Conceptual foundations: A bird's-eye view[*]

Jonathan Baron[†]   Tess Wilkinson-Ryan[‡]

1 Introduction

The term "behavioral law and economics" was apparently coined by Jolls, Sunstein and Thaler (1998).[1] The idea was that the economic approach to law, as represented in the field called "law and economics", had imported a kind of micro-economic theory that assumed that agents were rational in a certain sense. Yet this sort of theory itself was under challenge, at first from psychology but ultimately from economists themselves. The challenge was not so much against the concept of rationality, although that too had been challenged by philosophers. Rather, it was that, in fact, agents did not *behave* rationally much of the time. If law were going to use economic theory as the basis of analysis, then it would have to acknowledge these behavioral findings and somehow incorporate them.

   In this chapter, we shall try to outline the conceptual foundations of the field that has emerged. We shall concentrate on the "behavioral" concepts, many of which have come from psychology.[2] In an effort to be concrete, we shall illustrate our points with examples, with the hope that we do not tread too much on the territory of other chapters in this book.

2 Where is behavior relevant?

The field covers several kinds of decision makers: courts (judges and juries), individuals (bringing a lawsuit, committing a crime, making and honoring contracts, etc.), and lawmakers (regulators,

---

[*]Citation: Baron, J., & Wilkinson-Ryan, T. (in press). Conceptual foundations: A birds-eye view. In K. Zeiler & J. C. Teitelbaum (Eds.), *Research handbook on behavioral law and economics*. Northampton, MA: Edward Elgar.

[†]Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104–6241, U.S.A. Email: baron@psych.upenn.edu.

[‡]University of Pennsylvania Law School

[1]It was analogous to "behavioral decision theory". Now we also have "behavioral public finance" and behavioral many other things.

[2]Some have come from experimental economics, which psychologists tend to regard as a subset of experimental social psychology, constrained by various rules such as forbidding deception.

legislators, and, ultimately, voters). Naturally, some of these decisions have been studied by psychologists, economists and legal scholars who lack an interdisciplinary perspective (e.g., Dhami & Ayton, 2001, for courts, and Becker, 1968, for crime). Unfortunately for the student, this work cannot be ignored, but we shall not emphasize it here. We shall emphasize the general framework provided by "behavioral decision theory", also known as the field of "judgment and decision making" (JDM), which we think is a better term).

3   Normative, descriptive, prescriptive

JDM concerns three types of theories or models. Normative models are standards that researchers use to evaluate judgments and decisions. In some cases the model is simply "the right answer", as when people are asked to estimate quantities such as cost or frequency of mishaps. Such models are based on *correspondence* between a judgment and some real state of the world. In other cases, the right answer depends on a person's values and beliefs, so the only possible criterion is one of internal consistency, or *coherence* (Hammond, 1996; Dunwoody, 2009). Some judgments, such as probability judgments, can be evaluated both ways. A weather forecaster who says that the probability of precipitation is .5, rain .3, and snow .3, is incoherent. If events she says have .9 probability consistently occur with a frequency of .7, her judgment is mis-calibrated, lacking in correspondence.

Descriptive models explain judgments and decisions. Of particular interest are models that explain why judgments are systematically non-normative. Judgments can be non-normative because of random error, or because they are systematically *biased* in one direction. An example of a bias is that the sum of probabilities assigned to subsets (rain and snow) usually exceeds the probability assigned to the whole set (precipitation). Fox and Tversky (1998) have developed a mathematical model that explains this bias and other judgments, based on the idea that subjective probabilities are distorted.

Much of the tension in the JDM field concerns the issues of correspondence, coherence, and biases. One common argument is that, in real environments, the processes that lead to incoherence are sometimes best at achieving correspondence (Gigerenzer, Todd, et al., 1999).

Prescriptive models are ways of fixing or avoiding non-normative judgments and decisions. Several methods are relevant here, including the design of law itself, and its implementation through the wording of jury instructions, contracts, etc. The JDM field also is concerned with tools for the formal analysis of judgments and decisions, such as aggregation of judgments to make a prediction or analysis of risk regulation in terms of costs and benefits, where the benefits are often subjective. A recent innovation in prescriptive theory is the idea of *decision architecture*, the idea of designing decisions themselves so that people are more likely to choose the best option

according to normative theory (Thaler & Sunstein, 2008).

We should not assume that all biases need to be corrected or worked around. Some may be side effects of psychological processes that are optimal in typical environments. Yet, on the other hand, we should not assume that relevant legal environments are typical. Processes that usually serve us well may break down when we are asked to assess tort penalties for unusual harms, for example.

4   Normative models

Normative models in behavioral law and economics start with economic theory itself, particularly micro-economic theory, which is concerned with choices and transactions among decision makers. Microeconomics itself, however, is a mix of all three types of models, as well as general frameworks for expressing more specific models, such as demand curves, indifference curves, elasticity, and the marginal rate of substitution. The main normative models involve maximization of utility, profit, or wealth. An important example is expected-utility theory, which applies to decisions in which outcomes are uncertain.

Often these normative models are also assumed to be descriptive models, but this is where behavioral data often says otherwise. Micro-economics also includes much of game theory, which is a normative model for strategic interaction of multiple parties, in which outcomes depend on the decisions of more than one person.

In addition to standard micro-economics, analysis often relies on welfare economics as a normative model, particularly concerning distributional questions. Welfare economics is concerned with maximization of some measure of social welfare in a society

Here we concentrate on the models that are most useful in the behavioral side. In the psychology of judgment and decisions, other normative models include statistics and logic.

4.1   Utility and probability

The most standard normative models in this area are those involving utility and probability. Utility is a numerical measure of good, or, for some, of the extent to which goals are achieved. Utility theory, actually a set of theories, specifies that the best option in a choice is the one with the highest utility. Utility is, however, a function of outcomes, and typically the outcomes of a decision are uncertain. In this case, we want the highest *expected* utility (EU). The expected utility of an option is $\sum_i p_i u_i$, that is the sum over all possible outcomes i of, for each outcome, its probability times its utility. We can think of it is the average utility if the decision were repeated. Probability is a function of individual beliefs, and utility is a function of individual values, so, often, the only way we have of determining whether people follow this model is to test coherence over several

decisions. (See section 5.5.1.)

In EU theory, we usually assume that the utility of money is "marginally declining" or "concave." This means that a given difference in money in your bank account, say $1,000, has a bigger effect on utility when the amounts of money are small. A difference between $1,000 and $2,000 has more effect the difference between $1,000,000 and $1,001,000, or even the difference between $2,000 and $3,000. If you plot a graph of utility against money, it is concave when viewed from the bottom; its slope is decreasing as you move to the right. A consequence of this assumption is that we are averse to risks. If you have a choice of $2,000 or a 50/50 gamble on $1,000 or $3,000, you will take the $2,000, because the potential utility loss of $1,000 is greater than the potential gain, and they are equally likely.

Formal development of utility theory (e.g., Krantz, Luce, Suppes & Tversky, 1971) shows how we can assign numbers to outcomes or options, representing their utility. Conformity to utility theory in its various forms implies conformity to several subsidiary principles. One is transitivity: if you prefer option A to B, and B to C, then you must also prefer A to C. Transitivity implies that utility is an ordinal scale: any set of utility numbers that accounts for a set of choices cannot yield different orderings of the choices in the set. Obviously, transitivity is a property of numbers themselves.

But utility theories require not only that we can order outcomes but also that we can order *differences* of pairs of outcomes. Often we are confronted with choices between, say A1 together with B1 vs. A2 with B2. We must make a trade-off. If A1 is better than A2 to a greater extent than B2 is better than B1, then we should choose the first pair. Thus, the utility scale must be more than ordinal. It must be an interval scale, allowing the comparison of intervals. Differences must be transitive too.

Various mathematical theories have started with these principles (plus others) and have derived theorems concerning the representation of utility and its uniqueness. Uniqueness refers to what transformations can preserve the representation. An ordinal scale is preserved by any monotonic transformation. An interval scale is preserved by any linear transformation, that is, by multiplication and addition. Utility is an interval scale, but not a ratio scale. Ratio scales cannot be transformed by addition because have unique zero points. Interval scales do not. Utility is thus a measure like time, with no natural zero point, but not like mass. Utility is always relative. That is all we need, because we are always comparing options to one another, so we need only differences.

Various theories derive the interval property in different ways. For example, one (non-obvious but important) way of deriving expected utility theory involves the "sure-thing principle" (Savage, 1954), which is based on an analysis of decisions into options (e.g., umbrella or not), uncertain states of the world (e.g., rain or sunny), and consequences, which are a function of both the option and the state. The sure-thing principle states roughly that if, for a given state of the world, the

consequence is independent of the option chosen, the nature of that consequence does not affect the choice of options. This principle implies, ultimately, that option differences in different states of the world are independent of each other, so they may be added.

Interestingly, and non-trivially, the coherence principles of probability theory can be derived from EU itself. The two main principles are the addition rule and the multiplication rule.[3]

Another relevant version of utility theory concerns decisions over time. Here the outcomes may occur at different times, and the utility of later outcomes may be discounted. We set aside the question of when such discounting is rational, and why it might not be in some cases. But we must assume that the relative utility of outcomes (or their differences) does not itself change with the passage of time alone. An implication is that, if you prefer L (a large reward) at time T2 to S (a smaller reward) at (an earlier) time T1, this preference should not change as time passes and you move closer in time to T1. Decisions are independent of when they are made, so long as the outcomes and their dates are the same. This principle of delay independence implies that discounting must be exponential: over a given interval of time (like a month), the utility of an outcome at the end of the interval is a constant percentage of its utility at the beginning of the interval. Economists often assume exponential discounting for money, even though the utility of money is not generally assumed to be a linear function of the amount.

4.2   Utilitarianism and welfare economics

A natural extension of EU is utilitarianism, a normative model for decisions that affect many people, such as laws and regulations. Utilitarianism holds that we should maximize utility over all people. Famously, it implies that harm is justified only if it is necessary to increase total good. This was Bentham's (1843/1948) justification for legal punishment, and the idea was fully incorporated into economic theories of punishment (Becker, 1968; Shavell, 2004).

The extension from EU to utilitarianism is natural because we can think of many decisions either as decisions affecting one random person under uncertainty or as social decisions affecting many people. For example, the decision about whether to require a vaccination involves a trade-off between disease prevention and side effects. From the individual perspective, each person has some probability of getting the disease if not vaccinated, or of side effects if vaccinated. We can use EU theory to find the best option for that random person. From the social perspective, we have numbers of people instead of probabilities, but the analysis is the same, and the results should agree.

Formally, utilitarianism treats people as independent in much the same way in which EU treats states of the world as independent. We can compare options by looking at the differences between

---

[3]If A and B are mutually exclusive, then $p(\text{A or B}) = p(\text{A}) + p(\text{B})$; and $p(\text{C and D}) = p(\text{C}|\text{D})p(\text{D})$. The expression $p(\text{C}|\text{D})$ is the conditional probability of C given D, the belief we would have in C if we knew D.

the two options for each person, some of which can be negative, and adding.

Welfare economics has taken a broader view (Adler, 2012). Instead of maximizing total utility, it seeks to maximize a *social welfare function*, a function that computes the welfare of society from the welfare of all the individuals. This theory retains the assumption that nothing else matters, but it includes utilitarianism as a special case. Almost all theories assume that the utility of money (or its contribution to social welfare) declines as the amount of individual income or wealth increases. A hundred dollars means more to a poor person than to a rich one. Thus, the general approach is consistent with some sort of redistribution.

As Shavell (2004) points out, for most legal issues, it does not matter whether we apply utilitarianism, some other social welfare function in common use, or even the most naive method of aggregating welfare, the total of economic wealth. This argument is stronger if we assume (as Shavell does) that distributional concerns should not affect the law, except in one place, namely taxation (broadly defined so as to include negative taxation and social insurance). Thus, only in taxation does it matter whether we try to maximize (for example) utility or wealth.[4] Taking this point of view, Kaplow and Shavell (2002) analyze a number of legal principles. They point out that non-welfarist legal principles, such as those based on fairness, can generally be understood as leading to violations of ex-ante Pareto optimality, that is, application of such rules could possibly make some people worse off, in terms of their individual expectations, while making nobody better off.[5] They take their argument to support, in broad terms, the economic theory of law as a normative basis.

There are other normative models. One concern negotiations, which is theoretically central to the analysis of contracts. All of those that we consider are consistent with utility theory in general. Some of the models for decisions with multiple parties, however, do not assume (as utilitarianism and social-welfare functions usually do) that utilities of different individuals can be compared, so they apply some weaker standard such as Pareto optimality, which holds that, if outcome A is better than B for one person and worse for nobody, then A should be preferred.

---

[4]One argument for Shavell's assumption is that the attempt to craft legal rules with distributional concerns in mind is a crude tool. For example, we might tilt the balance toward plaintiffs in tort law on the ground that they are generally poorer than defendants. But that is not always true. Another argument is that, if some otherwise-optimal law does tilt toward the rich, we can optimally correct this effect by making taxes more progressive. An argument on the other side is that such corrections do not happen in the real world, so we are stuck with a second-best system. Also, an appropriate correction through progressive taxation for one type of case might be large and thus an over-correction for another type, or it might reduce incentive enough so as to reduce total utility.

[5]The same principle of ex-ante Pareto optimality is also inconsistent with non-utilitarian social-welfare theories (Adler, 2012), but Kaplow and Shavell do not go down this road.

4.3 Game theory

One of these weaker theories, which is quite influential, is game theory (Luce & Raiffa, 1957). Game theory applies to situations in which several people make decisions. The outcome for each person depends on the decisions of several people (which may include the person in question). Games like chess and checkers are an obvious example, with two players, but many games have more players. The theory is rich and complex, with applications to many practical problems such as the design of markets (Roth, 2008) and negotiation (Raiffa, 1982).

Game theory has provided a set of concepts that are applied more broadly in the study of decision making. One is the distinction between cooperative and non-cooperative games. Cooperative games are of particular interest because they include prisoners' dilemma games and the extension of those to multiple players. In these games each player has a choice between two options, one of which is best for the player, defection, and the other of which is better for everyone else, cooperation. (The "better for everyone else" may mean that some good is provided to everyone if some minimal number cooperate, which could be everyone. Or it could mean that the total is larger, if we are willing to sum across people.) A great variety of games of this type have been studies and analyzed. They include what are called social dilemmas, public-goods games, n-person prisoners' dilemmas, and commons dilemmas; these terms overlap extensively in their meaning. Examples of cooperative acts are paying taxes voluntarily, respecting the property of others, restraining one's own pollution, restraining use of common resources such as water or fish, and limiting family size. Arguably, participation in democratic government (e.g., voting) is also a form of cooperation (Baron, 2012). Laws often punish defection, thus reducing the self-interested benefit of defection, possibly to the point where it is disappears. In some cases the law does not fully deter defection by reducing its expected benefit to zero; for example, penalties for under-paying income taxes, taking into account the chance of getting caught, rarely reduce the expected-value of underpayment to zero. An interesting fact is that people often cooperate spontaneously when the cost of doing so is not too great (Dawes & Thaler, 1988). Cooperation is often enforced through social norms (Bicchieri, 2006).

When game theory is taken as a normative model, it typically assumes that people act in their narrow self-interest. Thus, defection is "rational" in social dilemmas. Often defection is also justified by EU theory of a sort that considers only self-interested utility (hence ignores altruism, the utility for other people's achievement of their goals), but EU and game theory do not quite always agree. For example, suppose some good is provided only if everyone cooperates, and, in this case, each player would be worse off by defecting (because the good is large enough to make up for the cost of cooperation). Game theory in this case would typically say that the game has two equilibria: everyone cooperates or everyone defects, but it would not specify the implications of this fact for each player's choices. EU theory could do so, however, by taking into account the

probabilities of each player for what the other players will do. If you think it is sufficiently likely that everyone else will cooperate, then you should cooperate.

Utilitarianism, when applied to similar cases, will usually prescribe cooperation, unless the benefit of cooperation is too small, given the number of other cooperators, or unless the theory gives special consideration to the decision maker's cost of cooperation. It could give such special consideration, for example, by assuming that an individual's willingness to sacrifice self-interest is limited and that the decision must be made so as to maximize utility for all, taking this limit into account as if it were an external fact (Baron & Szymanska, 2010).

5 Descriptive theory

Economic theory is often descriptive as well as normative. It is used to make predictions about effects of economic interventions, and to some extent it is useful. For example, considerable evidence supports the economic theory of the relation between crime and punishment, which holds that crime is deterred more by higher penalties or higher probability of apprehension (e.g., Fisman & Miguel, 2007).[6] But we concentrate here on predictions that do not follow naturally from economic theory.

5.1 Heuristics and biases

Kahneman and Tversky (1972) introduced the idea of heuristics as a way of explaining biases (systematic departures from normative standards). For example, one bias is the neglect of base rates in probability judgment, which can result from the "representativeness heuristic". In a classic study (Kahneman & Tversky, 1973), subjects were given a description of "Tom W", which resembled a stereotypical computer-science graduate student. They were asked to rank the probabilities that Tom was in a number of different graduate fields, including computer science and "social sciences". Another group was asked to rank Tom's similarity to the typical student in each field, that is, the extent to which Tom was representative of that field. A third group was asked to rank the sizes of the different fields. Of interest was that the probability rankings matched the similarity (representativeness) rankings almost perfectly and had essentially no relation to the sizes of the different fields, although the subjects judged these sizes quite accurately. Normatively, the sizes, the base rates, matter. (In the extreme case, if the field had only one student, Tom would not be very likely to be in that field even if he matched the typical student fairly well.) But people ignore

---

[6]Fisman and Miguel examined parking tickets by United Nations staff, finding that national origin affected the number of tickets when penalties were absent because of diplomatic immunity, but when immunity was removed all tickets decreased dramatically. It should be noted that Robinson and Darley (2004) claimed to find minimal effects of penalties on crime, but their analysis was mostly limited to small changed in the law.

these base rates. Arguably, base rates are even excluded by law when they are normatively relevant, as in the exclusion of evidence of prior convictions.

The term "heuristic" was invented by the mathematician George Polya (1945), who thought of a heuristics as a weak method, something that might help in solving a problem. It is a rule that can be applied easily when we are stuck on a problem. Kahneman and Tversky recognized the value of heuristics but argued that they could also mislead, especially in cases, unlike mathematics, where the answer they yielded could not be easily checked.

Legal scholars were quick to seize on the idea of heuristics to explain various phenomena in the judicial system, but this literature has tended to rely heavily on the first few heuristics that Kahneman and Tversky and others enumerated. Many others have been proposed in the last few decades (Baron, 2008).

More recently, Sunstein (2005) has argued that heuristics are used for moral judgments, and these moral heuristics often have legal implications. These heuristics cause biases, if the normative model of utilitarianism is used as a standard.[7]

An example of a moral heuristic is "do no harm", a rule against causing harm through action, even to prevent greater harm (which would result from omission) (Baron, 1996). The resulting bias is called "omission bias", a bias in favor of harmful omissions over less harmful acts (Ritov & Baron, 1990; Spranca, Minks, & Baron, 1991; Baron & Ritov, 2009). Omission bias is probably behind opposition to (and laws against) active euthanasia. It may also underlie the reluctance of regulatory agencies to approve drugs that may have harmful side effects. It is arguably built into the law, in the absence of "bad Samaritan" laws in most jurisdictions. An exception is tort law, where negligence is typically an omission, not an act. Omission bias is reduced in cases where the decision maker has responsibility for the welfare of the potential victim.

People show great variation in omission bias. Some do not show it at all in many situations. Much of this variation seems to be related to how people think about causality. A "but for" (*sine qua non*) view of causality would not show any bias. The bias is found when people think about causality in physical terms. If a vaccine causes a side effect, we can imagine a sequence of causal steps between the injection and the disease. If failure to vaccinate causes a disease, no such a series of steps between behavior and outcome are apparent.

5.1.1 Attribute substitution, and the isolation effect

Many heuristics work through what Kahneman and Frederick (2002) call "attribute substitution." The idea is that we estimate some value by using another value that is correlated with the first, usually because the latter is easier to judge. Thus, very young children who have trouble count-

---

[7]Sunstein does not endorse this standard but almost all of his examples of biases are biases as defined by it.

ing beyond 5, when asked which of two rows of coins has more coins, will pick the longer row regardless of the number (Lawson, Baron, & Siegel, 1974).

The "isolation effect", another general mechanism for heuristics, is that people attend to what is available "on screen" and tend to ignore what they could easily infer (Kahneman & Lovallo, 1993).[8] When evaluating policies, the intended purpose of the policy is what is on screen, and people often ignore secondary effects. For example, they prefer taxes on businesses to income taxes, but when they are asked who pays the business tax, many people switch (McCaffery & Baron, 2006).

Both mechanisms may be involved in judgments about risk. When people are asked how to allocate funds to reduce risks, and when they are given options in terms of proportion of risk reduced, they tend to allocate more money when the proportion is greater, regardless of the absolute size of the risk. This results in larger allocations to large reductions in tiny risks (e.g., McDaniels, 1988). This effect may account for a substantial part of the discrepancies in risk regulation noted by Breyer (1993) and others: some risks are under-regulated in terms of cost per life saved, while others are over-regulated. The latter tend to be risks that affect fewer people.

5.1.2 Simple heuristics and adaptive cognition

The idea of heuristics is two-sided. Presumably most heuristics exist because they have adaptive value.[9] They may be faster, easier, less demanding of information, or capable of producing an answer when nothing else is available. Gerd Gigerenzer and those he has inspired have demonstrated many cases where simple heuristics outperform more elaborate ways of making judgments (Gigerenzer et al., 1999). For example, in one famous set of experiments, German subjects were as good as Americans at saying which of two American cities was larger. The Germans knew very much less, but they could do well at the task by picking the single city that they recognized, when they recognized only one (Goldstein & Gigerenzer, 2002; Hoffrage, 2011).

5.1.3 Two-systems theory

The idea of heuristics meshes with a kind of psychological theory that has been around for centuries, based on the idea that our cognitive machinery has two ways of doing things: System 1 is fast, automatic, and perhaps impulsive; System 2 is slow and reflective (e.g., Evans, 2008; Kahneman, 2011). The distinction is perhaps best illustrated in trick problems, such as those used by

---

[8]The isolation effect is essentially the same as "focusing" (Legrenzi, Girotto, & Johnson-Laird, 1993) and "choice bracketing" (Read, Loewenstein & Rabin, 1999).

[9]Such functionalist or teleological explanations are not the only ones that can be given. Nor do they imply that adaptive value works through biological evolution. It can also work through cultural evolution, individual learning, or individual invention.

Frederick (2005). For example, "If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?" System 1 says "100 minutes". Some people stop there. But System 2 can correct this error. Close to the idea of System-1 is the role of emotion, drives, and motives, such as sex, anger, fear, and joy. Although it is tempting to think that System 1 is the source of biases and System 2 is the source of rational thinking, the evidence for this view is weak. Biases and errors may well arise from difficult-to-use rules, which are applied by System 2 (insofar as the distinction is clear).

One of the questions for contemporary JDM research is when people rely on System 1 vs. System 2 (see, e.g., Milkman, Chugh, and Bazerman, 2009). People are apparently more likely to rely on System 2 reasoning when they believe they will be asked to give reasons for their choices (Tetlock 1985), when they make choices about events in the far rather than the near future (Milkman, Rogers & Bazerman, 2008), and when they are making judgments in a foreign language (Keysar, Hayakawa & An, 2012).

5.1.4   Impulsiveness, hot cognition, and self-control

Economics thinks of preferences in a very straightforward way: if you did it or chose it, then you preferred to do it or choose it. This is the theory of revealed preferences. What makes the notion of preferences complicated from a psychological standpoint is that people make all kinds of impulsive or unfortunate decisions that they would never have planned to make and that they regret almost immediately. For example, a smoker might, all things considered, prefer not to be a smoker. But when the next nicotine craving hits, he smokes a cigarette. The economic view of this behavior is that whatever his professed preference, his choice to smoke reveals that he preferred to smoke that cigarette. From a psychological view, it is not clear that we ought to identify the decision to smoke as the real preference and ignore the agent's explicit goals or subsequent regret. When a choice is made often determines how it is made, in ways that present a challenge for straightforward ideas about preference and self-control.

Impulsiveness describes the value of now, the disproportionate weighing of outcomes in the current period, such that an outcome with a smaller but sooner benefit trumps one with a later, bigger benefit (Lynch & Zauberman, 2006). Smoking now has a small immediate reward; quitting smoking has a big long-term health benefit. The idea of hot vs. cold cognitive states (Metcalfe & Mischel, 1999) is related to impulsiveness. A hot state is one in which the outcome is immediate and highly salient or vivid. Decisions made with the prospect of immediate gratification are made in a hotter state than those that ask people to consider future rewards, but a hot state is not just about time preference. It also describes the individual's visceral attraction to the reward, which is made greater by states of deprivation—hunger, thirst, exhaustion, and drug withdrawal cause people to reason very differently about the value of food, water, sleep, and opioids, respectively (Loewen-

stein, 1996). Of particular importance to any discussion of legal action is the phenomemon called the "hot/cold empathy gap" (Loewnstein, 1996). The hot/cold empathy gap describes the consistent failure of individuals in a cold state to predict how they will behave in a hot state. This has implications for a wide spectrum of individual decision-making, from the everyday (hungry shoppers choose more junk food than sated shoppers) to the criminally offensive (men in an aroused state report greater willingness to use deception or force with a partner than men in a non-aroused state).

Behavioral law and economic studies of impulsiveness and hot cognition have focused on legal approaches to reducing the probability or harm of impulsive choices. Sin taxes and warning labels are common policy approaches for reducing poor consumption choices, but the existing studies suggest that these approaches will fail, because decision-makers in a hot state are quite insensitive to the costs, financial or otherwise, of immediate gratification (Lynch & Zauberman, 2006). On the other hand, cooling-off periods have gotten quite a bit of traction with behavioral law and economics scholars (Sunstein & Jolls, 2005). Cooling-off laws give consumers some amount of time after a purchase to correct for decisions that now cause regret.

5.1.5   Outrage, retribution, and punishment

A utilitarian view of punishment is largely concerned with deterrence (and also incapacitation and rehabilitation, when these are relevant), but for most individuals the motive for punishment is retribution. The psychological study of anger, disgust, and punishment has been highly influential for legal scholars, particularly for criminal law. In some respects, this is an odd area for behavioral law and economics, insofar as it is an odd area for law and economics, too. Although Becker (1968) and Posner (1973) have written highly influential articles on optimal deterrence, many if not most legal scholars advocate an essentially deontological theory of state punishment. In other words, the most influential normative theory of punishment, at least in the criminal law context, expressly disavows utilitarianism. Psychological research suggests that such theories are supported by the intuitive judgments of many people (Robinson & Darley, 2000).

Anger and outrage.   The psychology of anger begins with the observation that anger is a physiologically intense emotion; that is, it is characterized by a high level of arousal. Cognitive appraisal (e.g., this person intends to do me harm) and physiological response (heart racing, body temperature rising) have dynamic effects—when you think you've been wronged, you experience anger arousal, but the physiological experience of anger may also drive attributions of blame. Anger is an emotion that motivates action, and often the action it motivates is retaliation. One way for an angry person to reduce the aversive aroused state is to retaliate; for example, counter-aggression noticeably reduces blood pressure (Baker & Schaie, 1969).

Feelings of anger or outrage may push people to punish less carefully. Punishment, of any reasonable target, becomes the goal, at the expense of accuracy (Goldberg, Lerner, & Tetlock, 1999). Angry people are more likely to attribute harm to a person rather than a situation (Keltner, Locke & Audrain, 1993), more likely to infer culpability from a set of ambiguous facts (Goldberg et al., 1999), and more likely to rely on superficial cues or stereotypic judgments (Bodenhausen, 1994).

Retaliation. When people are mad, they retaliate. Psychological research on anger suggests that anger is a highly aversive state that people really want to stop experiencing. One of the most effective ways to dissipate angry feelings is to act on them; this is retaliation. In a classic Ultimatum game (Thaler, 1988), one of two subjects is randomly chosen to receive $10. This subject can offer some amount to his partner, the Receiver, who can accept or refuse. If the Receiver refuses, the money goes back to the experimenter. Economic theory based on self-interest alone implies that the Receiver should accept any amount greater than 0. Yet many Receivers turn away stingy non-zero offers, leaving both players with nothing. One way of explaining this finding is that the Receiver sees the stingy offer, gets mad, and fights back with the only tool available. Indeed, this explanation is supported by Xiao and Houser (2011) who find that if Receivers are permitted to communicate their anger to selfish Senders, the Receivers will then take the low offers.

There are multiple theories, some complementary to each other, to explain the psychology of anger and revenge. Evolutionary psychologists argue that revenge and retaliation are adaptive, insofar as one who exacts revenge for slights against him will deter others from taking advantage (Tooby, Cosmides, & Price, 2006), and indeed there is evidence that the threat of retaliation inhibits aggression (Baron, 1973). Psychologists who ascribe to a theory of inequity aversion argue that "getting even" really is about restoring psychological equity (Berscheid, Boye, & Walster, 1968).

Citizens can also impose punishments on one another without a legal judgment. A party to a contract who feels taken advantage of might breach knowing that expectation damages will not fully compensate the non-breaching party for the hassle, much less the insult, caused by breach (Wilkinson-Ryan, 2011). Disputants in a tort suit might refuse to keep negotiating if they think the other party is being clearly unfair—even if settlement is cheaper than litigation (Babcock & Loewenstein, 1996). Divorcing spouses may force litigation as a means of punishing a violation of the marriage contract (Wilkinson-Ryan & Baron, 2006).

Third-party punishment. Third-party punishment is a response to a first person's moral violation, against a second-person victim, by a third person who did not suffer direct harm from the violation. From a behavioral point of view, the interesting point is that people will impose punishments on perceived wrongdoers, even when the punisher has no obligation to punish and when the

punishment is costly to the punisher (Fehr & Gachter, 2002).

Arguably, third-party punishment by a state is central to law itself. In the absence of something like a government, retaliation by individuals or groups (such as kin) is common, and the invention of a legal system is seen as a reform, preventing cycles of violence and replacing impulsive action with something like due process (Diamond, 2008). However, governments, once they exist, try to prevent both second-party punishment (retaliation, "taking the law into your own hands") and third-party punishment ("vigilante justice"). Subjects in economic experiments on third-party punishment are thus in an awkward position, because of social norms against third party-punishment outside of the law. The experiments may depend on making it clear that the law has no role.

Aside from criminal law, citizens are asked to levy punishments on wrongdoers when they decide punitive damages in tort actions. From the point of view of potential parties to litigation, the ability to predict the magnitude of a damages award bears both on whether or not to take certain precautions, and, ultimately, whether or not to settle a claim (Polinsky & Shavell, 1997).

Translating a sense of moral wrongfulness or outrage into an award of money is a particularly complex decision, and it leads to unusually erratic judgments (Kahneman, Schkade, & Sunstein, 1998). The outrage model suggests that subjects arrive easily at an appropriate severity of punishment, and thus have a shared "punitive intent." But assigning a dollar value to the punitive intent becomes more difficult, and jurors are apt to take into account a variety of factors, including the size of the offending firm and the extent of harm suffered by the plaintiff.

Pointless punishment and neglect of deterrence   Baron and Ritov (1993) have described a phenomenon they call "pointless punishment", wherein people are willing to impose punitive damages on tortfeasors irrespective of the consequences of the punishment for the social good—that is, people punished misdeeds, even in scenarios in which the punishment itself would yield net harms to society. In that case, such punishment was retributive, but many subjects, in other scenarios, were sensitive to whether or not the punishment could deter future harm.

The questionnaire used by Baron and Ritov asked respondents to imagine that the United States had a legal system (much like New Zealand's), in which separate decisions were made about penalties against injurers and compensation for victims (Baron & Ritov, 1993). The cases, one involving a death caused by a vaccine and the other involving sterility caused by a birth-control pill, were designed so that they would probably not meet a negligence standard for liability. Subjects (including retired judges) were asked about appropriate penalties under two conditions: a finding of liability would lead the producer to make a safer product; or it would lead the producer to remove the product from the market, which would mean that only a less safe alternative would be available. (Arguably, this is what happened in the U.S. in the 1970s and 80s, with respect to some vaccines.)

Some subjects thought that the penalty should be less, or zero, when its deterrent effect was perverse, that is, leading to worse outcomes. Most subjects thought that the penalty should be the same. Baron and Ritov (1993) asked subjects if they had heard of the deterrence principle as a justification of punishment and, once it was briefly explained, whether they agreed with it. Some subjects had not heard of it: Of these, about equal numbers accepted and rejected it once it was explained. Other subjects had heard of it: Of these, some rejected it, and some accepted it. Those who rejected the principle thought that retribution was a better way to apply the law. In sum, many people who apply nonutilitarian principles do not know what they are rejecting. A brief explanation of the utilitarian approach will be persuasive to some of these people, but not to others.

Of interest in this study were the individual differences in whether deterrent effects were considered, and in willingness to consider them. We might conclude that "on the average, people levy pointless punishments," but that obscures the fact that the average does not represent everyone and that people are somewhat malleable.

5.2   Other-regarding preferences

Economic models of human decision-making often assume rational, self-interested agents. In this section we take up the latter part of that formulation for a moment, the idea that people's preferences are driven by self-interest. The assumption that people are self-interested is often an assumption of convenience, if only because there are individual differences and some people really are only self-interested, and the law must take them into account. But there may be situations in which this assumption does more harm than good, where making law for the "knaves" crowds out intrinsic virtuous motivation (Frey, 1997).

5.2.1   Prosocial behavior and fairness

There is ample evidence that most people have other-regarding preferences. Some examples of this are very intuitive; it is quite natural to express a preference for one's children to be happy and successful, even when there is no clear material benefit to oneself. It is more difficult to explain typical results from the classic dictator game, in which a participant in an experiment is given $10 and told to allocate some or none of it to an anonymous second player. Instead of keeping all of the money, most players give away a non-trivial sum, even if the experimenter cannot identify who made which allocation, and even if the players will never know one another's identities (Forsythe, Horowitz, Savin, and Sefton, 1994; Engel, 2010).

Unlike economics, psychology starts with the premise that empathy and prosocial behavior are adaptive and, indeed, integral to normal psychological development. The psychological study of other-regarding preferences has long been oriented around the question of when and why people

help one another. Emotional accounts argue that humans have a kind of reflexive aversion to seeing other humans in need (Piliavin, Dovidio, Gaertner, & Clark, 1981), or more generally, that the human capacity to empathize is central to prosocial behavior (Coke, Batson, & McDavis, 1978). There are also norm-based explanations suggesting that people help because they prefer to follow an internalized social norm that favors helping and sharing, in turn avoiding "self-sanctions" like guilt and lowered self-esteem (Schwartz, 1973). Perhaps the most influential description of the helping/sharing norm is equity theory (Walster, Walster, & Berscheid, 1978) and inequity aversion (Fehr & Schmidt, 1999).

The idea that people are not entirely self-interested, that they may prefer to exchange some marginal welfare of their own in order to increase the welfare of others, has been very influential for private law contexts like negotiation and contract. For example, one of the core predictions of economists for legal behavior is that a party to a contract will breach the contract when breach is cheaper than performance (including the cost of paying expectation damages). Recent work suggests that many people would not breach for small to moderate increases in profit, in large part because they are sensitive to the expectations of the promisee and prefer not to disappoint them (Wilkinson-Ryan, 2010). Loewenstein, Thompson and Bazerman (1989) used negotiation experiments to track disputants' preferences for outcomes to self and other, and found that many subjects were willing to forego gains to themselves in favor of a more equitable allocation. In short, the identification of other-regarding preferences in experimental and field settings offers evidence about when negotiators, parties to a contract, or even disputants in a settlement might choose to forgo profit in favor of equity.

5.2.2 Moralistic and protected values

Just as people may have preferences for the material well-being of others, there is also strong evidence that they have "moralistic" principles for others. That is, some people would prefer that others follow a particular set of moral rules even when following those rules does not have good consequences for those affected or for the adherents of the rules in question, outside of the moralistic value itself. Baron (2003) found endorsements of moralistic goals across a number of policy domains, often without regard to the consequences of the behavior and the preferences of the other. The idea of moralistic values helps explain policy debates in a variety of arenas, including same-sex marriage, contraception and abortion, environmental regulation, pornography, and drug use. Holders of moralistic values often attempt to justify them in terms of consequences for someone, but such justifications often seem too weak to justify the strength of the value itself

Many of these moralistic values are also protected from trade-offs with other values (Baron & Spranca, 1997; Baron & Ritov, 2009; see also Fiske & Tetlock, 1997, for a similar idea, which they called "sacred values"). People will say, for example, that abortion is always wrong no matter

how great the benefits that would result from allowing a single abortion (and regardless of whether a majority of people think otherwise). Roth (2007) has argued that values of this sort, with popular support, find their way into laws and rules, such as prohibitions on selling organs. He argues that a feeling of repugnance lies behind such prohibitions, which he labels as "repugnant transactions."

5.3 Naïve theories

Unfairly neglected in recent literature is another approach, which holds that people have naïve theories. The idea comes from the study of child development. Children think that the sun goes around the earth, that the earth is flat. But naïve theories are also found in adults. Some people think that a thermostat works like an accelerator, so that, if you turn it way up, the room will heat up faster (Kempton, 1986). This approach is somewhat like the idea of biases, as compared to normative models. But the "true theories" to which judgments are compared are not seen so much as definitions of optimality that are good for all time but rather as just the most defensible current thinking. People undoubtedly have naïve theories of economics, of how the law works, and of their roles as citizens (Baron, 2012).

One area in which naïve theories are both abundant and consequential is consumer finance. Recent debates over regulation of financial markets have cited a number of behavioral anomalies that are arguably caused by naïve theories, especially in the high-stakes context of retirement planning (see, e.g., Fisch, 2010; Coates & Hubbard 2008). Bernartzi and Thaler (2001) documented the phenomenon of "naïve diversification", showing that most people think that portfolio diversification is essentially an investment of 1/n into each of n funds offered—they believe that spreading out investments is good, irrespective of the attributes of the options at hand. Similarly, many investors rely on an explicit theory that favors funds with strong past performance and ignores fund fees, which are deemed too small to matter (Wilcox, 2003), though in fact almost most retail investors are better off ignoring past performance and choosing the lowest-fee funds available.

5.4 Error

Most of the emphasis in the JDM literature is on biases, which are systematic errors, but unsystematic errors also matter in the law. One example is variability in awards of punitive damage (Sunstein, Kahneman & Schkade, 2005). Whatever the right answer about what these awards should be, juries are extremely variable, and the same happens with simulated juries. As in the case of capital punishment, questions arise about whether variable penalties are inherently unfair.

From a normative perspective, the prospective offender should consider the expected utility of the offense. For sufficient deterrence, the expected utility of the penalty should be greater—but

not too much greater—than the expected utility of the gain from the offense.[10] Thus, it should not matter whether the penalty is variable; all that matters is the expected utility of all the penalties. But it is possible that the variability is skewed. Errors in punitive damages are limited on the low size by 0, but not limited on the high side, so that the mean of these awards is higher than the median. But the median may come closer to what is correct. If so, the variability itself creates excessive awards.

One might argue that variability is unfair even if the mean were correct. The same might be said for the variability that results from taking probability of apprehension into account when assessing penalties (one of the arguments in favor of assessing punitive damages at all). From a utilitarian point of view, however, variability does not necessarily reduce total utility, and the alternative would be worse. If we treated all offenses as if apprehension rates were 100%, under-deterrence would be massive, and unfairness would still result from the fact that some offenders would get away with no penalty at all.

5.5 Formal models

JDM scholars sometimes apply formal, mathematical models to explain judgments and decisions. EU and probability theory are models of this sort, as are other models used in economics. But the emphasis in JDM is on models that explain biases.

5.5.1 Prospect Theory and related theories

Kahneman and Tversky (1979) proposed a modification of EU that accounted for most of the known biases, which they called Prospect Theory (PT). The EU of a prospect is $\sum_i p_i u_i$, over the possible outcomes i. PT replaces p_i with a transformation $\pi(p_i)$, where the π function exaggerates small differences in probability near 0 and 1 and pushes other probabilities toward .5 (or something a little lower), as if people tended to think of probability of an event in terms of three categories: "won't happen", "might happen", and "will happen". The effect is to make people overattentive to small differences in very low probabilities (and very high ones).

PT then replaces u_i with v_i, the "value function". Economists usually think of the utility of money in terms of wealth. But v_i is applied not to wealth but to gains and losses from a reference point, usually the status-quo. In EU, we usually assume that wealth has declining marginal utility. The value of an additional \$1,000 is greater to an average person than to a billionaire. (The billionaire thus suffers less pain from paying more in taxes, up to a point.) But, according to PT, when

---

[10]Over-deterrence has at least three disadvantages: it causes harm to offenders that is not justified by the harm prevented; it is costly; and it deters offenses that arguably should not be deterred, such as breaking into a cabin in the wilderness in order to avoid freezing to death.

people make decisions involving gains and losses from the status-quo, they think differently. The value function v has two features. One is that "losses loom larger than gains". This is of course perfectly reasonable; EU says the same thing about wealth. But for PT, the distinction between gains and losses is psychological. It depends on what you think of as the reference point.[11]

One result is that possession can determine the reference point and increase the apparent value of an endowment. In a classic study of this "endowment effect", some Cornell students were told that they owned a Cornell mug and were asked how much money they would require to give it up. Another group was told that they could have a choice of money or the mug, and they were asked how much money they would require to take the money rather than the mug. The only difference between the groups was thus what they were told about ownership. The "owners" demanded about twice as much money to give up (forego) the mug (Kahneman, Knetsch & Thaler, 1990).

Although PT attributes this effect to loss aversion, that is, to the shape of v, it is also possible that it results from a simple heuristic against change, i.e., a bias toward the status-quo. Loss aversion itself has been studied in other ways and turns out to be somewhat labile, depending on how it is measured (e.g., McGraw, Larsen, Kahneman & Schkade, 2010; Ert & Erev, 2013). Note that questions about a results explanation are not the same as questions about its reality.

The second feature of v is that differences near the reference point have more of an effect than differences farther away, regardless of whether they are gains or losses. This illustrates a general psychological principle of *diminishing sensitivity*, which also applies to the π function, particularly the decreased sensitivity to probability differences that are far from 0 or 1. Again, EU would normally imply diminishing sensitivity for gains, but not for losses. Because of the declining marginal utility of wealth, EU says the opposite for losses. The difference between losing $1,000 and losing nothing is less serious than the difference between $10,000 and $11,000 according to EU, but more serious according to PT. This is why people can be risk averse when gains are involved but risk seeking for losses. Faced with a choice of losing $100 for sure or a 50/50 gamble of losing $200 or nothing, people may choose the gamble because the difference between $-\$100$ and 0 seems larger than that between $-\$200$ and $-\$100$.

A third feature is that the reference point is malleable; it can change as a result of how a gamble is described or "framed". In the classic "Asian disease problem" (Tversky & Kahneman, 1981), subjects were told that an unusual Asian disease is expected to kill 600 people. Half the subjects were told that there were two programs: one would save 200 people, and the other had a .33 chance of saving 600. The other half of the subjects were told that with the first program 400 would die, and with the second there was a .67 chance that 600 would die. The first way of presenting the

---

[11]Tversky and Kahneman (1992) presented a second version of PT, which takes into account the ranking of possible outcomes, called cumulative prospect theory. For this exposition, we treat the two forms as equivalent, because they make the same major assumptions, but some of their predictions differ, with results favoring the later theory.

problem implied a reference point of the 600 lost, so the outcomes were gains. The second way used the absence of the disease as the reference point, so outcomes were losses. Subjects were risk aversion in the gain form, tending to choose the 200 lives, but risk seeking in the loss form. The two forms are, of course, equivalent. This result illustrates the effect of framing.

PT has been used to explain a number of choices and biases in the legal context. Zamir and Ritov (2010) used a set of experiments to show that loss aversion affects the preference for contingent-fee arrangements in lawyer-client relationships. Experimental "plaintiffs" preferred a contingent fee even if the expected fee was much higher in such a situation than in a non-contingent fee arrangement. Defendants, on the other hand, preferred fixed fees—because defendants choose between two purely negative gambles, and do not have the option to avoid a loss altogether. In contract, researchers have argued, and shown experimentally, that the contract serves as a reference point around which parties frame their expectations (Hart & Moore, 2008; Fehr, Hart and Zehnder, 2011). Parties are willing to accept certain costs built into the contract but respond negatively when they must pay the same costs at a counterparty's discretion under a flexible contract. In other words, the moment of contract resets the status quo and provides a frame for evaluating each particular outcome as a gain or a loss.

Many of the psychological insights of PT, and the findings that support it, have survived subsequent research: the malleability of reference points and the effects of framing; different risk attitudes in gains and losses; diminishing sensitivity; and some examples of loss aversion. However, PT may be misleading in its account of how people combine information about outcome probability and utility. In particular, PT (along many other theories) ascribes risk aversion/seeking and loss aversion to the shape of the utility (or value) function. An alternative type of model ascribes these effects to the weights rather than the utilities. The weights, psychologically, refer to the amount of attention paid to each possible consequence. Risk aversion and loss aversion result from more attention/weight to the worse possible outcomes. One such model is that transfer-of-attention exchange (TAX) model of Birnbaum (Birnbaum, 2008; Birnbaum, Johnson, & Longbottom, 2008), which accounts for choices among gambles in terms of differential attention to better and worse outcomes. Several results are inconsistent with PT and related models but consistent with the TAX model, e.g., PT does not explain some violations of stochastic dominance.[12]

As Birnbaum (2008) points out, we must be careful to distinguish phenomena from their theoretical accounts. Risk aversion is a phenomenon; concave utility is an explanation of it, but not the only explanation. Of course, true utility of money surely is concave — an additional $1,000 matters more to a poor person than to a rich one — but responses to gambles need not occur for this reason. (And the declining marginal utility of wealth cannot plausibly account for the amount

---

[12]Stochastic dominance says that if, for gamble G and all values of X, the probability of winning X or more is at least as great as the probability of winning X or more in (non-identical) gamble F, then G should be preferred over F.

of risk aversion found in laboratory experiments with small amounts of money.)

5.5.2 Psychophysics and psychophysical numbing

The idea of "declining marginal disutility" of losses illustrates a general psychological principle of *diminishing sensitivity*. Differences seem smaller when they are farther from a salient reference point. When we think about disasters that affect large numbers of people, such as wars, famines, hurricanes, and earthquakes, the numbers affected cease to matter so much as they get larger. This "psychophysical numbing" leads us to focus disproportionately on single victims and attend to little to the great mass (Slovic, 2007).

5.5.3 Hyperbolic discounting

Diminishing sensitivity may account for over-attention to the near future, as opposed to the distant future (Lynch & Zauberman, 2006). A delay of one day from now means more to us than the difference between 7 and 8 days in the future, yet, when the 7th day arrives, our perspective will change. Whatever the cause, people's discounting of future events falls off sharply when the events are close in time, roughly in the shape of a hyperbola. Yet economic theory, for good reason, says it should fall of with exponential decay. The phenomenon of hyperbolic discounting may help to explain the sorts of impulsiveness we described in section 5.1.4.

Anomalies in preferences over time are of particular relevance to legal decision-makers. Regulation of insurance, consumer financial products, retirement savings, and even contracting in general is often faced with evidence that individuals underweight the value of future events. Moreover, the story is not even that straightforward—recent evidence suggests intertemporal discounting is not symmetric for past and future. In a paper that brings hyperbolic discounting to prospect theory, Caruso, Gilbert, and Wilson (2008) show that people value future events more than equivalent events in the equidistant past. They demonstrate their theory with jury decision-making experiments. Subjects were asked, for example, to set damages for a person who had been injured and either had gone through a painful recovery or would undergo a painful recovery—and they awarded more when the recovery was imagined in the future rather than the past. For ordinary citizens making legal decisions and judgments, the temporal relationships of the relevant events may affect decisions even when timing should be irrelevant.

5.6 Myside bias, polarization, and cultural cognition

Another set of biases concern the irrational persistence of belief (Baron, 2008, ch. 9).[13] *Selective exposure* is the tendency to seek evidence that will support beliefs that are already strong. *Biased assimilation* is the tendency to downgrade contradictory evidence when it arrives, while taking supporting evidence at face value. These biases, together with the tendency not to look for alternative possibilities, are together called myside bias. When people on opposite sides of some issue are subject to these biases, the result is polarization, those on each side become more extreme. Mere presentation of evidence does not help, because of biased assimilation.

Kahan et al. (2011) have recently argued that myside bias is stronger when "cultural" factors support certain beliefs. For example, those opposed to government regulation because of their social environment are more likely to deny the role of human beings in causing global warming. Thus, people are motivated to hold beliefs consistent with their overall cultural perspectives. More generally, we might suppose that bias is stronger when a belief is more central to a person's identity (in the sense of Erikson, 1968), which is usually but not always related to a person's cultural milieu.

Excessive reading of the psychological literature on myside biases is a risk factor for pessimism. Before we sink into despair, we must remember a couple of things. First, the research is mostly about average effects. On the average, people have these biases. But studies of individual differences show extreme variability, to the point where we can be confident that many people do not display these biases at all (e.g., Stanovich & West, 1998). The reason they show up in average effects is simply that very few people, if any, show the opposite biases, such as over-weighing evidence against one's current beliefs. The average of zero and some positive effects is positive.

Second, myside bias is correlated with people's beliefs about what good thinking is. Those who are most subject to these biases also believe that good thinkers are loyal to their beliefs, for example (Baron, 1995). Such findings suggest that myside bias can be corrected by challenging people's naïve theories about what good thinking is. People might be taught to respect thinking that is "actively open-minded" (Baron, 2008).

6 Prescriptive approaches

The sort of debiasing just mentioned is one kind of prescriptive approach, that is, an answer to the question of what to do about biases and errors. The behavioral law-and-economics literature has suggested others.

---

[13]Much has been written about how "crazy" (or unwarranted) beliefs persist, but little about how they are first formed. People who believe that the attacks of 9/11 were a conspiracy by the U.S. government, or Israel, have persisted in this belief despite counter-arguments, but most people never formed such beliefs in the first place. Why not?

6.1 Decision analysis, cost-benefit analysis, and decision aids

One approach is to provide tools to help people analyze decisions formally. Governments have taken various steps to avoid reliance on human intuitive judgments. Central banks have largely removed the setting of interest rates from the hands of politicians. Several regulatory agencies of many governments use cost-benefit analysis, to varying degrees, to decide on regulations. Breyer (1993) and (in a more moderate form) Sunstein (2002) have advocated extensions of this techno-cratic approach. It could, for example, reduce the problem of over-regulating some things while under-regulating others.

Other tools for decision analysis could be used for individual decisions. Some decision aids of this sort are used in the design of guidelines for diagnosis and treatment in medicine, and also for individual decisions. In principle they could also be used to aid courts, especially when technical data are involved.

A relevant example from health care is the decision about fetal testing for Down syndrome (e.g., Ganiats, 1996). Some years ago, the major test for this was amniocentisis, which involves withdrawing fluid from the uterus and examining the cells. The test was not perfect, and as a side effect it sometimes caused miscarriages. But the probability of a Down syndrome birth increased with the mother's age. A widely adopted guideline for the test was that women over 35 should have it. The origin of this guideline was that, at this age, the probability of a test-caused miscarriage was equal to the probability of a Down-syndrome birth. The assumption was that a positive test result would lead to an abortion. If we apply EU theory to this decision, such a guideline would be appropriate if the disutility of a miscarriage were equal to that of a Down-syndrome birth. But that is rarely the case. Even crude assessment of individual utilities for these two outcomes would imply that some women should never have amniocentesis (even if they were willing to have abortions if the test were positive) and others should have it at much younger ages. This example shows how the use of simple decision analysis, in terms of expected utility, could yield better decisions for individuals. In this case, the incorrect decisions did not result from any particular known biases, but they were not optimal. (Now there are many more fetal tests available, and the decision is much more complicated. Individualized decision analysis would still help, but it would require a computer.)

A similar example is the allocation of funds to different investments in retirement programs. Many employers now offer default plans that are designed to be optimal for most employees, but employees with defined-contribution plans can usually choose from various options that vary in risk and expected outcome (e.g., stocks vs. bonds). Some of the companies that provide these investments attempt to measure something like the risk attitude of their customers, which utility theory would say is a function of the degree of curvature of the utility function for money. Such measurements may be less successful than those that might be used for amniocentesis (Baron,

2011).

6.2   Libertarian paternalism and the "decision architect"

Thaler and Sunstein (2008) proposed another prescriptive approach, which is to think of the design of decisions as they are presented to people, which they call decision architecture. Sometimes this involves simply changing the way in which options are presented, such as putting recommended options first on a list. It can also involve changing the options themselves. But the idea is to leave people free to choose while, at the same time, trying to shield them from harmful effects (on others as well as themselves) of decision biases. Some changes in design require legislation or regulatory decisions. Examples are requiring that employers that offer pension plans make diversified plans the default, while giving employees the option to switch to a riskier plan. People are biased toward the default, so this idea makes use of a bias to help people choose options thought by others to be better. Other examples of decision architecture include mandatory waiting (cooling off) periods (which counteract the tendency to make impulsive decisions), mandatory provision of options to un-do bad decisions, and mandatory provision of information. In some cases, such devices could replace what amount to legal prohibitions (e.g., against giving up the right to sue, or assisting suicide), thus increasing the choice available.

Recent research on nudges has focused extensively on the use of default rules (e.g., Sunstein, 2011, 2013). Perhaps that is the result of the fact that other sorts of nudges, e.g., sin taxes, cooling-off periods, and mandatory disclosure of certain kinds of information, have been known and used for some time, while the intentional use of defaults that people are free to override as a means of "light paternalism" is new. The idea of a bias toward the default should be distinguished from two other closely related biases: the status-quo bias and omission bias (Baron, 2008). Default bias does not require a status-quo. An example is that new motorists can be assigned to insurance policies with or without a limited tort provision. When this is done, they strongly favor whatever they are assigned (Johnson, Hershey, Meszaros, & Kunreuther, 1993). Omission bias is a bias toward inaction, but the term is limited to cases in which both action and omission cause harm. When both options result in improvements, actions are often favored over omissions, but this effect is generally small compared to the effect for harms, leading to an overall bias toward the default (Baron & Ritov, 1994).

Aside from the auto-insurance example, another example of an apparently successful nudge concerns organ donation. Countries where being an organ donor the default (especially those with hospitals that are willing to enforce the rules) have much higher donation rates than those with an opt-in rule (Johnson & Goldstein, 2003). Another good example is enrollment in pension plans and employer-provided health insurance by default, with the possibility of opting out (Sunstein,

2011).

Thaler and Sunstein argue that their proposals mostly "nudge" people in the direction away from biases and errors, while leaving people free to make bad choices, or, ideally, choices that are better for them than the default. They thus call their approach "libertarian paternalism". But economists might argue that they are sometimes increasing the cost of some options, if only the cost of waiting. Hence the line between coercion and liberty is not sharp (Baron, 2010).

6.3   Asymmetric paternalism and libertarian welfarism

Camerer, Issacharoff, Loewenstein, O'Donoghue, and Rabin (2003) suggest a different analysis that leads to most of the same conclusions. Many of the proposals favored by Sunstein and Thaler can be seen as preventing large harms to a few at the expense of very minor costs to many, hence the idea of "asymmetric paternalism." For example, compulsory waiting periods for marriage, divorce, or assisted suicide can prevent a small number of huge mistakes, for a relatively small cost to many. If the cost ratio is sufficiently extreme, measures of this sort have a utilitarian rationale.

Korobkin (2009) has extended the idea in another direction (although fully compatible with the idea of asymmetry), which is to consider social benefits as well as individual benefits.[14] He calls this libertarian welfarism. A prime example is organ donation (as noted earlier), where we can take advantage of bias toward the default in order to increase donation rates through presumed consent (Johnson & Goldstein, 2003).

7   Conclusion

The simple elegance of classical economic theory has been disrupted by findings from psychology and related fields. Note that the disruption is not in the normative application of economic theory but first in the descriptive part, and, as a result in new prescriptive ideas that do not arise from economics itself. To some it seems unfortunate that the findings of psychology seem more like a list of heuristics, biases, and models than any sort of cohesive theory like that of micro-economics. This is especially so now that the list of biases has increased several fold over the much smaller list available when legal scholars started taking this work seriously. We have tried here to point out some common themes in the research, which tie together various bunches of findings.

A general concern that often arises is that many of the prescriptions that come out of these new findings, such as those of libertarian paternalism, seem like tricks designed to manipulate irrational actors. Economic theory, by contrast, is all about incentives and is designed for rational

---

[14]Sunstein and Thaler include many examples in which the benefits are social, but they do not make a point of discussing the general idea.

actors. Irrational actors may not stay irrational, especially when they are repeatedly manipulated by the same tricks. An implication of this view is that we should spend more effort trying to make people more rational and not jump to the conclusion that irrationality is always with us. Indeed, large individual differences in most biases indicate that it is not universal.

Yet, the forces that produce irrational biases—still not fully understood—are likely to remain. Civilization and its army of educators will be waging a constant battle against these forces, so they will probably to continue to exist. Thus, the design of the law and its application may need to consider how it can deal with irrationality for the foreseeable future.

References

Adler, Matthew D. (2012). *Well-being and fair distribution: A framework for policy analysis*. New York: Oxford University Press.

Babcock, Linda, & Loewenstein, George (1997). Explaining bargaining impasse: The role of self-serving biases. *Journal of Economic Perspectives, 11*, 109–126.

Baker, John W., and K. Warner Schaie. 1969. "Effects of Aggressing 'Alone' or 'With Another' on Physiological and Psychological Arousal." *Journal of Personality and Social Psychology* 12: 80–86.

Baron, Jonathan. (1996). Do no harm. In D. M. Messick & A. E. Tenbrunsel (Eds.), *Codes of conduct: Behavioral research into business ethics*, pp. 197–213. New York: Russell Sage Foundation.

Baron, Jonathan. (2003). Value analysis of political behavior — self-interested : moralistic :: altruistic : moral, *University of Pennsylvania Law Review, 151*, 1135–1167.

Baron, Jonathan. (2008). *Thinkind and deciding (4th ed.)*. New York: Cambridge University Press.

Baron, Jonathan. (2010). Review of *Nudge: Improving decisions about health, wealth, and happiness*, by R. H. Thaler and C. R. Sunstein. *Journal of Behavioral Decision Making, 23*, 224-226.

Baron, Jonathan. (2011). Risk attitude, investments, and the taste for luxuries vs. necessities. *Frontiers in Cognition*. Special issue "Human preferences and risky choices" (Eds.: P. Van Schaik, P. Kusev, & A. Juliusson). *2*, 329. http://www.frontiersin.org/cognition/10.3389/fpsyg.2011.00329/full

Baron, Jonathan. (2012). The "culture of honor" in citizens' concepts of their duty as voters. *Rationality and Society, 24*, 37–72.

Baron, Jonathan, & Ilana Ritov. (1993). Intuitions about Penalties and Compensation in the Context of Tort Law. *Journal of Risk and Uncertainty, 7*, 17–33.

Baron, Jonathan. & Ritov, Ilana (1994). Reference points and omission bias. *Organizational Behavior and Human Decision Processes, 59*, 475–498.

Baron, Jonathan, & Ritov, I. (2009). Protected values and omission bias as deontological judgments. In D. M. Bartels, C. W. Bauman, L. J. Skitka, & D. L. Medin (Eds.), *Moral Judgment and decision making*, Vol. 50 in B. H. Ross (series editor), *The Psychology of Learning and Motivation*, pp. 133–167. San Diego, CA: Academic Press.

Baron, Jonathan, & Spranca, Mark (1997). Protected values. *Organizational Behavior and Human Decision Processes, 70*, 1–16.

Baron, Jonathan, & Szymanska, Ewa (2010). Heuristics and biases in charity. In D. Oppenheimer & C. Olivola (Eds). *The science of giving: Experimental approaches to the study of charity*, pp. 215–236. New York: Taylor and Francis.

Baron, Robert A. (1973). Threatened Retaliation from the Victim as an Inhibitor of Physical Aggression. *Journal of Research in Personality, 7*, 103–115.

Becker, Gary S. (1968). Crime and Punishment: An Economic Approach. *Journal of Political Economy76*, 169–217.

Benartzi, Shlomo, and Richard H. Thaler. 2001. Naïve diversification strategies in retirement saving plans. *American Economic Review* 91: 79–98.

Bentham, Jeremy (1948). *An introduction to the principles of morals and legislation.* Oxford: Blackwell Publisher. (Original work published 1843)

Berscheid, Ellen, David Boye, & Elaine Walster. (1968). Retaliation as a Means of Restoring Equity. *Journal of Personality and Social Psychology10*, 370–376.

Bicchieri, Cristina (2006). *The grammar of society: The nature and dynamics of social norms.* New York: Cambridge University Press.

Birnbaum, Michael H. (2008). New paradoxes of risky decision making. *Psychological Review, 115*, 463–501.

Birnbaum, Michael H., Kathleen Johnson, & Jay-Lee Longbottom (2008). Tests of Cumulative Prospect Theory with graphical displays of probability. *Judgment and Decision Making, 3*, 528–546.

Bodenhausen, Galen V., Lori A. Sheppard, & Geoffrey P. Kramer. (1994). Negative Affect and Social Perception: The Differential Impact of Anger and Sadness. *European Journal of Social Psychology24*, 45–62.

Breyer, S. (1993). *Breaking the vicious circle: Toward effective risk regulation.* Cambridge, MA: Harvard University Press.

Camerer, Colin F., Samuel Issacharoff, George Loewenstein, Ted O'Donoghue, & Matthew Rabin. (2003). Regulation for conservatives: Behavioral economics and the case for "asymmetric paternalism." *University of Pennsylvania Law Review, 151*, 1211–1254.

Caruso, Eugene M., D. T. Gilbert, & Timothy D. Wilson. (2008). A Wrinkle in Time: Asymmetric Valuation of Past and Future Events. *Psychological Science19*, 796–801.

Coates, J. C., & R. Glenn Hubbard. (2007). Competition and Shareholder Fees in the Mutual Fund Industry: Evidence and Implications for Policy. *Journal of Corporation Law33*, 151–222.

Coke, Jay S., C. D. Batson, & Katherine McDavis. (1978). Empathic Mediation of Helping: A Two-Stage Model. *Journal of Personality and Social Psychology36*, 752–766.

Dawes, Robyn M., & Thaler, Richard H. (1988). Cooperation. *Journal of Economic Perspectives, 2*, 187–197.

Dhami, Mandeep K., & Peter Ayton (2001). Bailing and jailing the fast and frugal way. *Journal of Behavioral Decision Making, 14*, 141–168.

Diamond, Jared (2008). Vengeance is ours. Annals of anthropology. *New Yorker*, April 21, pp. 74 ff.

Dunwoody, Philip T. (2009). Theories of truth as criteria in judgment and decision making. *Judgment and Decision Making, 4*, 116–125.

Engel, Christoph. (2011). Dictator games: A meta study. *Experimental Economics14*, 583–610.

Erikson, Erik H. (1968). *Identity, youth and crisis*. New York: Norton.

Ert, Eyal, & Erev, Ido. (2013). On the descriptive value of loss aversion in decisions under risk: Six clarifications. *Judgment and Decision Making, 8*, 214–235.

Evans, Jonathan St. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology, 59*, 255–278.

Fehr, Ernst, Oliver Hart, & Christian Zehnder. (2011). Contracts as Reference Points-Experimental Evidence. *American Economic Review, 101*, 493–525.

Fehr, Ernst, & Klaus M. Schmidt. (1999). A Theory of Fairness, Competition, and Cooperation, *The Quarterly Journal of Economics, 114*, 817–868.

Fiske, Alan P., & Philip E. Tetlock. (1997). Taboo trade-offs: Reactions to transactions that transgress the spheres of justice. *Political Psychology, 18*, 255–297.

Fisman, Raymond, & Edward Miguel. (2007). Corruption, norms, and legal enforcement: Evidence from diplomatic parking tickets. 115 *Journal of Political Economy*, 1020–1048.

Forsythe, Robert, Joel L. Horowitz, N. Eugene Savin, & Martin Sefton. (1994). Fairness in Simple Bargaining Experiments, *Games and Economic Behavior, 6*, 347–369.

Fox, Craig R., & Amos Tversky (1998). A belief-based account of decision under uncertainty. *Management Science, 44*, 879–895.

Frederick, Shane (2005). Cognitive reflection and decision making. 19 *Journal of Economic Perspectives*, 24–42.

Frey, Bruno S. (1997). A constitution for knaves crowds out civic virtues. *Economic Journal, 107*, 1043–53.

Ganiats, T. G. (1996). Justifying prenatal screening and genetic amniocentesis programs by cost-effectiveness analysis. *Medical Decision Making, 16*, 45–50.

Gigerenzer, Gerd, Peter M. Todd, & the ABC Research Group (1999). *Simple heuristics that make us smart*. Oxford: Oxford University Press.

Goldberg, Julie H., Jennifer S. Lerner, & Phillip E. Tetlock. (1999). Rage and Reason: The Psychology of the Intuitive Prosecutor. *European Journal of Social Psychology, 29*, 781–795.

Goldstein, Daniel G., & Gerd Gigerenzer (2002). Models of ecological rationality: The recognition heuristic. 109 *Psychological Review*, 75–90.

Hammond, Kenneth R. (1996). *Human judgment and social policy: Irreducible uncertainty, inevitable error, unavailable injustice.* New York: Oxford University Press.

Hart, Oliver, & John Moore. (2008). Contracts as Reference Points. *Quarterly Journal of Economics, 123*, 1–48.

Hoffrage, Ulrich (2011). Recognition judgments and the performance of the recognition heuristic depend on the size of the reference class. 6 *Judgment and Decision Making*, 43–57.

Johnson, Eric J., & Goldstein, Daniel G. (2003). Do defaults save lives? *Science, 302*, 1338–1339.

Johnson, E. J., Hershey, John C., Meszaros, Jacqueline, & Kunreuther, Howard (1993). Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty, 7*, 35–51.

Jolls, Christine, & C. R. Sunstein. (2006). The Law of Implicit Bias. *California Law Review, 94*, 969–996.

Jolls, Christine, C. R. Sunstein and R. Thaler (1998). A behavioral approach to law and economics. 50 *Stanford Law Review,*, 1471–1550.

Kahan, Dan M., Maggie Wittlin, Ellen Peters, Paul Slovic, Lisa Larrimore Ouellette, Donald Braman, Gregory N. and Mandel. (2011). The tragedy of the risk-perception commons: Culture conflict, rationality conflict, and climate change. Temple University Legal Studies Research Paper No. 2011-26; Cultural Cognition Project Working Paper No. 89; Yale Law & Economics Research Paper No. 435; Yale Law School, Public Law Working Paper No. 230. Available at SSRN: http://ssrn.com/abstract=1871503 or http://dx.doi.org/10.2139/ssrn.1871503

Kahneman, Daniel (2011). *Thinking, fast and slow.* New York: Farrar, Strauss and Giroux.

Kahneman, Daniel, & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment*, pp. 49–81. New York. Cambridge University Press.

Kahneman, Daniel, Jack L. Knetsch, & R. H. Thaler (1990). Experimental tests of the endowment effect and the Coase theorem. *Journal of Political Economy, 98*, 1325–1348.

Kahneman, Daniel, & Daniel Lovallo (1993). Timid choices and bold forecasts—a cognitive perspective on risk-taking. *Management Science 39*,17-31.

Kahneman, Daniel, David Schkade, & C. R. Sunstein. (1998). Shared Outrage and Erratic Awards: The Psychology of Punitive Damages. *Journal of Risk and Uncertainty, 16*, 49–86.

Kahneman, Daniel, & Amos Tversky (1972). Subjective probability: A judgment of representa-

tiveness. *Cognitive Psychology, 3*, 430–454.

Kahneman, Daniel, & Amos Tversky (1973). On the psychology of prediction. *Psychological Review, 80*, 237–251.

Kahneman, Daniel, & Amos Tversky (1973). On the psychology of prediction. *Psychological Review, 80*, 237–251.

Kahneman, Daniel, & Amos Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263–291.

Keltner, Dacher, Kenneth D. Locke, & Paul C. Audrain. (1993). The Influence of Attributions on the Relevance of Negative Emotions to Personal Satisfaction. *Personality and Social Psychology Bulletin, 19*, 21–29.

Kempton, W. (1986). Two theories of home heat control. *Cognition, 10*, 75–90.

Keysar, Boaz, Hayakawa, S. L., & An, S. (2012). The foreign-language effect: Thinking in a foreign tongue reduces decision biases. *Psychological Science, 23*, 661–668.

Korobkin, Russell. (2009). Libertarian welfarism. *California Law Review 97*, 1651–1685.

Krantz, David H., R. Duncan Luce, Patrick Suppes, & Amos Tversky, (1971). *Foundations of measurement* (Vol. 1). New York: Academic Press.

Lawson, Glen, Jonathan Baron, & Linda S. Siegel. (1974). The role of length and number cues in children's quantitative judgments. *Child Development, 45*, 731–736.

Legrenzi, Paulo, Vittorio Girotto, & Philip N. Johnson-Laird.(1993). Focusing in reasoning and decision making. *Cognition, 49*, 36–66.

Loewenstein, George F. (1996). Out of Control: Visceral Influences on Behavior, *Organizational Behavior and Human Decision Processes, 65*, 272–292.

Loewenstein, George F., Leigh Thompson, & Max H. Bazerman. (1989). Social Utility and Decision Making in Interpersonal Contexts. *Journal of Personality and Social Psychology, 57*, 426–441.

Luce, R. Duncan, & Howard Raiffa (1957). *Games and decisions: Introduction and critical survey*. New York: Wiley.

Lynch, J. G., & Gal Zauberman. (2006). When Do You Want It? Time, Decisions, and Public Policy. *Journal of Public Policy and Marketing* 25: 67–78.

McCaffery, Edward J., & Baron, Jonathan. (2006). Isolation effects and the neglect of indirect effects of fiscal policies. *Journal of Behavioral Decision Making, 19*, 1–14.

McDaniels, Tim L. (1988). Comparing expressed and revealed preferences for risk reduction: Different hazards and question frames. *Risk Analysis, 8*, 593–604.

McGraw, A. Peter, Larsen, J. T., Kahneman, Daniel, & Schkade, D. (2010). Comparing gains and losses. *Psychological Science, 21*, 1438–1445.

Metcalfe, Janet, & Walter Mischel. (1999). A hot/cool system analysis of delay of gratification:

Dynamics of willpower. *Psychological Review, 106*, 3–19.

Milkman, Katherine L., Chugh, Dolly, & Bazerman, Max H. (2009). How Can Decision Making Be Improved? *Perspectives on Psychological Science, 4*, 379–383.

Milkman, Katherine L., Todd Rogers, & Max H. Bazerman. (2008). **Harnessing our inner angels and demons: What we have learned about want/should conflicts and how that knowledge can help us reduce short-sighted decision making, Perspectives on *Psychological Science, 3* 324–338.

Piliavin, Jane A., J. F. Dovidio, Samuel L. Gaertner, & Russel D. Clark. (1981). Emergency Intervention. New York: Academic Press.

Polinsky, Mitchell, & Steven Shavell, (1997). On the Disutility and Discounting of Imprisonment and the Theory of Deterrence, NBER Working Papers 6259.

Polya, George (1945). *How to solve it: A new aspect of mathematical method*. Princeton: Princeton University Press.

Posner, Richard A. (1973). *Economic Analysis of Law*. Boston: Little Brown.

Raiffa, Howard (1982). *The art and science of negotiation.* Cambridge, Mass.: Harvard University Press.

Read, Daniel, George L. Loewenstein and Matthew Rabin (1999). Choice bracketing. *Journal of Risk and Uncertainty, 19*, 171–197.

Ritov, Ilana, & Baron, Jonathan. (1990). Reluctance to vaccinate: omission bias and ambiguity. *Journal of Behavioral Decision Making, 3*, 263–277.

Robinson, Paul H., & Darley, John M. (1997). The Utility of Desert. *Northwestern University Law Review* 91, 453–499.

Robinson, P. H., Darley, J. M. (2000). Testing lay intuitions of justice: How and why? 28 *Hofstra Law Review*, 611–634.

Robinson, P. H., & Darley, J. M. (2004). Does criminal law deter? A behavioral science investigation, 24 *Oxford Journal of Legal Studies*, 173–205. Roth, Alvin E. (2007). Repugnance as a constraint on markets. 21 *Journal of Economic Perspectives*, 37–58.

Roth, Alvin E. (2008). "What have we learned from market design?" Hahn Lecture, Economic Journal, 118 (March), 285–310.

Savage, Leonard J. (1954). *The foundations of statistics*. New York: Wiley.

Schwartz, Shalom H. (1973). Normative Explanations of Helping Behavior: A Critique, Proposal, and Empirical Test. *Journal of Experimental Social Psychology, 9*, 349–364.

Shavell, Steven (2004) *Foundations of economic analysis of law*. Cambridge, MA: Belknap Press of Harvard University Press.

Slovic, Paul (2007). "If I look at the mass I will never act": Psychic numbing and genocide. 2 *Judgment and Decision Making*, 79–95.

Spranca, Mark, Elisa Minsk, & Baron, Jonathan. (1991). Omission and commission in judgment and choice. *Journal of Experimental Social Psychology, 27*, 76–105.

Stanovich, Keith E., & Richard F. West (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General, 127*, 161–188.

Sunstein, Cass R. (2002). *Risk and reason: Safety, law, and the environment.* New York: Cambridge University Press.

Sunstein, Cass R. (2005). Moral heuristics (with commentary). 28 *Behavioral and Brain Sciences*, 531–573.

Sunstein, Cass R. (2011). Empirically informed regulation. *University of Chicago Law Review, 78*, 1349–1429.

Sunstein, Cass R. (2013). Deciding by default. *University of Pennsylvania Law Review, 162*, 1–57.

Sunstein, Cass R., Kahneman, Daniel, & Schkade, David (2005). Assessing punitive damages (with notes on cognition and valuation in law). *Yale Law Journal, 107*, 2071–2053.

Tetlock, Phillip E. (1985). Accountability: A social check on the fundamental attribution error. *Social Psychology Quarterly, 48*, 227–236.

Thaler, Richard H. (1988). The ultimatum game. *Journal of Economic Perspectives, 2*, 195–206.

Thaler, Richard H., & Cass R. Sunstein (2008). *Nudge: Improving decisions about health, wealth, and happiness.* New Haven, CT: Yale University Press.

Tooby, J., Leda Cosmides, & Michael E. Price. (2006). Cognitive Adaptations for n-person Exchange: The Evolutionary Roots of Organizational Behavior. *Managerial and Decision Economics, 27*, 103–129.

Tversky, Amos, & Daniel Kahneman (1992). Advances in prospect theory: cumulative representations of uncertainty. *Journal of Risk and Uncertainty, 5*, 297–323.

Walster, Elaine, G. William Walster, & Ellen Berscheid. (1978). *Equity: Theory and Research.* Boston: Allyn and Bacon.

Wilcox, Ronald T. (2003). Bargain Hunting or Star Gazing? Investors' Preferences for Stock Mutual Funds. *Journal of Business, 76*, 645–664.

Wilkinson-Ryan, Tess (2011). Breaching the Mortgage Contract: The Behavioral Economics of Strategic Default. *Vanderbilt Law Review, 64*, 1547–1583.

Wilkinson-Ryan, Tess, & Baron, Jonathan. (2009). Moral Judgment and Moral Heuristics in Breach of Contract. *Journal of Empirical Legal Studies, 6*, 405–423.

Xiao, Erte and Daniel Houser 2011 "Punish in Public", Journal of Public Economics.95, 1006-1017

Zamir, Eyal, & Ritov, Ilana, (2010). Revisiting the Debate over Attorneys' Contingent Fees: A Behavioral Analysis. *Journal of Legal Studies, 38*, 246–288.