Behavioral Economics as Part of a Rhetorical Duet: A Response to Jolls, Sunstein, and Thaler

Mark Kelman*

INTRODUCTION

It hardly seems surprising that Professors Jolls, Sunstein, and Thaler ("JST") have written such an interesting and provocative piece, attempting to bring the insights of "behavioral economics" to bear on the usual issues addressed by more conventional law and economics scholars. (Of course, it's plausible that I'm simply exhibiting the sort of "hindsight bias" they discuss, irrationally believing "I knew all along" that they *would* write an interesting piece because in fact they *did*.) Professor Thaler, after all, is surely one of the most thoughtful and imaginative of the leading scholars in behavioral economics. There is always a risk that one will be deemed to slight uncited works by highlighting the particular virtues of some pieces (or slight the many virtues of his coauthors, because they are newer to this field).[1] Still, I think if one wants to get a good feel for the merits of *experimental* behavioral economics, one cannot do much better than to look at Thaler's cowritten work on the endowment effect,[2] and if one is interested in what strikes me as an even more interesting empirical study of real-life, nonlaboratory labor market behavior that appears inexplicable from a conventional rational choice perspective, one simply cannot find a better piece of work than

* William Nelson Cromwell Professor of Law, Stanford University.

1. I need only remind readers of an old joke: The guilt-tripping mother gives her son two shirts for his birthday. He tells her he loves them, expresses all the appropriate gratitude. He comes down the next morning, wearing one of the new shirts. "So what's the matter with the other one?" she asks him.

2. Daniel Kahneman, Jack Knetsch & Richard Thaler, *Experimental Tests of the Endowment Effect and the Coase Theorem*, 98 J. POL. ECON. 1325 (1990).

Thaler's cowritten study of the labor supply functions of New York City cabdrivers.[3]

I have never called myself a "behavioral economist" (or any other kind of economist for that matter) but from the very beginning of my career, I've been concerned with many of the issues that JST identify. My third-year student paper at Harvard Law School[4]—the paper that got me my job here at

3. Colin Camerer, Linda Babcock, George Loewenstein & Richard Thaler, *Labor Supply of New York City Cabdrivers: One Day at a Time*, 112 Q.J. ECON. 407 (1997). While conventional rational choice theory unambiguously predicts that workers, including cabdrivers, would work more hours on days when hourly wage rates are higher—and wage rates change from day to day because nonremunerative cruising time declines when there is bad weather, subway breakdowns, conventions, etc.—cabdrivers in fact drive fewer hours on such days, setting implicit "daily income targets," consistent with the observation of behavioral economists that actors "bracket" decisions narrowly. *See id.* at 408-10.

4. *See* Mark Kelman, Addictive and Status-Oriented Consumption: A Preliminary Critique of Neo-Classical Choice Theory, Anti-Egalitarianism and Economic "Growth" (1976) (unpublished manuscript, on file with the *Stanford Law Review*). A discerning reader can note several things from the title alone. First and foremost, of course, there's more than a bit of what you'd have to call youthful grandiosity if you weren't a close enough friend of mine to go with "exuberance."

Second, I was not really interested simply in coming up with a more accurate description of human behavior than conventional economists had provided, even though much of the paper was an effort to *describe* what I was calling "addictive" and status-oriented consumption. Two of the chapters center on savings behavior and fertility. The basic descriptive claim in each was parallel to an "endowment effect" claim, though I did not use "endowment effect" vocabulary but "goods addiction" language. I used addiction language, trying to emphasize the point that in relationship to ordinary consumption goods, people behave in a fashion we often attribute to heroin addicts: Loss—of the addictive drug dosage, of the "familiar" goods bundle—causes "withdrawal" panic and great pain, though there is far less urgency about seeking "more." I argued in each chapter that people are very anxious when there is some chance that they will lose some of the consumption bundle they are accustomed to but are relatively indifferent to gains above such a baseline. Thus, they cut back discretionary expenditures, e.g., by having fewer children, not when they are poor in some absolute sense but when having more children creates some substantial chance that their consumption of familiar goods would fall below the expectationally set baseline. Similarly, I argued, only partly persuasively I think, that people did not save either to smooth out consumption over the life cycle or when their absolute or relative income was high, but rather they saved primarily "residually"—i.e., saved whatever they simply didn't bother to spend—in those periods when their income exceeded their customary consumption baseline, a baseline that lags in time behind income growth. It is not the consumption of goods as such that gives pleasure, but rather we consume those goods we are used to having because the loss of accustomed bundles causes pain asymmetric to the gains from new goods.

Though there was a great deal of descriptive material in the piece, my goal was pretty explicitly *political*. I viewed rational choice theory as a fundamentally complacent, *normative* theory: a theory that assumed that so long as people made choices within a given series of option sets constrained only by resource availability, they would be as well off as they could be. I was interested, above all, in noting that the choices that people made at times one, two, three, etc., might add up to a "lifetime" choice plan that they would *not* have chosen. I was interested in this point particularly because one of my main substantive political beliefs at the time was that we could ultimately "solve" the "conflict" over resources between rich and poor, both intranationally and internationally, only if and when richer people came to see that they were "approximately" satiated; that is, if they came to see that while at any moment they would always choose more goods over fewer goods, there were reasons to believe that a lifetime of increasing accumulation did not make them subjectively better off. Thus, one of the egalitarian messages was that we would not need typical inegalitarian incentives to produce (e.g., low marginal tax rates) to the extent that the rich were

Stanford—centered, *descriptively*, on two issues that JST highlight early in their account of the gap between traditional and behavioral economics. First, I was interested in what they call "bounded willpower,"[5] though I used somewhat different language (language of regret and unwanted habituation[6]). Second, I was interested in the gap between "opportunity cost" income and "out-of-pocket" income,[7] given my belief that people were very reluctant to lose the opportunity to continue to consume things they were used to consuming but were rather indifferent to gains.[8] I think, though, that JST are right to imply there are interesting distinctions between the approach I (and others they characterize as "[e]arly skeptics about the economic analysis of law"[9]) have taken to behavioral economics and the approach they take. JST's method, I think, is to try to develop a comprehensive theory of human behavior that will blend the insights of traditional rational actor models with the new theoretical refinements of behavioral economics. I will argue that this approach replicates the biggest flaws of the law and economics movement:

(1) It fails to recognize not only the inevitable scientific *partiality* of all models of human behavior, but the degree to which stories about behavior, whether rational actor stories or "richer" behavioral ones, are essentially interpretive tropes rather than full-blown verifiable or falsifiable theories; and

(2) it fails to recognize that behavioral and mainstream economics can both best be used as *approaches* to data that are inexorably ambiguous in their implications.

It would help behavioral economists both to integrate the partial "truth" in rational choice interpretations of their findings and to understand the very dramatic limits on the completeness of their theoretical accounts. In the same fashion, "rational choice" economists should read JST and recognize how much material the rational choice school runs over, roughshod, to

rather indifferent to the precise level of appropriation of goods; the antieconomic growth message was that economic growth created a "hedonic treadmill," increasing the perceived need for goods just as it created the capacity to satisfy them.

5. Christine Jolls, Cass R. Sunstein & Richard Thaler, *Behavioral Responses to Law and Economics*, 50 STAN. L. REV. 1471, 1479 (1998) [hereinafter "JST"]. This aspect of my paper was the explicit topic of the first article that I published once I was teaching here at Stanford. *See* Mark Kelman, *Choice and Utility*, 1979 WIS. L. REV. 769.

6. *See* Kelman, *supra* note 5, at 783-87.

7. *See* JST, *supra* note 5, at 1483-84.

8. This issue was the dominant theme in the second article that I published after coming to Stanford. *See* Mark Kelman, *Consumption Theory, Production Theory, and Ideology in the Coase Theorem*, 52 S. CAL. L. REV. 669 (1979).

9. *See* JST, *supra* note 5, at 1473.

maintain its quite stunning belief that very basic insights from elementary economics have swallowed up the remainder of the social sciences and humanistic interpretivism, both in terms of describing existing social institutions and in terms of prescribing narrow means to implement uncontroversial systemic goals and reach consensus about appropriate goals.

I am not sure I can convince the authors (or my readers) that it might ultimately hurt social scientific inquiry to believe one is constructing (or at least trying to construct) a fuller, more accurate general theory of human behavior rather than simply to recognize the usefulness of drawing on a variety of interpretive traditions in constructing stories about social practices and individual choices. What I will instead try to do, first, is simply to note briefly the ways in which rational choice theory and the critique of it are bound together in a form of rhetorical duet or ritualized dance. In this connection, I will note that maintaining a real awareness of the fact that the models do not really supplement one another, as JST hope, but rather stand in a relationship of irreconcilable mutual dependence, helps us recognize that rational choice theorists will inevitably be perfectly reasonable in believing that many of JST's observations can be interpreted as consistent with their paradigm. I will then discuss, in a bit more detail, the degree to which both economistic and noneconomistic critics of behavioral economists will reasonably respond that even where the behavioral observations are most difficult to square with rational choice theory, they are still more "anecdotal" than they are implications of an alternative "theory." Finally, I will return very briefly to the question of how we might best employ "behavioral" insights.

I. MOVES AND COUNTERMOVES: THE STYLIZED DEBATE BETWEEN RATIONAL CHOICE THEORY AND ITS CRITICS

In a piece that I am cowriting with David Elliot and Hilary Folger centering especially on the limits and virtues of the "hindsight bias" literature,[10]

10. *See* Mark Kelman, David Elliot & Hilary Folger, Hindsight Bias and the Rational Choice Paradigm (1998) (unpublished manuscript, on file with the *Stanford Law Review*). A narrower aspect of our hindsight bias project is described in Mark Kelman, David Elliot & Hilary Folger, Decomposing Hindsight Bias (1998) (unpublished manuscript, on file with the *Stanford Law Review*). In this second piece, my coauthors and I argue that "hindsight bias" should be analytically decomposed into three distinct forms of "bias," which we call "primary," "secondary," and "tertiary": Actors may raise their estimates of the prior probability of an event when given outcome information, they may misestimate what their estimates of prior probability would have been once they know outcome information, and they may judge others unreasonable when those parties did not make the same estimates of prior probability that those possessing outcome information now make. *See id.* at 2. In our view, the first of these is a "bias" rather than a rational use of additional information only to the extent that subjects underestimate the possibility that an outcome occurred as a result of sampling variation. Existing experimental evidence does not establish that they do so,

we note that rational choice theorists have a stylized set of responses to their critics, who themselves can best be understood as offering a few stylized criticisms. Here is our basic view of this rhetorical dialectic:

Rational choice theory may be viewed as that theory in which certain subjects (whom we call "principals") articulate whimsical, subjective, private, individual, and individuating ends that essentially come from within. Bounded by rationality strictures in their choice of particular ends only by relatively weak consistency constraints, these subjects then attempt to attain these ends using instrumentally appropriate means, essentially dictated by a universally perceptible external "reality" that each competent mind should ultimately be able to apprehend.[11]

In relationship to ends formation, cognitive psychologists and behavioral economists, though, believe first that principals violate even weak rationality constraints.[12] Perhaps more interestingly, they share with sociological and anthropological critics of rational choice theory the view that the picture of ends selection posited by rational choice theory is either needlessly thin or affirmatively misleading, insofar as it assumes that ahistorical, acontextual individuals simply "have" exogenous tastes. Classical choice theory assumes each principal assigns each option a unique value and that, given an offered set, the decisionmaker chooses the option of highest value, based on

---

and we provide some experimental evidence that they use outcome information *only* in situations in which it is rational to do so. *See id.* at 4-6.

11. Rational choice theory also posits another group of subjects (we call them "agents") who take their ends from principals who direct them. The agent seeks to follow the principal's commands, making (if rational) externally dictated appropriate particularistic judgments about what the principal would have wanted in the given case.

Typical psychological theories of context-dependence, for instance, question both the possibility that principals have stable (and even weakly rational) internal preferences *and* the idea that either principals or agents will be able to meet any of the goals they seek to attain in an instrumentally appropriate fashion, since they will be "distracted" by context and hence unable to stay focused on their narrowly defined tasks. Thus, as described in note 12 *infra*, a context-dependent decisionmaker's preferences between two goods are affected by the presence or absence of an irrelevant alternative. Similarly, a juror affected by context-dependence will grade crimes or set punishment levels depending on irrelevant facts (whether the given punishment is intermediate or extreme in a list of punishments, whether the given punishment is offered alongside a "similar" punishment to which it is clearly superior). Mark Kelman, Yuval Rottenstreich & Amos Tversky, *Context-Dependence in Legal Decision Making*, 25 J. LEGAL STUD. 287, 290-95 (1996).

12. The most commonplace "weak rationality" critique is that the choices are frame-dependent, so that choosers prefer A to B and B to A simultaneously or nearly simultaneously, depending upon how the choice options are defined or the context in which they are selected. Thus, context-dependent decisionmakers, given the choice between a Mark Cross pen and six dollars may prefer the six dollars to the Mark Cross pen and the Mark Cross pen to six dollars, nearly simultaneously, depending upon the presence or absence of an irrelevant third option, a pen that is markedly inferior to the Mark Cross pen. *See* Itamar Simonson & Amos Tversky, *Choice in Context: Tradeoff Contrast and Extremeness Aversion*, 29 J. MARKETING RES. 281, 287 (1992).

these brute, internal preferences. Every variety of critic of rational choice theory questions that assumption.[13]

Those operating *within* the rational choice tradition have frequently explored the barriers actors face in trying to find apt means to meet their ends efficaciously, but assume that just as the appropriate instrumentally rational path can be found by examining the world "external" to the chooser, so the *barriers* to rationality are external impediments which could be removed, at some cost. Most significantly, people lack complete information about the efficacy of certain strategies for meeting ends. For example, a particular consumer may value fuel economy but overestimate how many miles per gallon the car he purchases will actually get; a voter may be unable to ascertain which candidate will actually support his favored policies. Those criticizing the rational choice tradition are more prone to believe that important barriers to instrumental rationality are *internal*; we may provide, for instance, more information but principals will "misprocess" it.

Thus, the critics of rational choice theory typically reverse the rational choice theorist's conceptions of what is internal and what is external. For the rational choice theorist, ends (or values) are subjective products of the internal subject, *expressing* the unique, particularized nonuniversal self (in fact, our unique menu of ends is what most individuates us), while means are objective, universal, better-or-worse instruments to meet *any* particular ends.

---

13. Thus, cognitive psychologists may, for instance, emphasize the context-dependence view that the decisionmaker will evaluate an option differently depending on the presence or absence of a third option: Instead of evaluating an option based on its worth to him (e.g., the Mark Cross pen), the consumer evaluates either Mark Cross pens as part of a set with inferior pens or Mark Cross pens outside such a set. "Renegade" economists, influenced more by sociology and history than psychology, are prone to note that tastes are, at a minimum, path-dependent (people become highly attached to the things they can get and are used to without regard to some transhistorical taste for the option), culture-specific, and manipulable by those interested in sustaining demand for goods. *See generally* Herbert Gintis, *Consumer Behavior and the Concept of Sovereignty: Explanations of Social Decay*, 62 AM. ECON. REV., May 1972, Papers & Proceedings, at 267; Herbert Gintis, *A Radical Analysis of Welfare Economics and Individual Development*, 86 Q.J. ECON. 572 (1972).

Lewin argues that economists willing to entertain attacks on neoclassical methodology have been considerably more open to the attacks from psychology than from sociology. In her view, though, the sociological attacks are far more significant. *See* Shira B. Lewin, *Economics and Psychology: Lessons For Our Own Day From the Early Twentieth Century*, 34 J. ECON. LIT. 1293, 1299 (1996).

There might well be two reasons for the *relative* willingness to embrace psychological rather than sociological theory. First, the psychologists are methodologically individualistic in the same way that economists are: They, too, posit a world that can be understood as constructed from the interactions of autonomous individuals, with some set of motives (conscious or unconscious) and abilities (fully rational or "defective"). Second, most of the psychologists' critiques may, as we explore in the text, be themselves translated "economistically." What the psychologists typically identify are the heuristics people use given the "costs" of pursuing locally unboundedly rational strategies. Since economists have long emphasized the need for trade-offs given scarcity, it is not surprising to find that people don't "get" all the "rationality" they would want if it were costless to attain it.

Furthermore, the failure to adopt better means is a function only of external impediments since virtually all people (except those lacking minimal rational competence) have the capacity to agree on apt means once external impediments are removed. The critics instead believe that ends are often a product of external forces (context, framing, history, culture, "influence"), while means are often mischosen—despite the removal of all the external impediments to rational choice—because of internal processing deficits.

The rational choice theorists typically counter with four arguments, presented here in their essence. First, they say that the critics have often misinterpreted the ends that the subjects in fact seek in "observing" that they have failed to meet these ends. Second, they argue that while it may appear that the agents are acting irrationally, they are in fact processing incomplete information as well as it can be processed. (At times, the notion is that the allegedly irrelevant information is actually relevant to the particular case; at other times, rational choice theorists note that it is rational to use more general rules of thumb about when information is relevant, even if it results occasionally in the use of information that is not in fact probative.) Third, rational choice theorists say that while we may observe irrationality in particular settings, it may not be stable over time since either institutional forces or individual learning will overcome it over time.[14] Fourth, they claim that even if we observe behavior that does not meet the normative ideal of rational decisionmaking, no one can improve on that behavior.

It is far more difficult than JST seem to imply to spin free of this point/counterpoint dialectical dance. In this regard, it was particularly striking for me to see how cavalierly and readily JST accept the notion that there is some established hindsight *bias*.[15] Having attended to this particular issue in some detail,[16] I think it is fair to say that rational choice theorists would be quite reasonable to emphasize two arguments. The first is that hindsight bias theorists have done a very poor job proving that subjects fail to meet their ends rather than that they have ends distinct from those experimenters attribute to them. Thus, take for example the standard experiments in which mock jurors have "demonstrably" judged tort defendants "unreasonable" for failing to anticipate a result that in fact eventuated when their own ex ante judgment

14. JST make this point. *See* JST, *supra* note 5, at 1486, 1490. They then argue (correctly in my view) that it is not always true (noting, for instance, that criminals will not correct their cognitive errors and that equity premia may persist over time). *See id.* at 1485-86. The problem, the main problem I return to in Part II, in terms of JST's ambition to develop a "behavioral theory" is that it is almost impossible to predict when "correctives" will and won't operate in their view. They seem to emphasize that irrationality will not persist when there are profit opportunities made available by others' irrationality, but, as I note later, it would seem that an entrepreneur aware of the irrational equity premium could borrow money to establish funds that would give investors bond-like returns and invest this borrowed money in equities.

15. *See id.* at 1523-32.

16. *See* note 10 *supra*.

of the probability of that result would actually have been lower.[17] It is impossible to say with confidence that what the jurors did was misapply a Hand negligence formula[18] rather than substitute an alternative end (something closer to strict liability). In a number of other experiments purporting to demonstrate cognitive hindsight bias, in fact, it is absolutely lucid that the outcome information was not being used to reassess prior probabilities (e.g., of finding probative evidence in a warrantless search) but to come up with an alternative normative scheme (e.g., one in which only factually innocent property owners are entitled to sue under Section 1983 for Fourth Amendment violative searches).[19]

The second argument is that hindsight bias theorists have done a wholly inadequate job explaining when "primary hindsight bias"[20]—the term Elliot, Folger and I use for the tendency to increase one's estimate of the prior probability of an event once one sees the actual outcome—represents fully rational Bayesian updating, when it represents a "bias," and when it represents a useful heuristic. One may believe the prior probability of drawing a club from a deck of cards is one in four prior to seeing what cards are drawn; if an unbiased dealer draws 490 clubs out of 500 draws, though, it is plainly *rational* to believe the ex ante odds of drawing a club were considerably higher than one in four. If, though, an experimental subject sees four cards drawn and three are clubs, he would be underestimating the prevalence of sampling variation to declare that the ex ante odds of drawing a club were three in four. If prior probabilities were harder to calculate than they are in this example, it might be useful to assume that one's best evidence of what was likely to happen was to observe what in fact happened; the use of such a heuristic might itself be rational.[21]

---

17. *See, e.g.*, Kim A. Kamin & Jeffrey J. Rachlinski, *Ex Post ≠ Ex Ante: Determining Liability in Hindsight*, 19 LAW & HUM. BEHAV. 89 (1995) (discussing such an experiment).

18. Judge Learned Hand's famous formula runs: "[I]f the probability [of injury] be called P; the injury, L; and the burden [of adequate precautions], B; liability depends upon whether B is less than L multiplied by P; i.e., whether B < PL." United States v. Carroll Towing Co., 159 F.2d 169, 173 (2d Cir. 1947) (Hand, J.).

19. *See* Dorothy K. Kagehiro, Ralph B. Taylor, William S. Laufer & Alan T. Harland, *Hindsight Bias and Third-Party Consenters to Warrantless Police Searches*, 15 LAW & HUM. BEHAV. 305, 310-12 (1991) (finding experimental subjects are more likely to think cotenants had a right to authorize searches of the suspect's home when contraband or illegal evidence is found). While the authors attribute this attitude to hindsight bias, the fact is that the cotenant's authority in no way turns on whether he is reasonably suspicious of the presence of evidence or contraband, so that it is quite clear that the experimental jurors simply substitute a norm scheme in which the "guilty" are stripped of civil rights protection for one in which we judge the lawfulness of the search.

20. *See* note 10 *supra*.

21. Such a heuristic could be rational, in fact, in two distinct senses. First, it might reduce the aggregate amount of error in judgment that the subject makes; that is, he will estimate probabilities more accurately over a range of cases if he uses this technique than if he uses some alternative technique. Second, in a weaker but still rational sense, it may lead to a higher rate of errors in judgment, but the costs of the increase in judgment errors might be lower than the cost of using a more

JST do acknowledge the third and fourth "standard" rational choice responses to those fixed on breaches of rationality: that institutions develop some correctives for individual irrationality,[22] and that irrationality may be difficult to correct.[23] Of course, those more devoted to the rational choice perspective will question whether JST have dealt adequately with the force of these objections, let alone the two immediately preceding arguments.[24]

---

particularistic factfinding method rather than employing the rule of thumb. JST believe heuristics are rational only in the second sense. *See* JST, *supra* note 5, at 1477-78. But JST explain neither why they think heuristics are not irrational even in that second sense (other than by asserting that they would not be adopted otherwise) nor why they are not rational in the even stronger first sense I mention (other than by assuming that subjects always have better particularistic factfinding techniques available to them).

22. In this regard, they draw Jeffrey J. Rachlinski, *A Positive Psychological Theory of Judging in Hindsight*, 65 U. CHI. L. REV. 571 (1998). Rachlinski notes that judges may develop techniques to counter their hindsight bias, e.g., overcoming the problem in patent of believing all inventions were obvious, once made, by focusing on "secondary considerations" like the commercial success of the new invention. *See id.* at 613-15.

23. Ultimately, though, JST believe that hindsight bias could best be dealt with by altering the burden of proof (to a clear and convincing evidence standard). *See* JST, *supra* note 5, at 1530-31. I think their suggestions smack of just the sort of false scientism that besets the paper generally: Their claim that they can calibrate the gains from eliminating overdeterrence that will occur if such a move is made against the losses from underdeterrence seems to me nothing more than an assertion of faith. The problem may be worse than they anticipate if, as Elliot, Folger and I believe to be the case, hindsight bias in trials may exist in some classes of cognitively similar cases and not in others, on the basis of affective/emotional facts (e.g., sympathy or antipathy to those who were asked to make judgments of future probability) that are unrelated to *both* rational choice theory and behavioral economics. We discuss this problem in Kelman et al., Hindsight Bias and the Rational Choice Paradigm, *supra* note 10, at 66-69.

24. Rational choice theorists will doubtless question other aspects of the JST story. Space limitations (and the expectation that Judge Posner might have his own set of questions) prevent me from doing more than noting a couple of the problems. First, JST's suggestion that behavioral economists are in a unique position to deal with the magnitude of price effects is quite peculiar. *See* JST, *supra* note 5, at 1485. Behavioral economists, after all, hardly invented the concept of price elasticity.

Second, in the discussion of the ultimatum game, JST imply that rational choice theorists assume that utility maximizers will invariably maximize money income; if they did, of course, penny offers by the Proposer would be accepted. *See id.* at 1489-93. But this appears to involve precisely the sort of needlessly narrow misspecification of the rational agent's ends that sophisticated rational choice theorists caution against. Even ignoring (just for simplicity's sake) the amount the Proposer is willing to pay to feel he is a just person, a Proposer will know that he must offer more than a trivial amount to make selfish gains as long as he knows that the Responder will sacrifice something to punish someone who is ungenerous. The Responder who behaves this way does so because the psychic gains from such punishment are greater than the goods she could gain with the money; she is likely to punish the ungenerous Proposer particularly, given endowment effects, when she need only sacrifice "opportunity cost income." (I return, in Part II, to whether JST have given us a particularly helpful theory of fairness that will permit us to predict in some general sense how much people will pay to punish which sorts of breaches of what kinds of norms.)

II. INCOMPLETE THEORY

What is ultimately perhaps more bothersome about JST's piece is that they are so unself-critical about the degree to which behavioral economics can better be seen as a series of particular counterstories, formed largely in parasitic reaction to the unduly self-confident predictions of rational choice theorists, than as an alternative general theory of human behavior. Again and again, the authors seem to confuse discordant observations for a countertheory and evade questions about the gaps in the behavioral picture, seemingly believing, quite wrongly in my view, that acknowledging these gaps would fatally wound their enterprise.

Let me take just a few examples of explicit or implicit behavioral theories that JST advance, and comment on what strikes me as their manifest incompleteness as theories.

A. *Persistence of Irrationality*

First, irrationality of particular agents will be of little moment where (1) arbitrage is possible, (2) most activity is undertaken by professionals, and (3) there is opportunity for learning.[25]

Obviously, it is important to JST to be able to "predict" the appropriate domain of their theory, if they believe it is a descriptive theory that will predict outcomes well. In fact, they argue that behavioral economics is especially germane to law and economics because law often concerns itself with conditions in which irrationality will survive. It is not at all clear, though, that one could identify the situations in which JST would believe that irrationalities will persist ex ante. There appears to be a set of ex post explanations available to behavioralists to account for situations in which "rational choice" results obtain but no general theory of when to expect them.

For example, JST refer to the "equity premium puzzle."[26] Equities appear to have a long-term rate of return so far in excess of the riskless rate of return that it could not be explained by mere risk aversion; instead, it can better be explained by several distinct forms of individual irrationality that behavioral economists typically highlight.[27] But JST give nothing resem-

---

25. *See id.* at 1486. See, especially, their comment in relationship to why markets don't discipline criminals: "Outside of financial markets (and not always there), those who engage in low-payoff activities lose utility but do not create profit opportunities for others." *Id.*

26. *See id.* at 1485.

27. In the view of behavioral economists, the equity premium compensates stockholders for the risk of suffering a loss over the sort of short time horizon that imperfectly rational agents (who typically bracket decisions narrowly) generally utilize in evaluating economic events. If investors evaluate the returns on their investment portfolios once a year *and* are heavily loss averse (losses from a reference baseline are twice as significant in utility terms as dollar equivalent gains), then investors would be roughly indifferent between stocks and bonds, even though this implies an

bling a satisfactory reason to explain why entrepreneurs do not come along to "borrow" investors' money on terms approximating the terms that bond issuers make available and invest that borrowed money in equities, bidding for investors' money until the return on "bond-like" funds asymptotically approaches the (aptly) risk-adjusted return on equities. There appears to be a profit opportunity for professionals with perfectly adequate access to data to learn about long-term equity performance.

Naturally, it is always possible to argue that there are "not enough" such opportunities to "correct" the irrationality—in precisely the same way that rational choice theorists can invariably argue that their academic enemies have not *adequately* understood the complexity of the good the chooser seeks in situations in which a chooser has seemingly chosen irrationally.[28] But both rational choice and behavioral theories are better understood as "attitudes" rather than theories if they tell us no more than what one will *try* to do with discordant data.

B. *Fairness and Punishment*

Second, people will be willing to pay something (in forgone income or in out-of-pocket income) to appear fair themselves and will pay something to punish those who are unfair.[29]

Once more, JST are right to note, contrary to income-focused variants of rational choice theory—though not to variants in which there is no a priori limit on whether people gain utility from spite, admirable behavior, or anything else—that people will (sometimes) sacrifice (some) money (out-of-pocket or in forgone receipts) to punish the unfair or to behave in ways that

---

enormous gap in expected returns to the two financial instruments. *See* Shlomo Benartzi & Richard Thaler, *Myopic Loss Aversion and the Equity Premium Puzzle*, 110 Q.J. ECON. 73, 73-76 (1995).

28. Thus, the deeply committed rational choice theorist can argue that the person who seeks a good in a particular social context (e.g., a heavily advertised product) does not irrationally prefer the good because of the context but seeks an alternative good (the good attached to the social meaning the ads have helped create). Even where the context is purely cognitive, rather than social, the committed rational choice theorist may argue that tastes are not inconsistent but complex. The Mark Cross pen selected over six dollars when an inferior pen is offered is simply not the same good as the Mark Cross pen chosen over six dollars alone. Rather, it is a pen whose superiority to other goods is atypically salient (in an informational sense) or a good that should be chosen to demonstrate to the experimenters that one is smart enough to distinguish a bad from a good product.

29. *See* JST, *supra* note 5, at 1489-97. Note especially carefully the ambitions to theoretical completeness the authors demonstrate. After describing the practice of Ithaca road stand farm operators (who leave a cash box on the table in which they expect customers to put money when they buy produce, with a small slit which would make it difficult to remove money), JST proffer a view one of them previously expressed: "We think that the farmers who use this system have just about the right model of human nature." *Id.* at 1493 (quoting Richard H. Thaler & Robin Dawes, *Cooperation*, *in* RICHARD H. THALER, THE WINNER'S CURSE: PARADOXES AND ANOMALIES OF ECONOMIC LIFE 6 (1992)). But, JST continue, "We don't recommend that Mercedes dealers adopt the road-side stand selling technique." *Id.* at 1494.

they believe respectable people behave. The notion, though, that this critical observation constitutes a "theory" of this behavior is puzzling: It does not begin to answer any of the questions that such a theory would.

Put generally, it gives us little or no sense of (1) what sorts of otherwise-selfishly irrational behavior must be performed, by what classes of people, to gain self-respect,[30] (2) what sorts of treatments by others will be deemed sufficiently unfair to generate spite,[31] or (3) what, to echo a charge they make in attacking rational choice theory, is the actual *magnitude* of these effects.[32]

---

30. Thus, it may help, in a *critical* way, if one is unduly wedded to the rational choice-based instinct that people are selfish money maximizers, to know that during divorce and alimony controversies, some payers will offer more money than they would likely be forced to pay given the desire to appear like people who are "kind" or "giving" to former spouses (though this of course will be balanced—across cases? in each case?—by spite and revenge towards these same former spouses that might lead to "irrationally" low offers). But JST don't begin to tell us how we might think about the conditions in which that motive is likely to emerge, or fail to emerge, or whether it will be manifested differently where the spouse is asked to pay for the former spouse either to gain skills necessary to become less dependent, or to support a lavish consumption style, or to prevent dire poverty (though each of these "voluntary" payments could be said to meet the preference to be the kind of person the spouse wishes to be).

31. JST believe they have given some content to this issue by noting that persons believe (1) that close to even splits of undeserved receipts are acceptable, (2) that disproportionate divisions of goods to which there are differential entitlements are legitimate, and, most significantly, (3) that sales at prices above "reference transaction" prices are unfair or illegitimate. *See* JST, *supra* note 5, at 1496. But each of these observations is far more question-begging than the authors imply. In the absence of an independent theory of when receipts are deserved outside of narrow experimental settings in which the party given the right to divide receipts is truly designated randomly, we cannot predict when a division will be seen as the division of merited or unmerited holdings. Will offers to divide inherited property upon marital dissolution be viewed differently—and in what direction, let alone magnitude—from offers to divide savings from past earnings? When will parties believe sellers are engaged in price-gouging and when will these same parties believe they are simply observing changes in market prices? Why do people tolerate rapid escalation in housing prices, based on local scarcity, without engaging in transaction-thwarting, self-defeating, spiteful behavior against sellers, but turn on ticket scalping? When will they believe shortage-based price changes and explicit or implicit auctions are legitimate (there are auctions after all, for nonduplicable paintings or celebrity signatures, so we don't seem to have a general principle that *cost-based* pricing is the inevitable reference transaction)?

32. If we know (most) people (in Ithaca) will pay one dollar for corn they could readily take (though given rational risks of apprehension, a purely selfish utility maximizer might pay the dollar), but wouldn't pay $40,000 for a Mercedes they could take (with no fear of apprehension?), we've got some pretty broad parameters to work within. But would we be able to guess how much the average person would pay for such a good or the distribution of offers? Does anything in behavioral economics give us any systematic idea how much more people will pay (or forgo) to avoid what is legally defined as theft (taking the corn, taking money the farmer collected in the box) than to avoid what is conventionally defined as rude behavior (not tipping someone one will never see again, taking more than one's "fair share" of a buffet at a public event)? Does the in-kind/in-cash distinction matter in either or both of these examples? Does behavioral economics give us systematic evidence about which people will typically pay (forgo) more and which less, so that we might predict where the Ithaca system of sale is and is not stable?

More generally, the fact that behavioral economists have so few explanations of the psychological phenomena they highlight makes it extremely difficult to imagine how the theorist will deal with "new" cases. Take the endowment effect, for instance: There seems to be little doubt that there is some significant aversion to the sale of things one physically possesses, some asymmetry

C. *Lawmaking and Fairness Norms*

Third, in explaining the actual content of law, it is important to recognize that laws that enforce fairness norms will often be enacted, even when doing so meets neither the selfish material ends of any organized constituency nor efficiency ends.[33]

---

between asking and offering prices. But there is really no evidence at all about the roots of this observed behavior. Do people form equivalently strong attachments to goods instantly or does the reluctance to sell grow over time? To what degree do people seek to escape the process of calculating by adopting a rule of thumb that they don't evaluate the opportunity costs associated with holding assets, while when in their "purchasing" mode, they must calculate in any case? To what degree do people seek to escape the world of marginal trade-offs (and thus approximate the sense of satiation or "post-economic" existence) by deeming as many domains off-limits to calculation as they can, decommodifying even already-owned commodities because the commodity form is itself problematic? To what extent do people close transactions so as to avoid a spiral of obsession about each decision they might have made? To what extent do people "endow" either income or specific commodities whose possession signals status, so that their loss threatens status degradation?

If we could gain a genuine understanding of such roots (and I doubt we can get very far in that direction), we would not simply satisfy our abstract intellectual curiosity but we would better be able to "predict" what would and wouldn't be part of a decisionmaker's "endowment." For many of us, our "instinct" is that a "typical" gambler would more readily gamble his first day's winnings in Las Vegas (i.e., would not be loss averse to these) than gamble an inheritance he learned of while on the Vegas vacation, which in turn he would more readily gamble than a wage increase he learned of. But none of these "instincts" derive from a clear picture of what drives the endowment effect.

Similarly, neither the cognitive nor the affective origins of what Elliot, Folger and I call "primary" hindsight *bias* (the tendency to increase estimates of the ex ante probability of events that actually occurred, even when the ex post information is not in fact probative of ex ante probabilities; *see* note 10 *supra*) are clearly understood. Cognitively, it may be difficult to imagine the counterfactuals that did not occur. As an affective matter, people may take comfort in believing that what happened was inevitable, rather than a "roll of the dice." Having little or no insight into the bias' roots, it is impossible to know how strong one thinks the effect will be in any particular case, or even whether it will be universally present.

33. *See* JST, *supra* note 5, at 1508-15. I want to ignore for now the fact that there is little reason, looking at JST's examples, to believe they have demonstrated that the laws they describe either are not efficient, or, if inefficient, fail to meet the needs of organized interest groups. Thus, for instance, it is hard to see why (those whom conventional public choice theorists would probably describe as) politically well-organized ticket sellers would ever favor scalping: Presumably, they have chosen the optimal selfish level of monopoly rent extraction given long-term pricing strategies and cannot possibly gain anything by allowing others to extract rents they've chosen not to exploit. A refusal (by law or custom) to allow auctions for scarce hot Christmas presents to develop, *see id.* at 1512, may be potentially Pareto efficient: Each parent is spared implicit or explicit accusations that he could have gotten the hot item if he'd only cared enough, and would willingly pay to be spared that accusation were a market in nonaccusations feasible (but, of course, it's not, not just for typical transaction costs reasons but because paying not to be allowed to pay for the hot item destroys the "good" the payer seeks, immunity from the options that markets give).

Instead, my focus here is on the wholly unsystematic nature of the "theory." In this regard, the flaw is duplicated by JST's account of "availability entrepreneurs." *See id.* at 1519. JST give neither an account of why or how the entrepreneurs pick the particular incidents to publicize that they pick (are they irrationally duped too? or exploiters of a readily duped Congress and public?), nor an account of when such efforts might fail or how trade-offs are made when there are multiple "unduly" available crises with which to cope.

Once more, there is an implicit "counterstory" to both of the prevalent conventional accounts of political behavior here.[34] But it is not itself a story about political behavior that is even lucid, let alone compelling or complete. It appears that JST argue essentially that legislators indeed seek reelection above all,[35] but that they can be reelected so long as they satisfy voter preferences, which are themselves often based on fairness norms[36] rather than a desire either to correct genuine market failures or to seek selfish material gains. To the degree JST advance a theory, then, it appears to be a theory of a perfectly functioning preference-responsive democracy, in which there is some modest (counterintuitive?) insight into the preferences that voters will manifest. But there is little explanation of why JST think legislators are such excellent agents, or whether they are more typically good agents when translating "fairness" norms than translating other voter preferences, or when they believe legislators sacrifice voter sentiment either to organized rent-seekers or to their own ideological or technocratic beliefs. Again, it seems, what we've really got is another *outlook* to work with, not a developed positive theory of the content of law.

III. Conclusion: The Hubris Problem

In my view, the chief (intellectual) problem with rational choice theory has been its pretense to completeness. Rational choice theorists discarded observations and stretched to interpret discordant data in manners that were often wholly unpersuasive, sometimes merely banal and unhelpful, to maintain the illusion that human behavior and social organization were both, in descriptive or positive terms, readily explicable and, in normative terms, readily perfected. Behavioral theorists have done a wonderful job exposing

34. In the mainstream modern public choice account, readily organized rent-seekers dictate legislation by outbidding the widely dispersed victims of the rent-seekers, who are unable to overcome free rider problems to offer sufficient money and political services to legislators with a simple reelection motive. In the "public interest" tradition, legislatures correct market failures and redistribute so as to maximize some (distribution-sensitive) social welfare function.

35. In that sense, JST are like conventional public choice theorists and not like those who think legislators are significantly motivated by ideology, though I see no basis at all for them to adopt this view. For a fuller critique of this proposition, Mark Kelman, *On Democracy-Bashing: A Skeptical Look at the Theoretical and "Empirical" Practice of the Public Choice Movement*, 74 VA. L. REV. 199, 217-23 (1988) and sources cited therein.

36. Once more, one has the same criticism that I raised above: There is nothing resembling a theory of when voters will find market transactions unfair and when they won't. Rent control regulations—embodying a refusal to allow sales of housing services above reference prices—are enacted in some places and not others. There is no significant political demand for price control for owner-occupied housing, though, even where there is rent control: Above-reference prices are common. The answer may be wholly circular: Above-historical rental prices must be deemed *more* unfair than above-historical purchase prices (and we know this because there is more demand for rent control).

this incompleteness and reminding us both how many distinct sources of understanding we have[37] and how limited that understanding really is.

We should understand that imbedding ourselves in the dialectic dialogue between rational choice theory and its critics will make us wiser users of the rich, inexorably overwhelming data with which we have to deal. It is *constructive*, not merely "skeptical," to anticipate more fully the range of plausible interpretations of human conduct, and constructive to understand that open-textured interpretivism is not the enemy of knowledge but of rigidity and self-delusion.

37. Children selecting Halloween candy "diversify" their candy portfolios when offered two pieces of two distinct candies at a single house, but select their favorite candy two successive times if offered one piece of candy at two successive houses. If their "desire" were to diversify their candy holdings, they should do so whether the choice was simultaneous or sequential. *See* Daniel Read & George Loewenstein, *Diversification Bias: Explaining the Discrepancy in Variety Seeking Between Combined and Separated Choices*, 1 J. EXPERIMENTAL PSYCHOL.: APPLIED 34, 45-46 (1995). But we have nothing resembling a *theory* to tell us why a single visit to a house might be an apt "mental bracket" in making diversification decisions while an evening's activity is not, just participant-observers' cultural instincts, of the form sharpened as much by introspection and literature as by anthropology.