The Jurisprudence of Experimental Law and Economics

Comment
by
DOUGLAS A. KYSAR

1 Introduction

With astounding speed and comprehensiveness, experimental approaches to law and economics have infiltrated the legal academy. Slightly more than a decade ago, one could find only a handful of studies that attempted to bring the insights and methodologies of experimental economics and cognitive and social psychology to law (see, e.g., KOROBKIN AND GUTHRIE [1994], RACHLINSKI [1994], KAMIN AND RACHLINSKI [1995], RACHLINSKI [1996], KOROBKIN [1998], RACHLINSKI [1998]). Today, on the other hand, one finds an array of experimental studies on legal topics being conducted at universities in North America, Europe, Israel, and elsewhere. New scholars are being encouraged into the field not only indirectly through the field’s salience and popularity, but also more directly and tangibly through the establishment of scholarship competitions and conferences aimed specifically at junior scholars. A new peer-reviewed journal, the Journal of Empirical Legal Studies, has become a leading outlet for empirical approaches to law, including experimental ones. This special issue of the Journal of Theoretical and Institutional Economics, co-edited by Christoph Engel of the Max Planck Institute for Research on Collective Goods and Urs Schweizer of the University of Bonn, further demonstrates the transnational appeal of experimental law and economics. The net result has been an explosion of interest in empirical approaches to law in general, and experimental approaches in particular.¹

Among the most significant work in this exciting new field has been a series of experimental studies on judicial officers jointly conducted by Chris Guthrie of Vanderbilt Law School, Jeffrey Rachlinski of Cornell Law School, and the Honorable Andrew Wistrich of the United States Federal District Court for the Central District of California (see, e.g., GUTHRIE, RACHLINSKI, AND WISTRICH [2001]; GUTHRIE, RACHLINSKI, AND WISTRICH [2007]). These studies have demonstrated that judges

¹ Reviews of the legal literature can be found in GUTHRIE [2003], RACHLINSKI [2003], and LANGEVOORT [1998]. Reviews of the behavioral economics and psychological literatures from which the legal studies draw can be found in CAMERER AND LOEWENSTEIN [2004] and GILOVICH AND GRIFFIN [2002], respectively.
employ cognitive heuristics in much the same manner as laypersons and that, at times, their use of such mental shortcuts – also like laypersons – leads judges to engage in biased decision making. These studies have been especially significant for two reasons. First, the researchers have been ingenious in their adaptation of existing cognitive and social psychological survey designs to the legal context – a feat that requires the authors to have both a sophisticated grasp of psychology and psychological methods, and a learned understanding of the subtleties and complexities of law. Second, and perhaps even more significantly, the researchers have gained access to several large groups of actual sitting judges and magistrates as subjects for their studies. The researchers have thus pulled off something of a social science coup – they have not only secured a rare and significant observational opportunity, they have taken good advantage of it.

The present study, “Heuristics and Biases in Specialized Judges: The Case of Bankruptcy Judges” (RACHLINSKI, GUTHRIE, AND WISTRICH [2007]), is no exception. In this study, the authors gained access to 113 sitting United States bankruptcy judges, a group that represented a remarkable 36.1% of the entire federal bankruptcy judiciary. This was an especially valuable research opportunity not only because of the unusual level of access to sitting judicial officers of any stripe, but also because the bankruptcy judiciary in the United States is widely viewed as an exceptional category of judges, one for which merit-based selection criteria largely trump more political appointment considerations, thereby resulting in a truly expert body of judicial officers. Thus, the group of judges recruited by Guthrie, Rachlinski, and Wistrich offered an opportunity both to study further the role of cognitive heuristics and biases in judicial decision making, as their previous research had done, and also to evaluate the performance of a specialized group of judges. As the authors point out, in other contexts, experts have a somewhat mixed record as to whether they develop better heuristics than lay people (pp. 168f.). Whether bankruptcy judges would deploy heuristic processes in the same manner, and with the same detrimental impact on judgment, as non-specialized judges therefore represented an important, well-framed research question. And again, with customary skill and insight, the researchers used the opportunity to great effect.

This comment offers a few questions and observations regarding the survey’s findings and non-findings, followed by a more general question regarding the future of experimental law and economics. As will be seen, the comment concludes on a note of friendly challenge and provocation, directed not only to Guthrie, Rachlinski, and Wistrich, but to all researchers engaged in experimental economic and psychological

2 Despite this extremely large pool of federal bankruptcy judges, the sample nevertheless does appear anomalous in that the participating judges self-identified as 76.6% Democrat and only 23.4% Republican (RACHLINSKI, GUTHRIE, AND WISTRICH [2007, p. 170]). Given that some of the authors’ most interesting findings have to do with political party affiliation, it would be helpful for the authors to provide more contextual information to explain the seemingly disproportionate share of Democrats in their sample.
approaches to law. This field of study has undergone rapid evolutionary advances over the past decade, progressing from largely speculative, descriptive analyses of how various findings from psychology and experimental economics might impact law, to specially adapted surveys of law student “convenience samples,” to entirely novel research investigations of jurors, judges, and other actual legal actors. While applauding this progress—and the researchers who have made it happen—this comment also questions whether the next evolutionary step in the field can occur within the confines of the social science paradigm, or whether instead the nature of law ultimately will force those who study it through experimental approaches to embrace a more pragmatic, jurisprudential self-conception of their enterprise.

2 Evidence of Biased Decision Making

Rachlinski, Guthrie, and Wistrich’s bankruptcy subjects exhibited evidence of biased decision making from two cognitive heuristics: anchoring and framing (RACHLINSKI, GUTHRIE, AND WISTRICH [2007, p. 181]). In both instances, however, the interpretation of the researchers’ results is made at least somewhat ambiguous by complicating factors in the survey design.

For instance, the anchoring survey question involved not one but two potential anchors, both of which—unlike more typical tests of anchoring such as the spinning of a numerical wheel or the recitation of one’s phone number—were logically relevant to the decision task at hand. Specifically, all subjects were instructed that, under controlling Supreme Court precedent, they should begin with the prime rate of interest (4.5%) when setting the interest rate for a restructured loan and adjust upward as appropriate to reflect the debtor’s risk of default. Half of the subjects also were given information regarding the interest rate specified in the original loan document (21%), although they were told the parties agreed that this original rate was “irrelevant” to the judicial determination of a new rate. As the authors predicted, the bankruptcy judges in the second condition specified a higher rate of interest (7.13%) than the judges in the first condition (6.33%) (p. 172). Thus, the “irrelevant” information of the original interest rate appeared to affect the judges’ decision making.

An earlier study by the authors provided a much cleaner test for anchoring, one that involved only a single anchor—a $75,000 minimum amount in controversy jurisdictional requirement—that was clearly irrelevant to the task of setting an appropriate award of damages for a catastrophic injury (GUTHRIE, RACHLINSKI, AND WISTRICH [2001, pp. 790f.]). In the present study, the higher interest rate—the original contract rate—only seemed to be irrelevant to the task at hand by virtue of the parties’ supposed agreement that it was. To be sure, the Supreme Court decision does prescribe a method for calculating the new interest rate that eschews the original contract rate as an express anchor, but it would go too far to say that the original rate is logically irrelevant to the evaluation of the debtor’s default risk. Although the protection of bankruptcy does change the debtor’s risk
profile considerably, it does not render irrelevant the implicit risk rating that was provided by the original contract rate of interest. Thus, rather than a bias toward an irrelevant anchor, evidence of the influence of the original contract rate information on the judges’ decision making may simply reflect the rational use by those judges of information that was available to them, but not to the judges in the first condition. It may, in other words, reflect a tacit form of objection to the Supreme Court’s decision in *Till v. SCS Credit*, which had claimed that the original interest rate was not useful information to the bankruptcy judge and which had rejected the proposal of creditors that restructured loan interest rates be anchored and adjusted downward from the original contract rate. Admitting this possibility of nullificatory behavior by judges in the second condition complicates the attempt to identify evidence of cognitive bias. What appears to be evidence of irrational anchoring on irrelevant information may actually be a more subtle and normatively complicated form of judicial nonacquiescence.

The authors’ framing question was also quite complicated and may have been imperfectly constructed to capture the cognitive processes that the authors intended. In both conditions, judges were provided with a choice between bankruptcy reorganization plans, one of which provided a certain result and one of which provided an uncertain result that was probabilistically equivalent to the certain result (RACHLINSKI, GUTHRIE, AND WISTRICH [2007, pp. 173ff.]). By varying whether the outcomes were expressed as gains or losses, the researchers sought to uncover evidence of loss aversion – that is, a greater willingness to incur risk in order to avoid outcomes presented as losses than to obtain outcomes presented as gains, even when the outcomes are otherwise substantively identical. And, indeed, the authors did find that judges were less likely to prefer the certain, more conservative reorganization plan when it was described as offering a certain level of nonrepayment (or loss), as opposed to repayment (or gain) (p. 174).

As conceded by the authors, however, the survey instrument was amenable to an alternative interpretation in which the expected outcomes were not economically equivalent (pp. 174ff.). That is, because Plan A in both conditions only described what would happen to a certain portion of the unsecured debt, judges were free to infer that the remainder of the debt in Plan A might also be subject to some chance of repayment (in the gain condition) or non-repayment (in the loss condition). Indeed, using the Principle of Insufficient Reason with respect to the unspecified portion of the debt, one could easily infer that Plan A in the gain condition was worth $400,000 rather than only $200,000 (a sure $200,000 plus a fifty percent chance of repayment for the remaining $400,000), and that in the loss condition it was worth only $100,000 rather than $200,000 (a sure non-repayment of $400,000 plus a fifty percent chance of non-repayment for the remaining $200,000). Thus, it is not necessarily the case that “[b]oth versions of the choice are, in fact, economically identical” (p. 174). One could – consistent with rational choice assumptions – prefer Plan A to Plan B in the gain condition more than one prefers it in the loss condition.

In order to determine which explanation – the framing effect or the rational choice explanation – better accounts for subjects’ actual approach to the survey, one would
need a cadre of psychologists, linguists, and other experts, as well as an ample research budget. Raising the possibility of a rational choice explanation is therefore not intended to rule out the authors’ framing explanation, which indeed on reflection seems to offer the most compelling account of the observed behavior. It is instead offered simply in order to highlight the magnitude of the survey design challenge that researchers face when studying judicial decision making, a challenge that will be revisited at the conclusion of this comment.

3 Evidence of Non-Biased Decision Making

The authors found no evidence of biased decision making with respect to four questions that were designed to trigger, respectively, the omission bias, the effect of race on decision making, the effect of apologies, and the effect of terror management strategies (Rachlinski, Guthrie, and Wistrich [2007, p. 181]). Although this comment largely agrees with the authors that these various non-findings represent “good news” (p. 184), there are several questions and complications that deserve brief mention.

To begin with, the implications of the authors’ non-finding with respect to the omission bias are less clear than the remaining non-findings. In a variety of psychological experiments, subjects have been observed to qualitatively distinguish between, on the one hand, consequences that flow from human actions and, on the other hand, otherwise identical consequences that follow human omissions (that is, opportunities to act that are foregone). Generally speaking, consequences that flow from actions seem more worthy of blame and regret to individuals than those that follow omissions. Rachlinski, Guthrie, and Wistrich regard their failure to find such a distinction at work in the bankruptcy judges’ decision making to be desirable, as they had set out to design a question in which the distinction between actions and omissions would be irrelevant. Specifically, they tested bankruptcy judges’ willingness to discharge credit card debt that was used by a debtor to cover basic expenses. The debtor was unable to repay the credit card loans because he had lost savings through an investment in a mid-sized company that suffered a stock plunge following an announcement that the company was being made the subject of a securities fraud investigation (pp. 175f.). In one version of the question, the debtor inherited the stock directly from his brother and considered converting it into U.S. Treasury notes, but decided not to do so. In another condition, the debtor inherited the funds as U.S. Treasury notes and then affirmatively decided to convert them into the stock that eventually became worthless. Thus, in one condition, the debtor failed to avoid the financially disastrous investment, while in the second condition the debtor affirmatively acted to incur the disaster.

This study raises several interesting issues. First, a general – and testy – point about the omission bias: There is no such thing as an omission bias. Only strict impartial consequentialist-utilitarians such as the psychologist Jonathan Baron re-
gard these effects as indicative of bias. For others, including what is probably the majority of moral philosophers, longstanding conventions such as the act-omission distinction are an essential, indeed ineliminable aspect of moral reasoning. As Samuel Scheffler has shown, they are a foundational, meta-ethical feature of the very idea of moral agency, the very idea that human behavior has a normativity at all, in the sense that it has appeal to identifiable subjects of moral reasoning and responsibility (SCHEFFLER [2004]). Impartial consequentialist–utilitarians such as Baron instead force the individual to treat her actions as fundamentally mundane – that is, as being indistinguishable from all other elements of the complex causal order, including the indefinitely many “omissions” that are unavoidably created at every temporal instant of her life’s narrative. It is difficult, within such a homogenized conception of the causal order, to imagine how an individual would maintain a strong sense that her actions matter, that she is in fact a distinctive moral being whose identity choices are worthy of special attention, as opposed to simply a part of the furniture of impartial causal optimization (KYSAR [2007]).

Of course, however fascinating this meta-ethical discussion might be, why should it matter to the present paper? First, with respect framing, one might expect to see less pronounced results in the bankruptcy judge study than in the classic Asian Disease problem, as the authors did, both because the bankruptcy subjects were not evaluating core moral categories such as life and death, and also because the judges were acting as judges – that is, as specifically role-differentiated decisionmakers, in which the expression of their agency reflected, not purely or even primarily their personal moral identity, but rather the system of rules and conventions within which the judge qua judge operates. In such a case, one might expect less manifestation of the deontological, agent-relative, “not through me” brand of moral reasoning, of which the act-omission distinction is the best example, but which also is related to the kind of loss aversion demonstrated by the Asian Disease problem. Individuals may be more willing to incur risk in the domain of losses in that classic problem in part because they have been confronted with decision tasks that unavoidably implicate their moral agency, rather than an institution’s, and that do so with respect to the primary moral categories of life and death. “First, do no harm,” goes the familiar adage.

Second, with regard to the omission bias question, the survey instrument designed by Rachlinski, Guthrie, and Wistrich may not have been squarely aimed at an act-omission distinction. The authors attempted to avoid including facts in the hypothetical that would go to the state of mind of the investor, because they wanted to create a problem in which the nature of the behavior, whether action or omission, would have had no probative value with respect to culpability or fraud. Generally speaking, actions are thought to represent stronger probative evidence of intent or state of mind than omissions, because they are more clearly deliberately chosen and less subject to alternative explanations. The researchers therefore needed to exclude the possibility that the debtor was engaged in fraud or otherwise culpable conduct when deciding whether or not to invest (or maintain an investment) in the mid-sized
company. If fraud were a possibility, judges might rationally determine that the debtor in the action condition was less deserving of discharge than the debtor in the omission condition. This complication perhaps explains why the authors described U.S. treasury bonds as potentially offering “a better return” rather than “a safer investment vehicle,” which would have been the more natural way to describe a comparison of government bonds and private equity investments. Of course, if the investor discounted expected return for the possibility of a securities fraud investigation or other financial disaster, then the two descriptions would amount to the same thing. But, if the investor were thinking in those terms, then the judge would have had reason to believe that the investor was less innocent than the materials portray, in which case the question would no longer have offered a clean cut at an act-omission bias.

To the extent that the researchers tried to eliminate these state of mind issues – which they did by setting up a fact scenario that should not reasonably have been thought of as fraudulent by any survey respondent – their question came to resemble standard regret avoidance studies, in which researchers test the ease with which unfortunate outcomes and experiences can be cognitively “undone” as a means of evaluating the incidence and severity of regret attributed to a bad outcome. How readily, in other words, can subjects conjure a counterfactual scenario in which a misfortune that has been suffered is instead avoided? The answer to that question influences significantly the level of regret than individuals ascribe to an unfortunate situation. Regret avoidance is related to the act/omission distinction because actions generally are easier to cognitively “undo” than omissions. However, there does not appear to be an obvious normative gloss that can be put on the phenomenon. If, for instance, the researchers had asked the judges to predict what level of regret the debtor felt at his situation, and the judges in the active investment condition stated a higher level of regret than those in the passive condition, would that finding have been considered evidence of bias, or just affirmation of the manner in which individuals seek to impose narrative continuity onto the epiphenomenal collection of choices and actions that they call, by convention, their lives? In short, a basic level of confusion persists with respect to the omission bias question: What exactly did the researchers expect to find with this question and, if they had found it, would they have been justified in calling it a “bias”? Like the omission bias, Rachlinski, Guthrie, and Wistrich also found no effect for race, apologies, and terror management – three additional possible sources of undesirable emotion influence (RACHLINSKI, GUTHRIE, AND WISTRICH [2007, p. 181]) that have been found to be powerfully active in other decision making contexts. The lack of influence of racially identifying information on the discharge of student loan debt represents a desirable – and unfortunately surprising – non-finding. Because undue influence of race has been observed in other aspects of judicial decision making (p. 176), further research is merited here in order to glean what features of Rachlinski, Guthrie, and Wistrich’s decision task or sample pool led to the desirable non-finding.
Similarly, the failure to observe an effect of apology on the judges’ decision making is desirable in the context of a decision in which apology should have no substantive weight. Further research on the role of apology in judicial decision making should follow this important study and should exploit the kinds of factors identified by the researchers as likely to enhance or subdue its influence (pp. 179f.). In addition, it would be valuable to examine why gender, political party, and years of experience on the bench failed to have an effect with respect to the apology and terror management question, given that the ultimate decision task for that question was in many respects similar to the question on student loan debt, where such effects were observed. One of the authors’ most significant findings was that the bankruptcy judiciary may be more politically-oriented in its decision making than popular wisdom suggests (p. 183). It is worth investigating further, then, why the role of political party affiliation seems only to have appeared in response to certain decision tasks, despite the structural similarity of those tasks to other questions in the study.

Finally, as the authors note in the paper (pp. 178f.), it is hard to predict what form a terror management effect would have taken, even if it were observed. An essential aspect of the judicial officer’s role is often to recognize and reconcile multiple, conflicting deeply held values. Thus, to the extent that invocations of mortality trigger the attempt to manage terror through reaffirmation of fundamental values, it is unclear which fundamental values a judicial officer would reaffirm, those of her own or those of the broader governmental structure within which she is situated. If the former, a question arises regarding whether and to what extent such personally affective decision influence is appropriate for a judicial officer. If the latter, a further question arises regarding which, from among multiple and often competing fundamental values, would predominate. These complications go to a final, more general, and perhaps more significant question about the future direction of law and psychology, behavioral law and economics, and other experimental approaches to law.

4 “Bias” in the Context of Legal Decision Making

The early research program of Daniel Kahneman and Amos Tversky was ingenious and revolutionary because it used simple survey instruments to reveal insights about very general, but previously underappreciated or unrecognized aspects of human cognition (KAHNEMAN AND TVERSKY [1974]). The program did so primarily by comparing, against rational choice ideals, the observed behavior of laboratory subjects in response to decision problems that were sufficiently narrow and controlled to permit only a single inference about what cognitive processes were driving observed behavior. Importantly, the rational choice benchmark for comparison was chosen, not because rational choice theory provides a desirable conception of behavior for psychologists (although it does for some psychologists, most prominently Baron (see, e.g., BARON [1998] and [1993]), but rather because
it provides a stable, concrete, and tractable decision making framework against which to compare subject behavior, thereby allowing for the revelation of specific, implicit information about the way in which individuals perceive, evaluate, and decide.

For Kahneman and Tversky, the study of cognitive biases therefore was much like the study of optical illusions: By examining ways in which unobservable processes such as cognition or sight go awry under certain stylized conditions, researchers hoped to infer information about how, more generally, such processes are mentally structured. Importantly, although this methodology has widely been interpreted by economists and legal academics to suggest a normative policy program – one of altering decision environments and otherwise nudging individuals to make decisions that better comport with rational choice ideals (see, e.g., Sunstein and Thaler [2003], Camerer et al. [2003]) – no such program was originally envisioned by the proponents of the research. And although even psychologists now engage in a hearty debate regarding whether individuals in the world make “good” or “bad” decisions through the use of heuristics (Kysar et al. [2006]), no such debate can fruitfully be undertaken without some basic level of agreement among participants regarding what normative benchmark serves to identify “good” and “bad” decision making. Because such a task is not normally thought of as the proper province of social science, it is therefore important to distinguish between positive and normative uses of psychology and experimental economics. Conflating the two may lead both to bad science and to bad policy.

Unfortunately, however, the experimental methodology of Kahneman and Tversky is complicated in the legal context in two significant ways – ways that suggest the analysis of law might necessarily become mired in the normative thicket that positivistic conceptions of social scientific research instead attempt to avoid. First, the decision tasks faced by judges and other legal actors typically resist ready simplification. That is, unlike the kinds of decisions that Kahneman and Tversky sought to evoke through their stylized tasks, legal decision making often involves the balancing of incommensurable factors, the nuanced accommodation of alternative moral frameworks, the unavoidably imperfect reconstruction of historical facts, and, more generally, the messy, pluralistic practices that are associated with pragmatist thought. Thus, in order for an experimental survey instrument to be descriptively rich enough to get at a distinctively legal judgment, it might by necessity become too complicated to target only a single identifiable cognitive process in the manner that Kahneman and Tversky sought to achieve. Put differently, stripping the decision task down to a level at which its results offer unambiguous evidence of cognitive processes may also strip away what makes the decision task a specifically – or an interestingly – legal one.

Second, despite numerous heroic attempts (see, e.g., Kaplow and Shavell [2002]), no one has yet provided a single compelling normative framework for the analysis of law and policy. Moreover, the safest assumption going forward is that no one ever will. Thus, in most judicial contexts, it will be difficult to identify a relevant decision making benchmark akin to rational choice theory to use as a comparison
to “reveal” cognitive processes. As a result, in most cases observed behavior will be amenable to multiple interpretations, some of which may be consistent with the findings of experimental economists and psychologists in more general settings, but others of which will not be. Even the comparison of a simple interest rate-setting judgment against the dictates of a Supreme Court decision turns out to evade a single straightforward interpretation in which subjects are or are not seen to be complying with precedent. Similarly, the evaluation of a debtor’s failure to act – as opposed to a debtor’s action – appears initially to be a simple test of whether bankruptcy judges suffer the influence of the “omission bias.” In fact, however, the exercise turns out to implicate one of the deepest and most puzzling aspects of all of moral philosophy.

In short, the legal decision making environment may be too complex – too polycentric and multivalent – to support the Kahneman and Tversky research methodology, at least with the level of purity and clarity that those researchers were able to achieve with respect to the ultimate interpretation of results. If legal values are necessarily plural, if legal facts are necessarily uncertain, and if legal judgments are necessarily pragmatic, then any attempt to replicate the Kahneman and Tversky research methodology within the context of realistic legal decision tasks may either fail to yield determinate cognitive inferences or fail to capture what is distinctively legal about the assigned tasks.

All of which raises a provocative question: Can the new law and psychology and experimental law and economics movements actually fulfill their aspiration to social science rigor, objectivity, and tractability, or should they instead become something more like a pragmatic jurisprudence, in which legal behavior is not only observed, but also simultaneously described, critiqued, and refined in an explicitly normative pursuit of human flourishing through law? The pragmatist tradition is premised on a curious blend of skepticism and faith in the power of ideas that renders it especially suitable as a paradigm within which to study and evaluate legal decision making. It accepts the inherent limitations of any human knowledge structure, the built-in biases and blind spots that render any particular academic discipline – whether economics, psychology, or otherwise – inherently and irredeemably incomplete. Yet at the same time, pragmatism holds that humanity’s best hope for comprehending and resolving social conflict is precisely to examine it through many diverse theoretical and empirical lenses. So conceived, pragmatism offers a vision of knowledge acquisition and social negotiation that rejects the stark fact/value dichotomy characteristic of Enlightenment science and that instead accepts pluralism and fallibilism as unavoidable features of the human condition.

As argued in this comment, normative presuppositions already are implicitly at work in much of the social scientific study of law. Why not make the normative agenda both more transparent and more catholic by embracing pragmatism? Why not position law and psychology and experimental economics as primary behavioral engines of a new legal pragmatism, one that proves worthy of the name by refusing to privilege any single value framework, such as rational choice theory or welfarism? In light of the kinds of difficulties identified in this comment, such a pragmatic jurisprudential framework may offer a better conceptual and rhetorical fit for the
projects of law and psychology and experimental law and economics than the purely positivistic framework of social science.

Or, at least, we should study the possibility.

References


Douglas A. Kysar
Cornell Law School
Cornell University
Myron Taylor Hall
Ithaca, NY 14853-4901
U.S.A.
E-mail:
douglas-kysar@lawschool.cornell.edu