COMMENT ON NOLL AND KRIER, "SOME IMPLICATIONS OF COGNITIVE PSYCHOLOGY FOR RISK REGULATION"

*COLIN F. CAMERER**

We have known about systematic violations of the expected utility (EU) theory of choice for almost forty years, since Maurice Allais got Jimmie Savage to violate his own "sure-thing principle" (or "independence axiom") while making hypothetical choices over lunch in Paris. Savage was victimized by some combination of wine and intuition. The wine's effect is gone, but the intuition is not: devotion to EU sometimes produces unappealing choices.[1]

For half of the forty years since Allais's choice demonstration, psychologists have been actively studying judgment, too. The reigning paradigm focuses on simplifying "heuristics" people use to lighten their cognitive load but which produce systematic violations of normative theory ("biases").[2] Violations occur even when subjects are experts about normative rules (for example, statisticians) or in substantive areas (for example, doctors). When subjects are paid for accuracy, violations are sometimes reduced slightly, but rarely eliminated. Given the robust evidence, it is high time to study how these violations affect the institutional world of households, firms, markets, and bureaucracies outside the psychology lab. Noll and Krier take a useful step in this direction.[3]

* Professor, Decision Sciences Department, Wharton School of the University of Pennsylvania.

[1] See Martin Weber & Colin F. Camerer, Recent Developments in Modelling Preferences under Risk, 9 OR Spektrum 129 (1987), for a review.

[2] See Judgment under Uncertainty: Heuristics and Biases (Daniel Kahneman, Paul Slovic, & Amos Tversky eds. 1982).

[3] Roger G. Noll & James E. Krier, Some Implications of Cognitive Psychology for Risk Regulation, in this issue.

[*Journal of Legal Studies*, vol. XIX (June 1990)]

I. Putting Prospect Theory to Work Productively

Noll and Krier spend much of their article trying to derive implications from the value and decision-weighting functions of prospect theory. But many policy situations are too broad or too narrow for prospect theory. Most choices involve distributions over several attributes (so-called multiattribute utility), typically lives and dollars. Prospect theory is too narrow for these choices. Also, much policy debate revolves around judgments of likelihoods ("How lethal is Alar?") or riskiness ("Can we tolerate the possibility of nuclear catastrophe?") rather than around choices; prospect theory is then too broad.

Even when it is appropriate, deriving specific predictions from prospect theory is a struggle because choices depend both on the shape of utility (or "value") and on how probabilities are weighted; often these two interact and make predictions indeterminate. In my view, prospect theory is a skeleton of empirical facts over which a theory is draped. It might be easier to get policy-level predictions directly from the empirical facts. For example, framing effects are an empirical fact. They occur because choices depend on the reference point from which gains and losses are evaluated. As Noll and Krier point out, public debates about risk will then depend on whether possible outcomes are evaluated as gains (relative, say, to an inferior technology) or losses (relative to a reference point of zero risk). As reference points shift, public attitudes toward risk shift, too. The fickleness of public attitudes is thus an important indirect implication of prospect theory; it depends on the empirical existence of framing effects, not on which frame people choose (which is good, because frame choices are hard to predict).[4]

Another empirical fact with many implications is loss aversion. Loss aversion is powerful by itself: it can explain why people dislike risks in which they could gain or lose, without invoking curvature of utility functions. Loss aversion can also explain "endowment effects": the large, systematic deviation between buying and selling prices. (Selling prices are higher because giving something up causes a loss, which is more aversive than the gain from buying; so people must be paid more to sell.) In studies on siting a hazardous waste repository in Nevada, most people said they would not allow a site within one hundred miles unless they were paid $5,000 a year for twenty years (their selling price). The same people said they would only pay one hundred dollars a piece to move the facility away once it was located nearby. These large buying/selling price discrepancies

[4] See Baruch Fischhoff, Predicting Frames, 9 J. Exp. Psych.: Learning, Memory, and Cognition 103 (1983). (It is hard to infer a frame for choices that is consistent with subjects' judgments of which frames were most "natural.")

are experimentally robust—they tend to shrink in magnitude, but not disappear, when subjects are highly motivated and experienced.[5] They are also disturbing: which price is a better measure of value?

The tenacity of endowment effects may also help explain why entitlements like Social Security are virtually impossible to cut from the federal budget and why imposing new risks (for example, nuclear power) is difficult while comparable familiar risks (smog or earthquakes in Los Angeles) are accepted, even, until recently, joked about.

II. Other Directions

Let me sketch four other categories of thinking in behavioral decision theory that might help us understand and regulate risk.

1. *Intertemporal Choice.* Most risks unfold over time. The standard generalization of EU in such settings is "discounted utility." For a comparison to present outcomes, future outcomes are weighted by a discount factor that reflects the ratio of marginal utilities of consumption then and now. Presumably the discount factor is less than one because of "impatience," probability of death, and so on.

In several recent studies, many anomalies in intertemporal choice have emerged that are surprisingly like the anomalies in EU.[6] For example, people are more impatient (discount rates are higher) for short delays than for long delays.[7] If two Thai dinners next year are as good as one Thai dinner now, then two dinners in thirty years and one dinner in twenty-nine years should be equivalent. For most people they are not (they prefer the two dinners in thirty years). The size of short-term discount rates might explain why citizens are so impatient to reduce risk and solve problems by spending a lot now (for, say, National Guard patrols of crack-infested District of Columbia neighborhoods) rather than spreading spending over future years, when it might be used more efficiently.

Another anomaly in intertemporal choice is that speeding up risks is more painful than delaying them is pleasant—much as selling is more painful than buying is pleasurable. The delay-advancement asymmetry may explain why citizens are extremely frustrated by slowly developing

[5] See, for example, David S. Brookshire & Don L. Coursey, Measuring the Value of a Public Good: An Empirical Comparison of Elicitation Procedures, 77 Am. Econ. Rev. 554 (1987); David W. Harless, More Laboratory Evidence on the Disparity between Willingness to Pay and Compensation Demanded, 11 J. Econ. Behav. & Org. 359 (1989).

[6] See George Loewenstein, Anomalies in Intertemporal Choice: Evidence and an Interpretation (working paper, University of Chicago, Graduate School of Business, 1988).

[7] This "immediacy effect" (special preference for immediate outcomes) parallels the "certainty effect" (special preference for certain outcomes) in prospect theory.

technology that always seems to get delayed (for example, testing of AIDS drugs).

2. *Difficulty of Learning.* Psychologists who study judgment are mostly pessimistic about the ability of people to learn from experience. For instance, expert clinical psychologists with decades of practice are no better at diagnosis than graduate students and little better than novices (though they are more confident and quicker).[8] Predictions of simple statistical models about student academic success, recidivism of criminals, mental illness, financial performance, and the like are invariably better than predictions of experts.[9]

Why do people not learn more from experience? Studies show that learning is difficult unless feedback is clear, frequent, and quick. Weathermen get such feedback; they do learn to make good forecasts.[10] For chief executive officers and policymakers, learning is difficult because important risks unfold over long periods of time and feedback is unclear. It is hard to learn about risk-reducing programs like promotion of safe sex, changes in the drinking age, or tighter security at airports. Extremely successful programs, such as polio vaccination, cause risks to disappear so entirely that the feedback from them may be forgotten.

Especially maddening are systems with long lags between the initiation of a program and substantial effect (like environmental cleanup, birth control, or educational programs). In long-lag systems, optimal strategies are hard to derive and are often counterintuitive; human performance in such systems (studied experimentally)[11] is terrible.

The main implication of this pessimistic view of learning is that program evaluation must be careful, frequent, and quick for policymakers and citizens to learn from results.

3. *Ambiguity about Probability.* To a decision theorist, most risks are not risks at all. When the probability distribution of outcomes is unknown, such situations are usually called "uncertain" (Frank Knight's term) or "ambiguous." The classic thought experiment about ambiguity was proposed by Daniel Ellsberg.[12] Consider a bingo cage K (for

[8] See Howard Garb, Clinical Judgment, Clinical Training, and Professional Experience, 105 Psych. Bull. 387 (1989).

[9] See Colin F. Camerer, General Conditions for the Success of Bootstrapping Models, 27 Org. Behav. Human. Perf. 411 (1981); Robyn M. Dawes, David Faust, & Paul E. Meehl, Clinical versus Actuarial Judgment, 243 Science 1668 (1989).

[10] See Sarah Lichtenstein, Baruch Fischhoff, & Lawrence D. Phillips, Calibration of Probabilities: The State of the Art to 1980, in Kahneman, Slovic, & Tversky eds., *supra* note 2.

[11] See John D. Sterman, Testing Behavioral Simulation Models by Direct Experiment, 33 Mgmt. Sci. 1572 (1987).

[12] Daniel Ellsberg, Risk, Ambiguity, and the Savage Axioms, 75 Q. J. Econ. 643 (1961).

"known") with fifty red and fifty black balls and a cage U (for unknown) with one hundred red and black balls of unknown distribution. People often prefer to bet on a red draw from the K cage instead of a red draw from the U cage; the same is true for bets on black.

Suppose subjective probabilities are revealed by choices. Since people are typically indifferent between betting on red or black from the K cage, we conclude that P_K (red) = P_K (black) = .5. Preferring a red draw from K implies P_U (red) < .5;[13] preferring a black draw from K implies P_U (black) < .5, too. P_U (red) + P_U (black) is then less than one, violating the principle of additivity of complementary events.[14] Something must give: either additivity or the principle of inferring subjective probabilities from choices.

Ellsberg's elegant paradox shows that probabilities must carry two burdens at once: they must express likelihood and they must also express the amount of information, confidence, or "weight of evidence." In each cage, red and black seem to have equal likelihood, but we have more confidence that the likelihoods are .5 in cage K. One number cannot express both properties.

In some recent formal theory,[15] it has been shown that it is reasonable to use a variant of "minimax" to make decisions under ambiguity: choose the gamble that has the highest minimum expected utility, where the minimum is taken over all the possible distributions the probabilities might have.

Minimax thinking (freshly blessed by this theoretical justification) appears in studies of actuaries and underwriters, who attach higher selling prices when risks are ambiguous.[16] And, of course, changes in tort law—even possible changes—generate ferocious ambiguity that could (and should, according to the recent theory) scare insurance companies out of many markets because the risk of rule changes is not diversifiable.[17]

4. *Value Ambiguity and Contingent Weighting.* If probabilities can

[13] In subjective expected utility, probabilities can (only) be inferred from bets. Since the gain from winning a red bet from the U and K urns is the same, preferring the K urn reveals that P_K (red) > P_U (red).

[14] Additivity means $P(A \text{ or } B) = P(A) + P(B)$ if A and B are mutually exclusive (that is, $P(A \text{ and } B) = 0$).

[15] See Truman Bewley, Knightian Decision Theory: Part I (unpublished manuscript, Yale University, Cowles Foundation, 1986); David Schmeidler, Subjective Probability and Expected Utility without Additivity, 57 Econometrica 571 (1989).

[16] Robin M. Hogarth and Howard Kunreuther, Risk, Ambiguity, and Insurance, 2 J. Risk Uncertainty 5 (1989).

[17] See, for example, Patricia Danzon, Tort Reform and the Role of Government in Private Insurance Markets, 13 J. Legal Stud. 517, 539 (1984); George L. Priest, The Current Insurance Crisis and Modern Tort Law, 96 Yale L. J. 1521 (1987).

be ambiguous, can values be ambiguous, too? They can and are. Value ambiguity can explain buyer/seller discrepancies just as loss aversion does. Suppose people are not sure what hazardous waste risks are worth to them, but they have an ambiguous distribution of possible worths. When asked to pay to reduce risk, they name a price at the low end of the worth distribution. Asked their price for bearing more risk, they name a number on the high end of the distribution.

A dramatic class of "compatibility" effects are consistent with value ambiguity, too. Consider two risks, a "*P* bet" (for example, 95 percent chance of three dollars) and a "$ bet" (for example, a 30 percent chance of ten dollars). Asked to choose between the two bets, most people who pick the *P* bet instead of the $ bet later attach a higher selling price to the $ bet than to the *P* bet. These preference reversals are astonishing and robust.[18] (Money pumping of subjects[19] reduces the dollar magnitude of the pricing discrepancy, but not its frequency.)[20]

Psychologists now think that such reversals occur because people overemphasize the attribute that is compatible with the response mode they are using. When people choose, probability looms large (it is the more "prominent" attribute—the theory does not say exactly why—and thus gets more weight in choice). When they set a price, the dollar attribute looms large.[21] These data are especially bothersome because the choices are simple gambles with known probabilities and only one attribute of value (dollars). Imagine how arbitrary revealed preferences over complicated policy choices, with many attributes, might be.

The implication for risk assessment is that choices between risks will be different from patterns of preference inferred by "matching" attribute values until two risks are equally preferred.

[18] See Sarah Lichtenstein & Paul Slovic, Reversals of Preference between Bids and Choices in Gambling Decisions, 89 J. Exp. Psych. 46 (1971); David M. Grether & Charles R. Plott, Economic Theory of Choice and the Preference Reversal Phenomenon, 69 Am. Econ. Rev. 623 (1979); Paul Slovic & Sarah Lichtenstein, Preference Reversals: A Broader Perspective, 73 Am. Econ. Rev. 596 (1983); Amos Tversky, Paul Slovic, & Daniel Kahneman, The Causes of Preference Reversals, Am. Econ. Rev. (in press).

[19] In a money pump, or Dutch book, the experimenter would continually sell $ bets to subjects at high prices, switch the $ bets for *P* bets, then buy the *P* bets back from subjects at lower prices. This cycle leaves the subjects where they began—with no bets—except for the loss of wealth from buying high and selling low.

[20] See Joyce Berg, John Dickhaut, & John O'Brien, Preference Reversal and Arbitrage, in 3 Research in Experimental Economics (Vernon Smith ed. 1985); Marc Knez & Vernon L. Smith, Hypothetical Valuations and Preference Reversals in the Context of Asset Trading, in Laboratory Experimentation in Economics: Six Points of View (Alvin Roth ed. 1987).

[21] Amos Tversky, Shmuel Sattath, & Paul Slovic, Contingent Weighting in Judgment and Choice, 95 Psych. Rev. 371 (1988). (For compatibility effects to exist, people must be ambiguous about the proper trade-off between dimensions.)

Consider the comparison between two programs that both clean a polluted beach. Full cleanup costs $750,000 and makes the beach suitable for both sunbathing and swimming. Partial cleanup costs $250,000 but leaves the beach unsuitable for swimming.

	Cost ($)	*Outcome*
Full	750,000	can swim
Partial	250,000	cannot swim

In a choice between the two, 48 percent of people preferred full cleanup.[22] But policy situations sometimes require a matching of budget amounts to make programs equally attractive, rather than a choice between programs. A proponent of partial cleanup might ask, "Tell me how much a partial cleanup program would have to cost to make it just as good as full cleanup?" That is, match the value of C in the table below to make full and partial cleanup equally appealing.

	Cost ($)	*Outcome*
Full	750,000	can swim
Partial	C	cannot swim

According to contingent weighting, people will underweight cost relative to the extent of cleanup in making a choice (assuming cleanup is more important to people than money). But in the matching task (choosing C), the matched dimension of cost will be overweighted. People will state large values of C because the gap between $750,000 and C—the marginal cost of full cleanup—looms larger when matching. Then C will be greater than $250,000. But if full cleanup at $750,000 and partial cleanup at cost C are equally good, and $C > $250,000, then partial cleanup at a cheaper cost of $250,000 is certainly better than full cleanup at $750,000.[23] The matching response mode should therefore favor the more economical partial cleanup (at the initial costs of $750,000 and $250,000), and it did in experiments: 88 percent of matching-mode subjects expressed an implicit preference for partial cleanup (by choosing $C > $250,000), but only 48 percent preferred partial cleanup in a direct choice.

These data show that the way in which choices are made, directly or by

[22] *Id.* at 373.

[23] We assume that spending less is preferred to spending more and that choices are transitive. (Partial cleanup at cost C is indifferent to full cleanup at $750,000, and partial cleanup at $250,000 is preferred to partial cleanup at cost C, which implies that partial cleanup at $250,000 is preferred to full cleanup at $750,000.)

filling in attributes during debate and negotiation, makes a striking difference in what is decided.

III. Some Subtleties in Behavioral Decision Theory

Some of Noll and Krier's semantic usage is unorthodox. The orthodox definitions are sharper and more fruitful. Noll and Krier define "cognitive" decision theory as prospect theory and studies of judgment heuristics. We prefer the term "behavioral" decision theory because judgment errors result from motivation (what people want) as well as from cognition (how people think). Behavioral decision theory also includes research on learning, problem representation, errors in logic, modeling complex judgments, and more.

Noll and Krier call the "representativeness" heuristic "a tendency to reason by analogy to previous circumstances."[24] The original definition is deliberately more precise: people judge the likelihood that a sample was generated by a statistical process by how well the sample represents (or resembles) the process. (Representativeness may also be used to judge the likelihood of a value coming from a distribution, an object coming from a class, or an effect being due to a cause.)[25]

Representativeness causes mistakes by leading to violations of normative principles. Two examples follow: (1) Since prior probabilities play no role in the mental calculus of representativeness, people will fear hazards with low prior probabilities (like economic depression) when they see warning signs that are highly representative of the hazards. (2) If people expect the future to be representative of the past, they will ignore or misattribute the natural effects of regression toward the mean (when the past is extreme, the future will be less extreme). For example, if policies are enacted when a problem is especially severe, and its severity naturally regresses, the policy will *appear* to have worked even if it did not actually help.

The crucial axiom in expected utility is "independence": if two risks yield a common outcome with a common probability (and differ on other outcomes), preference between them should be independent of the common outcome. A related property is called "dynamic consistency": future choices should be consistent with plans made during previous choices. (More formally, choices at early nodes in a decision tree should assume choices at later nodes will be optimal.) Noll and Krier call the

[24] Noll & Krier, *supra* note 3, at text around note 10.

[25] Amos Tversky & Daniel Kahneman, Judgments of and by Representativeness, in Kahneman, Slovic, & Tversky eds., *supra* note 2. (The source Noll and Krier cite in their note 11 actually cites an earlier Tversky and Kahneman paper.)

latter principle "rational expectations,"[26] which usually refers to the more general assumption that agents in an economic theory understand that theory.

Their unorthodox labeling obscures the important difference between dynamic consistency (also known as "rational expectations") and independence. In expected utility theory, those two principles go hand in hand, so calling dynamic consistency the crucial property is odd (actually, independence is crucial) but harmless. But in other theories, the two principles need not coincide. In prospect theory, people who "isolate" future choices are dynamically consistent, but they violate independence. Then policy reversals will be uncommon even if people use prospect theory to make choices, exactly contrary to Noll and Krier's conclusion.[27]

IV. Conclusion

Evidence that judgments and choices often violate normative principles has aroused curiosity about the fundamental question of how individual judgments and choices aggregate to group, firm, market, and governmental decisions. Noll and Krier take up the governmental part of the question with gusto. Their analysis relies mostly on the value and decision-weighting (probability-transformation) functions of prospect theory. Producing direct predictions from these functions is hard; simpler analyses might be equally fruitful. Other topics in judgment research that might be useful, too, include choice over time, difficulty of learning, ambiguity about probabilities, and contingent weighting of dimensions. These ideas should help sharpen our understanding of the demand side of risk policy—that is, how people think about risk and its reduction.

[26] Noll & Krier, *supra* note 3, at sec. I.

[27] *Id.* at sec. IIA.