman 1974, Kahneman and Tversky 1979), who identify a range of heuristics that decision makers use to make judgments, and document biases that result from the use of such intuitive heuristics. We recently proposed Saliency Theory to unify the existing evidence on heuristic decision making (Bordalo, Gennaioli and Shleifer 2012, 2013, henceforth BGS). In this paper, we revisit experimental evidence on judicial biases from the perspective of this theory.

Saliency Theory builds on the premise that the valuation of a choice option occurs not in isolation, but in a comparative context. Decision makers contrast the features of the option in question to the features of choice alternatives or of "normal" situations that come to the decision maker's mind. For instance, the valuation of a premium good may fall if the good's high price (rather than its high quality) is salient, as when the good is presented together with cheaper alternatives or when the decision maker is accustomed to buying the same good

at lower prices. The cognitive underpinning of Saliience Theory is that human decisions are mediated by perception and attention. In evaluating a range of options, attention is drawn to unusual, extreme or salient attributes of each option. Because they draw attention, salient attributes receive more weight in the decisions than is justified by normative theory. Saliience Theory accounts for well documented behavioral patterns such as the instability of risk preferences (including preference reversals and Allais paradoxes), and more generally provides a novel and psychologically based account of context effects (such as decoy effects). By shaping what is perceived as normal and what is perceived as unusual or salient, context influences attention and decisions.

This logic seems useful to examine judicial behavior. After all, most judicial decision making is comparative. Judges compare the facts of the present case with facts of similar cases, legal precedents, statutes, personal experiences or expectations in similar situations, and so on. The fact that judicial decisions are made in context is partly by design; precedents are indeed supposed to anchor judicial decisions. At the same time, Saliience Theory suggests that these anchors may focus judicial attention on peculiar aspects of the case which, though salient, may not be the most useful or relevant for the legal decision.

In this paper, we first summarize Saliience Theory, and then apply it to several experiments bearing on the psychology of judicial decisions. In particular, we consider i) anchoring mechanisms in the setting of damages, ii) decoy effects affecting the choice among different legal remedies, and iii) framing of payoffs in shaping the choice of settlement versus litigation. The key papers our analysis draws on include Rachlinksi (1996), Guthrie, Rachlinski and Wistrich (2001), Kelman, Rottenstreich and Tversky (1996), and Viscusi (1999, 2001). We show that Saliience Theory provides a unified account of arguably disparate bits of evidence on judicial decision making, but also offers a new approach to positive analysis of damage awards in torts. In the concluding section, we consider some implications of Saliience Theory for such topics as legal evolution, legal procedure, and the stability of law.

2 Salience Theory

2.1 A Simple Example

We first illustrate the logic of salience in the context of attitudes towards risk. Consider the following experiments involving a choice between two risky prospects, or lotteries:

Experiment 1: Choose between the following two options:

$$L_1 = \left\{ \begin{array}{ll} \$1 & \text{with prob. } 0.95 \\ \$381 & 0.05 \end{array} \right. , \quad L_2 = \left\{ \begin{array}{l} \$20 \text{ for sure.} \end{array} \right. \quad (1)$$

Experiment 2: Choose between the following two options:

$$L_1 = \left\{ \begin{array}{ll} \$301 & \text{with prob. } 0.95 \\ \$681 & 0.05 \end{array} \right. , \quad L_2 = \left\{ \begin{array}{l} \$320 \text{ for sure.} \end{array} \right. \quad (2)$$

In both experiments, the risky option L_1 offers a small (5%) probability of a high payoff, and a high (95%) probability of a \$19 loss relative to the sure outcome given by L_2 , but the two lotteries have the same expected payoff. Experiment 2 simply adds \$300 to all the payoffs in Experiment 1.

A sample of 120 subjects participated in the two experiments (see BGS 2012b). In Experiment 1, 83% of the subjects chose the safe option L_2 , while in Experiment 2, 67% of the same subjects chose the risky lottery L_1 . Given that in each experiment the two options offer the same expected value, the same subjects exhibit risk aversion in the first experiment and are risk seeking in the second.

According to Expected Utility Theory, the stakes in either experiment are so low compared to lifetime wealth that people ought to be roughly risk neutral, choosing the risky lottery approximately 50% of the time. A small level of risk aversion would create stable preferences for the safe lottery L_2 in both experiments. Salience Theory, in contrast, predicts the observed pattern of risk aversion in Experiment 1 and risk seeking behavior in Experi-

ment 2.¹ In Experiment 1, the risky lottery L_1 's downside of \$1 feels very low relative to the sure payoff of \$20. This bad outcome is salient and subjects overweight it in their decision, leading to risk averse behavior. In Experiment 2, the risky lottery L_1 's downside of \$301 looks fairly similar to the sure payoff of \$320. The prospect of winning \$681 is more salient, leading to risk seeking behavior.

As we show in BGS (2013), the same logic of comparative evaluation can be applied to riskless choice. To give an example, imagine yourself in a wine store, choosing a red wine. You consider a French syrah from the Rhone Valley, selling for \$20 a bottle, and an Australian shiraz selling for \$10. You like French syrah better, you think it is perhaps 50% better. Yet it sells for twice as much. You decide the Australian shiraz is a better bargain and buy a bottle.

A few weeks later, you are at a restaurant, and you see the same two wines on the wine list. Both of them are now marked up by \$40, with the French syrah selling for \$60 a bottle, and the Australian shiraz for \$50. You again think the French wine is 50% better, but now it is only 20% percent more expensive. At the restaurant, it is a better deal. You splurge and order the French wine. Saliency Theory again predicts this reversal. At the store, the price difference between the cheaper and the more expensive wine is more salient than the quality difference, encouraging the consumer to opt for the cheaper option, whereas at the restaurant, after the markups, the quality difference is more salient, encouraging the consumer to splurge.

These examples illustrate how decision makers think in context when figuring out which of several choices represents a better deal in light of the options available. We next describe the saliency model, formalising the intuition behind such thinking.

¹Prospect Theory (Kahneman and Tversky, 1979) holds that people overweight the small 5% probability of the high outcome, favoring risk loving behavior in both experiments. To obtain the reversal of risk attitudes in Experiments 1 and 2, Prospect Theory requires a combination of probability weighting and decreasing absolute risk aversion in the value function, provided that the decision maker's reference point is the status quo, zero gains or losses. If instead the reference point is the sure gain, the two choice problems are identical and Prospect Theory cannot account for the evidence. A systematic comparison of Prospect and Saliency Theory is presented in BGS (2012).

2.2 The Model

A decision maker evaluates $N > 1$ alternatives in a choice context $\mathbf{C} = \{(a_{1k}, a_{2k})\}_{k=1, \dots, N}$, whose generic element is an option k identified by the value of its two attributes a_{1k} and a_{2k} . Attributes are measured in dollars and known to the decision maker. The choice context \mathbf{C} contains the alternatives of choice but can in addition contain “normal” choice options that are usually considered by the decision maker but are not currently available, or other options that are explicitly brought to the decision maker’s attention.

This framework is sufficiently general to accommodate several applications. In choice under risk, the choice context \mathbf{C} consists of different risky prospects with up to two different payoffs, in which case the attributes describe prospects’ payoffs (BSG 2012, and the framing example of Section 3.3). In riskless choice, the choice context consists of different consumer goods, in which case the attributes would stand for quality and price (BSG 2013). In a litigation example, the choice context consists of the loss and damages from different decisions available to the judge (see Section 3.1).

A fully rational decision maker values option k according to the linear utility function:

$$u(a_{1k}, a_{2k}) = \theta_1 a_{1k} + \theta_2 a_{2k}, \quad (3)$$

where $\theta_1, \theta_2 > 0$, and where weights add up to one, $\theta_1 + \theta_2 = 1$. If the decision maker chooses among risky prospects, θ_1 captures the probability of state 1 and $\theta_2 = 1 - \theta_1$ the probability of state 2. If the decision maker chooses among consumer goods characterized by quality and price, then θ_1 is the weight on quality q_k while θ_2 is the weight on price p_k (so that $a_{1k} = q_k$ and $a_{2k} = -p_k$). We assume for simplicity that in this case $\theta_1 = \theta_2 = 1/2$.

A salient thinker departs from Equation (1) by inflating the relative weight attached to the attribute that he perceives to be more salient. An attribute is salient for option k in the choice set \mathbf{C} if this attribute “stands out” relative to the other attributes of the same option. Formally, denote by (\bar{a}_1, \bar{a}_2) the reference option consisting of average attributes $\bar{a}_i = \sum_k^N a_{ik}/N$ in \mathbf{C} . The salience of attribute $i = 1, 2$ for a generic option k is then given by $\sigma(a_{ik}, \bar{a}_i)$, where $\sigma(\cdot, \cdot)$ is a salience function that satisfies the following properties:

Definition 1 *The salience function $\sigma(\cdot, \cdot)$ is symmetric, continuous and satisfies:*

1) *Ordering.* Let $\mu = \text{sgn}(a_{ik} - \bar{a}_i)$. Then for any $\epsilon, \epsilon' \geq 0$ with $\epsilon + \epsilon' > 0$ we have

$$\sigma(a_{ik} + \mu\epsilon, \bar{a}_i - \mu\epsilon') > \sigma(a_{ik}, \bar{a}_i). \quad (4)$$

2) *Diminishing sensitivity.* For any $a_{ik}, \bar{a}_i \geq 0$ and all $\epsilon > 0$ we have:

$$\sigma(a_{ik} + \epsilon, \bar{a}_i + \epsilon) < \sigma(a_{ik}, \bar{a}_i). \quad (5)$$

3) *Reflection.* For any $a_{ik}, \bar{a}_i, a'_{ik}, \bar{a}'_i \geq 0$ we have:

$$\sigma(a_{ik}, \bar{a}_i) > \sigma(a'_{ik}, \bar{a}'_i) \Leftrightarrow \sigma(-a_{ik}, -\bar{a}_i) > \sigma(-a'_{ik}, -\bar{a}'_i). \quad (6)$$

For a given option k , its attribute 1 is salient when $\sigma(a_{1k}, \bar{a}_1) > \sigma(a_{2k}, \bar{a}_2)$, while its attribute 2 is salient when $\sigma(a_{1k}, \bar{a}_1) < \sigma(a_{2k}, \bar{a}_2)$. Attributes 1 and 2 are equally salient when $\sigma(a_{1k}, \bar{a}_1) = \sigma(a_{2k}, \bar{a}_2)$.

Definition 1 encodes three key features of sensory perception. First, our perceptive apparatus is attuned to detect changes in stimuli. This is captured by ordering, which states that salience increases in contrast: option k 's value of attribute i , namely a_{ik} , is more salient the farther it is from the reference value \bar{a}_i of that attribute. In the consumer choice example, if a consumer good is much cheaper than average, then its low price is very salient. In the risky choice example, if in a given state of the world a risky lottery pays much more than all the other lotteries do, then that high payoff is salient for the lottery in question.

The second feature is that changes in stimuli are perceived with diminishing sensitivity (Weber's law). Formally, an attribute's salience decreases as the value of that attribute increases uniformly for all options. For instance, at higher price (or payoff) levels, given price (or payoff) differences are less noticeable and thus less salient. The difference between \$10 and \$20 looms larger than the difference between \$110 and \$120.

Third, our perceptive apparatus is sensitive to differences in magnitudes, or absolute values, so the impact of salience on positive and negative attributes is formally identical. In the case of consumer goods, quality and price are treated symmetrically. In the case of

lotteries, we attend to unusually large positive and negative payoffs equally.

These features interact to determine the salience of each option's attributes, but they can also be in tension. For instance, in the consumer goods case, if price levels and price dispersion both increase, then diminishing sensitivity suggests that price salience decreases while ordering suggests the opposite. Following BSG (2013) we resolve this tension by assuming that the salience function is homogeneous of degree zero:

A.0: The salience function satisfies ordering and homogeneity of degree zero, which is defined as $\sigma(\alpha \cdot a_i, \alpha \cdot \bar{a}_i) = \sigma(a_i, \bar{a}_i)$ for all $\alpha \neq 0$ and all $a_i, \bar{a}_i \neq 0$.

Homogeneity of degree zero characterizes salience in an intuitive way: an attribute's salience increases with the *proportional difference* in the value of that attribute relative to the reference level.² Together with ordering, homogeneity implies diminishing sensitivity and reflection of salience.

An example of a salience function satisfying homogeneity of degree zero is:

$$\sigma(a_{ik}, \bar{a}_i) = \frac{|a_{ik} - \bar{a}_i|}{|a_{ik}| + |\bar{a}_i|}, \quad (7)$$

for $a_{ik}, \bar{a}_i \neq 0$, and $\sigma(0, 0) = 0$.

Consider how salience distorts the valuation of an option in the choice set. Given a salience function σ , the decision maker – we call him a salient thinker – ranks the salience of the option's attributes and distorts their utility weights as follows:

Definition 2 *The salient thinker's valuation of option k enhances the relative weight attached to the salient attribute (while keeping the sum of utility weights constant). Formally:*

$$u_k^s = \begin{cases} \frac{\theta_1}{\theta_1 + \delta\theta_2} \cdot a_{1k} + \frac{\delta\theta_2}{\theta_1 + \delta\theta_2} \cdot a_{2k} & \text{if } \sigma(a_{1k}, \bar{a}_1) > \sigma(a_{2k}, \bar{a}_2) \\ \frac{\delta\theta_1}{\delta\theta_1 + \theta_2} \cdot a_{1k} + \frac{\theta_2}{\delta\theta_1 + \theta_2} \cdot a_{2k} & \text{if } \sigma(a_{1k}, \bar{a}_1) < \sigma(a_{2k}, \bar{a}_2) \\ \theta_1 \cdot a_{1k} + \theta_2 \cdot a_{2k} & \text{if } \sigma(a_{1k}, \bar{a}_1) = \sigma(a_{2k}, \bar{a}_2) \end{cases} \quad (8)$$

where $\delta \in (0, 1]$ decreases in the severity of salient thinking.

²In other words, upon a variation in an option's attribute a_{ik} , ordering dominates diminishing sensitivity if and only if the change in a_{ik} is proportionally larger than the induced change in the reference \bar{a}_k .

When evaluating an option k , the salient thinker pays more attention to, and overweights, its salient attribute. If attribute 1 is salient, its relative utility weight increases from θ_1 to $\theta_1/(\theta_1 + \delta\theta_2) > \theta_1$, while the relative weight of attribute 2 decreases from θ_2 to $\delta\theta_2/(\theta_1 + \delta\theta_2) < \theta_2$. Conversely, if attribute 2 is salient, its relative utility weight increases to $\theta_2/(\theta_2 + \delta\theta_1) > \theta_2$, while the relative weight of attribute 1 decreases to $\delta\theta_1/(\theta_2 + \delta\theta_1) < \theta_1$. As $\delta \rightarrow 0$, the salient thinker considers only the most salient attribute and fully neglects the non-salient attribute. If $\delta = 1$, the salient thinker pays equal attention to both salient and non-salient attributes, so his valuation is rational. Normalization of the utility weights ensures that valuation of the option (a_{1k}, a_{2k}) lies between a_{1k} and a_{2k} .

The salience model provides a novel and unified framework to explain a large set of anomalies in individual choice, both in a riskless setting (BGS, 2013) and in choice under risk (BGS 2012). We now consider two examples of the impact of salience on choice, which will bear on our subsequent analysis of judicial decision making.

2.3 Payoff Salience and the Instability of Risk Attitudes

The salience model provides a unified explanation for the instability of risk preferences illustrated by Allais paradoxes, framing effects, and preference reversals (BGS 2012). In fact, equation (8) shows that the decision weights associated to lottery payoffs are determined by the payoffs' salience, which in turn is shaped by the comparison to other lotteries in the choice set and by how payoffs are framed.

To illustrate the model, we now work through Experiments 1 and 2 of section 2.1. Each Experiment compares a risky lottery to a sure prospect. The risky lottery's downside occur in state 1, with probability $\theta_1 = .95$, and its upside on state 2, with probability $\theta_2 = .05$. In Experiment 1, the two options are $L_1 = (a_{11}, a_{21}) = (1, 381)$ and $L_2 = (a_{12}, a_{22}) = (20, 20)$. The reference payoffs are then given by $(\bar{a}_1, \bar{a}_2) = (\frac{20+1}{2}, \frac{20+381}{2}) = (10.5, 200.5)$. The salience of L_1 's downside is $\sigma(a_{11}, \bar{a}_1) = \sigma(10.5, 1)$, while the salience of its upside is $\sigma(a_{21}, \bar{a}_2) = \sigma(381, 200.5)$. L_1 's most salient payoff is its downside, because it exhibits a larger percentage difference from the relevant reference payoff, namely $10.5/1 > 381/200.5$.³

³In greater detail, $\sigma(a_{11}, \bar{a}_1) = \sigma(1, 10.5)$ which equals $\sigma(10.5, 1)$ by symmetry of the salience function. In turn, $\sigma(381, 200.5) = \sigma(381/200.5, 1)$ due to homogeneity. Then ordering ensures that $\sigma(10.5, 1) >$

Given this salience ranking, Equation (8) implies that the lotteries are evaluated as:

$$u^S(L_1) = \frac{0.95}{0.95 + \delta \cdot .05} \cdot 1 + \frac{\delta \cdot .05}{0.95 + \delta \cdot .05} \cdot 381, \quad u^S(L_2) = 20. \quad (9)$$

In evaluating lottery L_1 , the salient thinker increases the weight attached to its salient downside of \$1 and reduces the weight attached to its upside of \$381. As a consequence, his evaluation of L_1 falls below its expected value of 20. The valuation of the sure prospect L_2 is instead unaffected by salience, because it yields the same payoff in all states. As a consequence, in Experiment 1 the salient thinker is risk averse, choosing L_2 over L_1 .

Consider the behavior of the salient thinker in Experiment 2. The two options are $L_1 = (a_{11}, a_{21}) = (301, 681)$, and $L_2 = (a_{12}, a_{22}) = (320, 320)$, and the reference payoffs are $(\bar{a}_1, \bar{a}_2) = (310.5, 500.5)$. The salient payoff for L_1 is its upside, because state 2 features a higher percentage payoff difference than state 1, namely $681/500.5 > 310.5/301$. The two options are now valued as:

$$u^S(L_1) = \frac{\delta \cdot .95}{\delta \cdot .95 + 0.05} \cdot 301 + \frac{0.05}{\delta \cdot .95 + 0.05} \cdot 681, \quad u^S(L_2) = 320.$$

In evaluating lottery L_1 the salient thinker now increases the weight attached to its salient upside \$681 and reduces that of its downside \$301, so that $u^S(L_1) > 320$. Just as the experimental subjects, the salient thinker is now risk seeking and prefers L_1 to L_2 .

From this analysis we distill the following psychological intuition for the reversal in risk attitudes. The choice between the two lotteries requires individuals to trade off – for given probabilities – the \$19 loss against the \$361 gain of the risky lottery. In standard theory, individuals evaluate this tradeoff in absolute terms, making context independent choices. Salience theory, in contrast, captures the psychological principle that individuals evaluate costs and benefits not in a vacuum, but relative to the choice context they face. The \$19 loss of the risky lottery looks much smaller (and thus non-salient) when the payoff level is high at \$320 than when the payoff level is low at \$20. In contrast, the \$361 gain of the risky lottery looks quite large regardless of the payoff level. The shift in payoff levels changes the “anchor” to which lottery gains and losses are compared, explaining the shift from risk

$\sigma(381/200.5, 1)$, because $10.5 > 381/200.5$.

aversion in Experiment 1 to risk seeking in Experiment 2.

2.4 Wine and Decoys

Consider again the riskless choice between a French syrah and an Australian shiraz. At the store, the French syrah has attributes $(30, -20)$ and the Australian shiraz has attributes $(20, -10)$, where attribute 1 captures quality while attribute 2 captures price (both measured in dollars). In the store the reference wine is $(\bar{a}_1, \bar{a}_2) = (25, -15)$. In this case, price is salient for both the French syrah (i.e., $30/25 < 20/15$) and the Australian shiraz (i.e., $25/20 < 15/10$). The salient thinker focuses on price and chooses the cheaper Australian wine. Formally, $\delta \cdot 30 - 20 < \delta \cdot 20 - 10$.

At the restaurant, the French syrah has attributes $(30, -60)$, while the Australian shiraz has attributes $(20, -50)$. The reference wine at the restaurant is then $(\bar{a}_1, \bar{a}_2) = (25, -55)$. Quality is now salient for both the French syrah (i.e., $30/25 > 60/55$) and the Australian shiraz (i.e., $25/20 > 55/50$). The salient thinker focuses on quality and chooses the more expensive French wine. Formally, $30 - \delta \cdot 60 > 20 - \delta \cdot 50$.

The role of comparative evaluation is clear: the extra \$10 cost of the French syrah looks high when perceived in the context of low store prices, but it feels low when perceived in the context of high restaurant prices. As a consequence, the consumer is willing to pay \$10 more to get higher quality at the restaurant, but not at the store.

Comparative evaluation can also induce the so called decoy effect, which has been well documented in psychology and marketing (Huber, Payne and Puto 1983, Tversky and Simonson 1993): given a pairwise choice, adding an alternative dominated by one of the available options boosts the demand for the dominating option. This is at odds with standard theory, according to which the inclusion of irrelevant (in this case dominated) alternatives should not affect choice.

To see how this works, suppose that at the store a third, expensive wine $d \equiv (30, -30)$ is added to the choice set. This wine has the same quality of the French syrah but a higher price. The inclusion of this option would not affect the preferences of a rational decision maker. Not so for a salient thinker. In the new choice set, the reference wine is $(\bar{a}_1, \bar{a}_2) = \left(\frac{30+30+20}{3}, -\frac{30+20+10}{3}\right) = \left(\frac{80}{3}, -20\right)$. Against this new reference good, the French

syrah is now quality salient (i.e., $30/(80/3) > 20/20$), while the Australian shiraz continues to be price salient (i.e., $20/(80/3) < 20/10$). As a consequence, when the expensive decoy wine d is added, the salient thinker switches to choosing the expensive French wine. Formally, $30 - \delta \cdot 20 > \delta \cdot 20 - 10$.

Once again, this instability of preferences is due to comparative evaluation. Without the decoy, the French wine is evaluated in comparison to the cheaper Australian wine. In this context, the consumer focuses on the fact that the French wine is more expensive, so that it is perceived as a worse deal than the alternative. When the expensive decoy wine d is added to the choice set, the consumer compares the French wine to the decoy as well. In this context, the French wine is perceived as a better deal, so that the consumer focuses on its high quality and chooses it.

3 Salience Theory and Judicial Decision Making

Judicial decision making is a promising field for applying Salience Theory. Most judicial decisions involve the resolution of complex tradeoffs. Judicial evaluation of specific situations does not occur in a vacuum, and may be shaped by the litigants' positions and the judge's own experience, over and above the informational content of these inputs. A recent experimental literature has documented a variety of such contextual effects on legal decisions by jury-eligible persons as well as by judges (e.g. Rachlinski 1996, Guthrie, Rachlinski and Wistrich 2001, Viscusi 1999, 2001, Kelman, Rottenstreich and Tversky 1996). In this section, we show that the salience model provides a unified account of these disparate pieces of experimental evidence. This suggests that the logic of comparative evaluation captured by Salience Theory may play an important role on judicial evaluation more generally.

A skeptic might object that experimental evidence sheds little light on actual courtroom decision making. Experiments are necessarily simple, and fail to capture the complexity of actual cases. Judges in real world cases care more about getting the right answer than they do when answering experimental questions. They have the time and resources to get the information they need. Perhaps most important, judges are trained to cut through the forest of information, and to focus on the legally relevant facts that they can apply the law to.

We do not find these objections compelling. If anything, the complexity of actual courtroom experience would make judges more rather than less vulnerable to psychological biases than they are in streamlined experiments. In the courtroom, judges are bombarded with material that draws their attention away from legally relevant facts, including human aspects of the case, attorneys’ rhetoric, and introduction of precedents pulling in different directions. Judges are extremely busy, and must devote enormous effort to keep straight all the facts and legal nuance under consideration. Rather than cut through the forest of irrelevant disagreements, judges may look for a clearing in the forest that lets the light shine through. Yet such a clearing is likely to be delivered precisely by the salient facts of the case, which stand out and draw judicial attention, even when the salience of these facts is driven by legally irrelevant information. The salient facts or precedents enable judges to form quick intuitive assessments as to “who is right,” and to proceed to process further evidence through this lens. In balancing conflicting evidence, salience of particular facts can thus distort the judicial decision making process, and more so when the situation is more complex. In our view, then, the evidence from the relatively clean and straightforward experiments puts a lower bound on the influence of psychological factors on judicial decisions.

3.1 Irrelevant Information and the Setting of Damages

Judicial setting of monetary damages often involves a subjective comparison of the loss incurred by the plaintiff and the cost imposed on the defendant. By the comparative evaluation principle, these subjective quantifications can be influenced by bringing to a judge’s mind specific instances of compensation. The following experiments provide two concrete examples of this phenomenon.

Experiment 1. Guthrie, Rachlinski and Wistrich (2001) describe an experiment conducted with a subject pool of 167 federal magistrate judges, who were presented with the following case: “Suppose that you are presiding over a personal injury lawsuit that is in federal court based on diversity jurisdiction. The defendant is a major company in the package delivery business. The plaintiff was badly injured after being struck by one of the defendant’s trucks when its brakes failed at a traffic light. Subsequent investigations revealed

that the braking system on the truck was faulty, and that the truck had not been properly maintained by the defendant. The plaintiff was hospitalized for several months, and has been in a wheelchair ever since, unable to use his legs. He had been earning a good living as a free lance electrician and had built up a steady base of loyal customers. The plaintiff has requested damages for lost wages, hospitalization and pain and suffering, but has not specified an amount. Both parties have waived their rights to a jury trial". Judges were randomly assigned to either a "No Anchor" or an "Anchor" condition. Judges in the No Anchor group were asked: "how much would you award the plaintiff in compensatory damages?" Judges in the Anchor condition were also informed that "[t]he defendant has moved for dismissal of the case, arguing that it does not meet the jurisdictional minimum for a diversity case of \$75,000". These judges had to rule on the motion, and then they were asked "[i]f you deny the motion, how much would you award the plaintiff in compensatory damages?" The additional information here constitutes a (normatively irrelevant) anchor on the grounds that the plaintiff had clearly incurred damages greater than \$75,000. As a consequence, absent anchoring effects, judges should deny the motion to dismiss and there should be no difference in the damages awarded by judges belonging to the No Anchor and to the Anchor treatment.

Consistent with the motion being meritless, only two (2.3%) of the judges in the Anchor group granted the motion to dismiss the case. At the same time, the 66 judges in the No Anchor condition indicated that they would award plaintiff an average of \$1.25 million while the 50 judges in the Anchor condition awarded an average of \$882,000. Asking the judges to rule on a frivolous motion depressed mean damage awards by more than \$350,000 (or 29.4%) in this hypothetical case, pointing to a very strong anchoring effect.

Experiment 2. Viscusi (2001) describes an experiment conducted with a subject pool of jury-eligible subjects, who were presented with the following case: "A major auto company with annual profits of \$7 billion made a line of cars with a defective electrical system. This failure led to a series of fires in these vehicles that caused 4 burn deaths per year. Changing the design to prevent these deaths would cost \$16 million for the 40,000 vehicles affected per year. This safety design change would raise the price of cars \$400 each. The company

thought there might be some risk from the current design, but did not believe it would be significant. The company notes that even with these injuries, the vehicle had one of the best safety records in its class. The courts have awarded each of the victims' families \$800,000 in damages to compensate them for the income loss and pain and suffering that resulted. After these lawsuits, the company altered future designs to eliminate the problem."

Subjects were randomly assigned to two groups, which we call "Precautions Considered" and "Precautions Not Considered." In both groups, subjects were asked to indicate whether the court should award punitive damages and, if so, how much. However, in the Precautions Considered group, subjects were further told that the auto company had performed a corporate risk analysis and decided against taking precautions and changing the design. This additional information is again normatively irrelevant on the grounds that precautions are inefficient from an economic perspective: the \$16 million cost to the company to fix the faulty system greatly exceeds the expected loss (expected number of fires times compensation per fire), valued at roughly \$3.2 million.

In contrast to the efficient precautions analysis, a large majority of subjects in either group would award punitive damages in this case. Centrally to our argument, however, subjects in the Precautions Considered group were more likely to award punitive damages (96% vs 88%), and, conditional on this, awarded much higher damages: the median award more than tripled, from \$1 million in the No Precautions Considered group to \$3.5 million in the Precautions Considered group.

The authors of the two studies interpret their respective results in light of distinct and very specific mechanisms of decision making. To explain the effect of bringing the motion to dismiss to the judges' attention, Guthrie et al. (2001) suggest that being exposed to the motion may have caused the judges to consider the possibility that the true damages in this case were exceptionally low, so that their resolution of the uncertainty associated with the loss results in comparatively low damage estimates. To explain the effect of reminding judges about the failed possibility of taking precautions, Viscusi (2001) attributes the exceptionally high punitive damages to the subjects' "outrage" at the company's actions, and to their wish to "send a signal" to the company that "human life is more important than profits."⁴

⁴To get at what drives his results, Viscusi (2001) varies the level of risk and the cost per life saved in his

In contrast to these case-specific accounts, we argue that both examples arise from the same mechanism of comparative evaluation. Before the formal analysis, we briefly describe its logic. After reading the damage case, the judge (or juror) comes up with an assessment of the plaintiff’s loss. If this loss is salient in the context of the case, the judge sets high damages (potentially above the assessment) to reflect its perceived severity. If in contrast the loss is not salient, the judge sets lower damages (potentially below the assessment) to avoid an unfairly high penalty on the defendant. Consider the Guthrie et al. (2001) experiment: when the motion to dismiss suggests that the plaintiff’s severe injuries could be compensated with \$75,000 or less, any reasonable damage figure (which is necessarily much higher than \$75,000) is salient, and more likely to stand out relative to the assessed loss. As a consequence, judges select lower damages in the Anchor group compared to the No Anchor group. In the Viscusi (2001) experiment, when the jurors are reminded that no loss would have occurred had precautions been taken, a course of action the company contemplated and ultimately rejected, any estimated loss becomes more salient relative to the low cost to the company of avoiding it (regardless of efficiency considerations). As a consequence, jurors award higher damages in the Precautions Considered group than in the No Precautions Considered group.⁵

Saliency Theory offers a formal account of the judge’s decision making process. In setting damages, he evaluates a trade-off between two dimensions: an assessment of the plaintiff’s loss L (dimension 1) and the cost C to the defendant (dimension 2) of precautions or damage payments in court. A loss-cost package is then captured by the vector $(-L, C)$.

A judge wants to set compensation so as to make the plaintiff whole, namely so that the plaintiff is indifferent between having his loss compensated and having suffered no loss at all. Formally, the judge calculates the plaintiff’s willingness to accept $WTA(L)$ to bear the loss L , which denotes the minimum amount of money C the plaintiff is willing to accept in order to choose $(-L, C)$ over the status quo $(0, 0)$ in which no injuries are sustained and

hypothetical cases. These variables are not significantly associated with the damages awarded, leading to the conclusion that “all that matters is whether the company performed a risk analysis in advance.”

⁵Though Viscusi’s experiment pertains to punitive damages, we argue for interpreting the damages awards as compensation for the victims’ loss. Supporting this interpretation, when jurors were given different estimates of the value of life they adjusted the level of damages proportionately (which would not be warranted under a deterrent view of punitive damages). Similarly, the assessed damages are broadly insensitive to the probability of injury or the cost of precautions.

correspondingly no damages are received from the defendant. A rational judge sets:

$$\text{WTA}(L) = \min\{C \text{ s.t. } u(C - L) \geq u(0)\}$$

Assume, as in Equation (3), that the utility of the plaintiff is proportional to $C - L$, assigning equal weights ($\theta_1 = \theta_2 = 1/2$) to damages and loss, both expressed in dollars. Then, the rational judge sets $\text{WTA}(L) = L$, irrespective of the provision of meritless motions or of information about precautions.

In contrast, the salient thinker includes other information – such as the possibility to dismiss the first case or the expectation that precautions are taken in the second case – in the case’s context. Specifically, this context \mathbf{C} includes the plaintiff’s hypothetical choice options that are required to evaluate his compensation, namely the actual loss-cost package $(-L, C)$ and the status quo $(0, 0)$, but it also includes any specific loss-cost package $(-\tilde{L}, \tilde{C})$ that is brought to the judge’s attention. In this formalism, the motion to dismiss corresponds to the loss-cost package $\tilde{L} = L, \tilde{C} = 75,000$, while the auto company precaution corresponds to the package $\tilde{L} = 0, \tilde{C} = 16$ million. The willingness to accept computed by the salient thinking judge in context \mathbf{C} is then given by:

$$\text{WTA}^S(L|\mathbf{C}) = \min\{C \text{ s.t. } u^S(C - L|\mathbf{C}) \geq u^S(0, 0|\mathbf{C})\}$$

The judge still calculates the minimum damages C to be paid by the defendant such that the plaintiff chooses $(-L, C)$ over $(0, 0)$, but the superscript S now indicates that the judge’s tradeoff between losses to the plaintiff and costs to the defendant is shaped by salience. In line with Equation (8), when the loss is salient the judge overweighs it in the plaintiff’s utility function, and sets a higher WTA. When instead damages are salient, the judge overweighs the benefit of receiving damages in the plaintiff’s utility function, and sets a lower WTA.

Critically, whether the plaintiff’s loss L or a given cost C is salient depends on which specific loss-cost package $(-\tilde{L}, \tilde{C})$ the judge is reminded of.

Consider first the baseline No Anchor / No Precautions Considered conditions, in which no information on the motion to dismiss or on precautions considered by the company is provided. In this case, the choice context is simply $\mathbf{C} = \{(-L, C), (0, 0)\}$. The reference loss

in this context is $L/2$ and the reference damage cost is $C/2$. In this case, for any chosen level of damages C , losses and damages are equally salient because:

$$\sigma(L, L/2) = \sigma(1, 1/2) = \sigma(C, C/2),$$

which holds by homogeneity of degree zero. This implies that (all proofs in the Appendix):

Lemma 1 *In the baseline experiments (No Anchor / No Precautions Considered), salient thinking judges set the rational damages $C = L$.*

As implied by Equation (8), when damages and losses are equally salient, judicial valuation is not distorted and satisfies the strict liability principle.

Consider now the case where the choice context is $\mathbf{C} = \{(-L, C), (0, 0), (-\tilde{L}, \tilde{C})\}$, where $(-\tilde{L}, \tilde{C})$ is a generic a loss-cost package the judge is reminded of. This could be one of the experimental treatments above (motion to dismiss or information about precautions) or more generally it could be a precedent-setting damage award in a related case. In this context, the reference loss is $(L + \tilde{L})/3$, while the reference damage is $(C + \tilde{C})/3$. Thus, the plaintiff's loss L is more salient than a chosen level C of damages imposed on the defendant if and only if:

$$\sigma\left(L, \frac{L + \tilde{L}}{3}\right) > \sigma\left(C, \frac{C + \tilde{C}}{3}\right).$$

In the remainder of the analysis, we focus on the case where $\tilde{L} < 2L$ and $\tilde{C} < 2C$ (which both hold in the examples in question). In this case, it is immediate to show that the plaintiff's loss is salient if and only if:

$$\frac{C}{L} < \frac{\tilde{C}}{\tilde{L}}, \tag{10}$$

namely, whenever the ratio of the company's cost in terms of damages C relative to the plaintiff's assessed loss L is lower than the corresponding ratio in the "anchor" scenario $(-\tilde{L}, \tilde{C})$. Instead, the company's cost is salient when the damages C are high (relative to the loss) as compared to the anchor scenario $(-\tilde{L}, \tilde{C})$.

This mechanism has far reaching implications. When reminded of a stingy compensation package (small \tilde{C}/\tilde{L}) per unit of loss, the judge views even very small damages as salient,

which reduces the judge’s assessed WTA. When reminded of the possibility of avoiding the plaintiff’s losses altogether (large \tilde{C}/\tilde{L}), the loss incurred by the plaintiff is salient, which boosts the judge’s WTA. More generally, judges set damages as follows:

Proposition 1 *When the anchor $(-\tilde{L}, \tilde{C})$ is provided, the salient thinker sets damages:*

$$WTA^s = \begin{cases} \delta \cdot L & \text{if } \tilde{C}/\tilde{L} < \delta \\ \tilde{C}/\tilde{L} & \text{if } \delta < \tilde{C}/\tilde{L} < 1/\delta \\ 1/\delta \cdot L & \text{if } \tilde{C}/\tilde{L} > 1/\delta \end{cases} . \quad (11)$$

Roughly speaking, the salient thinking judge targets the share of the compensated loss C/L to the ratio \tilde{C}/\tilde{L} prevailing in the comparative situation. In this sense, \tilde{C}/\tilde{L} truly provides an anchor for setting compensation. When C/L is above the anchor \tilde{C}/\tilde{L} , the cost imposed on the defendant is the salient attribute of the case, which tends to induce the judge to reduce C . When C/L is below the anchor \tilde{C}/\tilde{L} , the plaintiff’s loss is the salient attribute of the case, which tends to induce the judge to increase C . This mechanism implies that if in the comparative situation the company bears a small cost $\tilde{C}/\tilde{L} < 1$, the presence of the anchor reduces damages below the initial assessment of the loss L . When instead in the comparative situation the plaintiff bears a small loss $\tilde{C}/\tilde{L} > 1$, the anchor boosts damages above the initial assessment of the actual loss L . Because the costs imposed on the company tend to be set in the same proportion to the loss as that in the anchor, Proposition 1 might be called a model of comparative proportionality of damages.

At this point we can revisit the two experiments. Guthrie et al. (2001)’s motion to dismiss provides an anchor where $\tilde{L} = L$ and $\tilde{C} = \$75,000$. The ratio \tilde{C}/\tilde{L} is equal to \tilde{C}/L , making any compensation C above $\tilde{C} = \$75,000$ salient. According to (11) this depresses the awarded level of damages, consistent with the experimental findings. The consideration of precautions in Viscusi’s (2001) experiment provides an anchor where $\tilde{L} = 0$ and $\tilde{C} = \$16$ million. Here the ratio \tilde{C}/\tilde{L} is extremely high, making losses salient for any positive damage C . This implies, by Equation (11), that the awarded damages are increased, which is also consistent with the experimental findings.⁶

⁶Of course, it is not always possible for judges to exactly match \tilde{C}/\tilde{L} . In particular, when the anchor

The intuition for these results is clear. When adjusting damages around an assessment of losses, the judge faces the difficult task of trading off the non monetary loss incurred by the plaintiff and the monetary cost imposed on the defendant. The principle of comparative evaluation implies that the presence of an anchor for costs affects how this tradeoff is resolved. Making judges think about precedents where sizeable losses receive zero or little compensation induces them to view even fairly low damages as generous, reducing their own estimate of a fair compensation. Making judges think about precedents where compensation was generous, or about counterfactuals which make the losses more salient, induces them to view fairly high damages as stingy, thus increasing their compensation recommendation.⁷

This discussion illustrates how the anchoring mechanism may be pervasive in legal decision making. Anchors can most naturally be formed by precedents, in which case salience provides a force of conformity to precedent preserving the stability of the law. But anchors can also arise from irrelevant information, as illustrated in Guthrie et al.’s experiment, or from judges’ expectations about, or counterfactuals to, the facts of the case, as illustrated in Viscusi’s experiment.

provides very little compensation, $(\tilde{C}/\tilde{L}) < \delta$ – the case of Guthrie et al.’s experiment – damages are salient but the judge still needs to compensate the non-salient loss $\delta \cdot L$. The fact that the plaintiff’s loss is non-salient, however, implies that damages are set below the assessment L . Matching the anchor is also not possible when it assigns a very high cost on the defendant per unit loss of the plaintiff – the case of Viscusi’s experiment. In this case, matching the ratio \tilde{C}/\tilde{L} when losses do occur would require the judge to impose arbitrarily high damages on the defendant. Instead, the judge limits the damages to be commensurate with the salient loss of $(1/\delta) \cdot L$. The fact that the plaintiff’s loss is now salient then implies that compensation is above the initial assessment L .

⁷In the context of riskless consumer choice, this finding parallels the famous beer experiment proposed by Thaler (1985, 1999). Subjects are told to imagine sunbathing with a friend on a beach in Mexico. It is hot, and the friend offers to get an ice-cold Corona from the nearest seller, a hundred yards away, to be brought back and consumed on the beach. Subjects are asked for their reservation price. In the first treatment, the nearest place to buy the beer is a beach resort. In the second treatment, the nearest place is a corner store. Many subjects would pay more for a beer from a resort than for one from the store, contradicting the fundamental assumption that willingness to pay for a good is independent of context.

As we showed in BGS (2013), the salience model explains the Thaler experiment through a mechanism similar to the one illustrated here. Because the salient thinker expects a high price for beer at the resort, he is willing to pay a higher price and still perceive the beer as a good deal (salient quality). In contrast, at the store, where beer is expected to be cheap, a high price would be very salient and the salient thinker is unwilling to pay it.

3.2 Legal Decoys

The logic of the previous section has further implications to areas of law where parties must choose between several courses of action. This is illustrated by the following hypothetical legal counselling case, proposed by Kelman et al. (1996) to undergraduate students at Stanford University:

“The Economics Department of a major university voted, two years ago, to recommend that your client, then an Associate Professor at the University, not be promoted to a tenured position. She claims that she was discriminated against on account of her gender. Your client is interested in (1) being compensated for wrongs done to her and in (2) having the University publicly admit wrongdoing in her case. At the same time, your client is very interested in the progress of women generally and wants (3) to do her part to push for affirmative action plans that would help women in Economics. The University counsel’s office has contacted you and asked you to communicate settlement offers to your client.”

Subjects had to advise their client on which settlement offer to accept. One group of subjects had to choose among two settlement offers (Choice Context 1). The first offer bound the University to an affirmative action plan for the Economics Department without admitting wrongdoing or paying damages. The second offer consisted of a public admission of wrongdoing and \$45,000 in damages. Another group of subjects (Choice Context 2) had to choose among the two offers above plus a third offer, consisting of a public admission of wrongdoing, plus a donation of \$35,000 in the client’s name to her favorite charity. Note that the third offer is clearly inferior to the second offer, because the plaintiff could always accept the \$45,000 damages of the second offer, give \$35,000 to charity, and keep \$10,000 for herself.

The behavior of rational subjects should be the same in both choice contexts. This is because the third settlement offer is dominated by the second (and thus irrelevant). In the experiment, however, behavior changed markedly from Choice Context 1 to Choice Context 2: while only 50% of subjects chose the second settlement offer in choice Context 1, 76% of the subjects chose the second settlement offer in Choice Context 2 when the third offer was included. In line with the decoy effect, the introduction of a dominated settlement offer

increased the valuation of the dominating offer.

The decoy logic of section 2.4 illustrates how salience accounts for this effect: the presence of the dominated third offer, with its lower monetary compensation, increases the salience of the dominating option's higher monetary compensation, and thus boosts the latter's valuation. Formally, suppose that the plaintiff values two broad categories of compensatory measures: non-monetary and monetary. Non-monetary measures include the University's public admission of wrongdoing and/or its pursuit of an affirmative action plan. Denote by g the defendant's dollar value of a public admission of wrongdoing, and by G the value of an affirmative action plan. Monetary measures obviously consist of damages, part of which may be valued less than free cash if earmarked to a specific use (charity in this case). Suppose that the defendant discounts each dollar earmarked to charity by $\phi \leq 1$.⁸ The three settlement offers are then described by:

$$S_1 = (G, 0), \quad S_2 = (g, D_2), \quad S_3 = (g, \phi \cdot D_3),$$

where attribute 1 captures non-monetary benefits, attribute 2 captures monetary benefits, and $D_2 = 45,000$, $D_3 = 35,000$.

If the plaintiff's utility function has $\theta_1 = \theta_2 = 1/2$ in Equation (1), she computes the overall utility of an offer by simply adding monetary and non-monetary attributes. In contrast, the salient thinker attaches a higher weight to the offer's salient dimension, be it its monetary or non-monetary component. For simplicity, assume that i) the plaintiff values the University-wide affirmative action more than the admission of wrongdoing, $G > g$, and ii) the plaintiff is not willing to pay \$45,000 out of her pocket for an admission of wrongdoing by the university, $g < D_2$. These assumptions are reasonable and have two implications. First, S_1 is the best offer from the viewpoint of non-monetary compensation and the worst offer from the viewpoint of monetary compensation. Second, the valuation of S_2 is higher if the damages it entails, rather than non-monetary compensation, are salient.

Consider the behavior of the salient thinkers. In Choice Context 1, which compares offers S_1 and S_2 , the reference non-monetary compensation is $(G + g)/2$ while the reference level

⁸When $\phi = 1$ the plaintiff would individually implement the same, or higher, charity donation of the settlement offer.

of damages is $D_2/2$. The salient attribute for offer S_1 is its total lack of damages, because $\sigma(0, D_2/2) > \sigma(G, (G + g)/2)$ by homogeneity of degree zero. The salient attribute for offer S_2 is its higher damages when:

$$\sigma(D_2, D_2/2) = \sigma(2, 1) > \sigma(g, (G + g)/2) = \sigma((G + g)/2g, 1), \quad (12)$$

which is satisfied if and only if $G < 3g$. If S_2 's damages are salient its evaluation increases (because high damages are the offer's best attribute, $D_2 > g$) so S_2 is more likely to be chosen over S_1 .⁹

Consider now the problem faced by subjects in Choice Context 2, when the decoy offer $S_3 = (g, \phi \cdot D_3)$, is added to the choice set. Now, when choosing between S_1 and S_2 , the salient thinker evaluates these offers in light of the reference attribute values $((G + 2g)/3, (D_2 + \phi D_3)/3)$. The salient attribute for offer S_1 continues to be its total lack of damages (because of homogeneity of degree zero), so its valuation is the same in both contexts. Intuitively, the fact that S_1 is the only option paying zero damages remains the distinguishing feature of this offer. On the other hand, the salient attribute of S_2 is its monetary compensation if and only if $\sigma(D_2, \frac{D_2 + \phi D_3}{3}) > \sigma(g, \frac{G + 2g}{3})$ which is satisfied if and only if $G < g \cdot \frac{7D_2 - 2\phi D_3}{D_2 + \phi D_3}$. Comparing this condition to (12) leads to the following result.

Proposition 2 *The set of conditions under which offer S_2 is chosen is strictly weaker in Choice Context 2 than in Choice Context 1 if and only if $D_2/D_3 > 5/4 \cdot \phi$.*

The introduction of the dominated offer S_3 increases the appeal of the dominating offer S_2 provided the damages paid by S_2 are sufficiently high relative to those paid by S_3 . The intuition is straightforward: when S_3 fares much worse than S_2 in term of damages, the relatively high damages of S_2 become salient, thereby boosting its valuation. It is easy to see that the condition in Proposition 2 is met by the experiment's monetary values $D_2 = 45,000$, $D_3 = 35,000$.

As with the anchoring effect of Section 3.1, the principle underlying the model's account of decoy effects in settlement offers is that of comparative evaluation. When solving a complex

⁹Formally, if $G < 3g$, monetary compensation is salient for both S_1 and S_2 , so S_2 is chosen if and only if $\delta \cdot g + D_2 > \delta \cdot G$. If instead $G > 3g$, monetary compensation is salient for S_1 and non-monetary compensation is salient for S_2 , so S_2 is chosen if and only if $g + \delta D_2 > \delta \cdot G$, which is a stronger condition.

problem such as choosing among multi-dimensional offers, individuals do not harness their own absolute metric for monetary and non-monetary compensation and apply it to the offers at hand. Rather, they evaluate different policies in context, by comparing their different features, and by focusing on each offer's most distinguishing aspect. When an inferior offer is added, the advantage of the dominating offer becomes salient, increasing its valuation.

3.3 Framing of Risky Prospects

We conclude our analysis by considering the role of salience in shaping risk attitudes in the legal context.

It is well known that individuals' risk preferences depend on whether payoffs are framed in terms of gains or losses, even though such frames are normatively irrelevant (Kahneman and Tversky, 1979). Rachlinski (1996) investigates this issue in a domain where risk preferences are central to judicial process, namely the decision of whether to settle or litigate. A simple litigation problem was presented to law students, half of whom played the role of counsel to the plaintiff and half the role of counsel to the defendant. Subjects had to advise their client on whether to settle or litigate. In particular, “[p]laintiff-subjects had to choose between a certain \$200,000 settlement offer and a 50% chance of winning \$400,000 at trial (and a corresponding 50% chance of winning nothing). Defendant-subjects had to choose between paying a \$200,000 settlement offer and facing a 50% chance of losing \$400,000 at trial with a corresponding 50% chance of losing nothing.” Rachlinski found that 77% of plaintiff-subjects advised to settle, but only 31% of defendant-subjects advised to settle. Plaintiff-subjects were risk averse, defendant-subjects risk seeking.

This behavior is inconsistent with Expected Utility Theory. Both plaintiff- and defendant-subjects choose between a safe amount and a mean preserving spread around it. Under common risk preferences, plaintiffs and defendants should both prefer to settle (if risk averse) or to litigate (if risk seeking). The conventional explanation for the observed instability of risk preferences relies on Prospect Theory's S-shaped value function, which is assumed to be concave for gains and convex for losses. This induces risk aversion for plaintiffs (who choose between a sure gain and a risky gain) and risk seeking for defendants (who choose between a sure loss and risky loss).

Saliency Theory can explain this instability of risk attitudes as a consequence of context dependence: changing the framing from gains to losses changes the lottery outcome perceived to be salient, thereby changing risk preferences.¹⁰ Saliency thus allows us to provide a novel unified psychological explanation for framing, anchoring and decoy effects.

To see how the saliency model can generate framing effects, consider the choice faced by plaintiff-subjects in Rachlinski’s experiment. In this frame, the available options are a sure gain coming from settlement $S = (200, 200)$ and a risky litigation $R = (0, 400)$ (for simplicity, in this analysis we write the payoffs in units of thousands of dollars). Each option is described by their payoffs in the two states of the world, which occur with probabilities $\theta_1 = \theta_2 = 0.5$. The plaintiff loses the litigation in state 1, wins in state 2, while the settlement gives the same payoff in both states. The salient outcome for litigation is the loss state 1 because $\sigma\left(0, \frac{200+0}{2}\right) > \sigma(1, 400/300)$ (due to homogeneity of degree zero). In the spirit of Experiment 1 of Section 2.3, because the downside is salient, litigation is valued less than its expected value of 200. As a consequence, plaintiff-subjects prefer settlement (a sure gain of 200) to litigation, exhibiting risk averse behavior.

Consider now the choice faced by defendant-subjects. In this frame, the available options are a sure loss coming from a settlement $S = (-200, -200)$ and a risky litigation $R = (0, -400)$, where $-$ consistent with the convention above $-$ in state 1 the defendant wins and the plaintiff loses. As before, $\theta_1 = \theta_2 = 0.5$. The salient outcome of litigation is now the winning state 1 because $\sigma\left(0, \frac{-200-0}{2}\right) > \sigma(1, 400/300)$ (again due to homogeneity). Because the litigation’s upside (getting a payoff of 0) is salient, litigation is valued above its expected value of -200 . As a consequence, defendant-subjects prefer litigation to settlement (a sure loss of 200), exhibiting risk seeking behavior.

Changing how payoffs are framed, from the gain to the loss domain, alters risk preferences by changing the nature of the salient outcome. In the gain domain, the downside risk of getting zero is salient, while in the loss domain the upside risk of getting zero is salient. Given that the problem faced by defendants simply subtracts \$400,000 from the problem faced by

¹⁰In other settings, both plaintiff and defendants stand to gain from the litigation process, as for example in copyright sharing decisions. When all payoffs are positive, a salient thinker may still shift from risk seeking to risk averse behavior depending on the framing of the gains, as in the example of Section 2.3. As we explained there, this behavior is incompatible with Prospect Theory.

plaintiffs, the framing effect shown here has the same nature of the context dependent shift in risk attitudes discussed in Section 2.3.

Even though the experiment just discussed involves legal counseling of plaintiffs or defendants, the same effects are likely to apply to the judiciary, given that judges often play an active role in supervising settlement talks. Guthrie et al. (2001) design an experiment with 167 federal magistrate judges, also documenting a shift in pro-litigation attitudes as the problem is reframed from gains to losses. But the issue is broader, as many judicial decisions involve choice among risky prospects. For example, Guthrie et al. (2001) document that judges are more likely to vote for a risky reorganization plan when creditors' payoffs are framed as losses rather than as gains.

4 Conclusion

Judicial decisions of necessity must balance various aspects of the case, ranging from the severity of misconduct and damages to the nature of the penalties. The process of reaching such balance or proportionality is vulnerable to the influence of salience: the idea that certain aspects of the case stand out in judges' attention and as a consequence influence judicial decisions. Of course, there are many legal mechanisms that try to limit salience or other psychological influences, most importantly precedents. The subtle point here, however, is that via the mechanism of comparative evaluation, these very mechanisms can shape judicial decision over and above their original intent. Indeed, the experimental evidence suggests (though does not prove) that salience can induce even experienced judges to extrapolate too much from past cases or counterfactual scenarios, distorting standard legal reasoning.

We have illustrated our approach using the analysis of experimental evidence, but here we consider some broader implications of our framework. In many situations, past cases anchor judicial decisions. This may explain why lawyers spend so much time relating the facts of the case to their favored precedents. But salience may also render judicial decisions inconsistent across cases, especially when driven by salient but legally irrelevant facts. Plaintiffs or defendants might be conspicuously unappealing individuals, prompting the court to find ways to put them down. The plaintiff might have suffered extreme harm with no evidence

of negligence, encouraging courts to nonetheless find responsibility so harm is not so salient relative to compensation. Such hard cases draw judicial attention to the features that should perhaps matter little, but end up mattering more.

A side benefit of salient thinking is that it encourages legal evolution, which usually takes the form of distinguishing cases from precedents (see e.g. Gennaioli and Shleifer 2007). Cases that differ from precedents along a salient dimension are likely to be those in which precedent leads to decisions whose unfairness or disproportionality if current law is applied stands out. The evolution of exceptions to the Economic Loss Rule, described by Niblett, Posner, and Shleifer (2010) has some of that character. Of course, salience can also encourage distinguishing based on legally irrelevant features of the case, and not just efficient legal evolution. To the extent that they shape the salience of the facts, cases that come to the judge's mind take on great significance (Radin 1925).

It should not be surprising the legal procedure has evolved to constrain the influence of salience and other cognitive biases. The rules of evidence control what is presented to judges and juries, and an elaborate body of law has developed to exclude information that might be prejudicial. The legal system seems highly conscious of the influence of salience; we doubt, however, that this influence is wholly eliminated.

References

- Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2012. "Salience Theory of Choice under Risk." *Quarterly Journal of Economics* 127 (3): 1243 – 1285.
- Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2012b. "Salience Theory of Choice under Risk." NBER working paper 16387.
- Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2013. "Salience and Consumer Choice." *Journal of Political Economy*, forthcoming.
- Epstein, Lee, William Landes, and Richard Posner. 2013. *The Behavior of Federal Judges: A Theoretical and Empirical Study of Rational Choice*. Cambridge: Harvard University Press.
- Frank, Jerome. 1973. *Courts on trial: Myth and Reality in American Justice*. Princeton University Press.
- Gennaioli, Nicola, and Andrei Shleifer. "The Evolution of Common Law." *Journal of Political Economy* 115 (1): 43 – 68.
- Guthrie, Chris, Jeffrey Rachlinski and Andrew Wistrich. 2001. "Inside the Judicial Mind." *Cornell Law Review* 86 (4): 777 – 830.
- Huber, Joel, John Payne, and Christopher Puto. 1982. "Adding Asymmetrically Dominated Alternatives: Violations of the Regularity and Similarity Hypothesis." *Journal of Consumer Research* 9 (1): 90 – 98.
- Kahneman, Daniel and Amos Tversky. 1979. "Prospect Theory: an Analysis of Decision under Risk." *Econometrica* 47 (2): 263– 292.
- Kelman, Mark , Yuval Rottenstreich, and Amos Tversky. 1996. "Context-Dependence in Legal Decision Making." *Journal of Legal Studies* 25 (2): 287 – 318.
- Niblett, Anthony, Richard Posner, and Andrei Shleifer. "The Evolution of a Legal Rule." *Journal of Legal Studies* 39 (2): 325 – 547.
- Posner, Richard. 2008. *How Judges Think*. Cambridge: Harvard University Press.
- Rachlinski, Jeffrey. 1996. "Gains, Losses, and the Psychology of Litigation." *Southern California Law Review* 70 (1): 113 – 186.

- Rachlinski, Jeffrey, Andrew Wistrich and Chris Guthrie. 2013. "Altering Attention in Adjudication." *UCLA Law Review* 60: 1586 – 1618.
- Radin, Max. 1925. "Theory of Judicial Decision or How Judges Think." *American Bar Association Journal* 11: 357 – 362.
- Thaler, Richard. 1985. "Mental Accounting and Consumer Choice." *Marketing Science* 4 (3): 199 – 214.
- Thaler, Richard. 1999. "Mental Accounting Matters." *Journal of Behavioral Decision Making* 12: 183 – 206.
- Tversky, Amos, and Daniel Kahneman. 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science* 185 (4157): 1124 – 1131.
- Tversky, Amos and Itamar Simonson. 1993. "Context Dependent Choices." *Marketing Science* 39 (10): 1179 – 1189.
- Viscusi, W. Kip . 1999. "How Do Judges Think about Risk?" *American Law and Economics Review* 1 (1): 26 – 62.
- Viscusi, W. Kip . 2001. "Jurors, Judges and the Mistreatment of Risk by the Courts." *Journal of Legal Studies* 30 (1): 107 – 142.

Appendix: Proofs

Proof of Lemma 1. In the No Anchor condition, the choice set is $\mathbf{C} = (-L, C), (0, 0)$. In this context, the reference loss is $L/2$ and the reference damage cost is $C/2$. Homogeneity of degree zero of the salience function then implies $\sigma(L, L/2) = \sigma(1, 1/2) = \sigma(C, C/2)$, so that for any chosen level of damages C , losses and damages are equally salient. As a consequence, $u^S(C - L|\mathbf{C}) = (C - L)/2$, while $u^S((0, 0)|\mathbf{C}) = 0$. Then WTA^S equals $\inf\{C \text{ s.t. } C - L \geq 0\} = L$. Therefore, the salient thinking judge sets the rational level of damages $C = L$. ■

Proof of Proposition 1. The reference loss level in $\mathbf{C} = \{(0, 0), (-L, C), (-\tilde{L}, \tilde{C})\}$ is $\bar{L} = (L + \tilde{L})/3$, while the reference level of damages is $\bar{C} = (C + \tilde{C})/3$. Thus, the salience of losses and damages of option $(-L, C)$ are, respectively

$$\sigma\left(1, \frac{1 + \tilde{L}/L}{3}\right), \quad \sigma\left(1, \frac{1 + \tilde{C}/C}{3}\right)$$

Recall from the text the assumption $\tilde{L} < 2L$ and $\tilde{C} < 2C$ (which holds in the examples). It follows that losses are salient when

$$\frac{\tilde{C}}{\tilde{L}} > \frac{C}{L} \tag{13}$$

Note that as C varies in the range $(\tilde{C} \cdot \frac{1}{2}, \tilde{C} \cdot \frac{L}{\tilde{L}})$, it can take values larger or smaller than the reference damage \bar{C} .

Recall the definition of willingness to pay:

$$WTA^S = \min\{C \text{ s.t. } u^S(C - L|\mathbf{C}) \geq u^S(0, 0|\mathbf{C})\}$$

Consider first the case where the awarded damages are salient, so that they fully compensate for losses when $C = \delta L$. This compensation is indeed salient when the reverse of (13) holds. Thus, $WTP = \delta L$ whenever $\tilde{C}/\tilde{L} < \delta$.

Consider now the case where $\tilde{C}/\tilde{L} > \delta$. If the damages were set at $C < L \cdot \tilde{C}/\tilde{L}$ then the loss would be salient, so that the damages can be hiked up all the way till L/δ , as long as the

saliency ranking does not change: $WTA = \min \left\{ L \cdot \frac{1}{\delta}, L \cdot \tilde{C}/\tilde{L} \right\}$. Therefore, $WTP = L \cdot \tilde{C}/\tilde{L}$ for $\tilde{C}/\tilde{L} < 1/\delta$ and $WTP = L/\delta$ for $\tilde{C}/\tilde{L} > 1/\delta$ (and $\tilde{C} < 2C$, namely $\tilde{C} < 2L/\delta$). ■

Proof of Proposition 2. Because the valuation of offer S_1 is constant in both contexts, the inclusion of the decoy offer $(g, \phi \cdot D_3)$ relaxes the conditions under which S_2 is chosen if and only if it relaxes the condition that S_2 's upside – its monetary compensation, or damages – is salient. In context 1, S_2 's damages are salient if and only if $G < 3g$. In context 2 (with the decoy), S_2 's damages are salient if and only if $G < g \cdot \frac{7D_2 - 2\phi D_3}{D_2 + \phi D_3}$. Thus, the decoy makes it more likely for S_2 to be chosen if and only if $g \cdot \frac{7D_2 - 2\phi D_3}{D_2 + \phi D_3} > 3g$, namely $D_2/D_3 > 5/4 \cdot \phi$. ■