

Article

Government Regulation of Irrationality: Moral and Cognitive Hazards

Jonathan Klick[†] and Gregory Mitchell^{††}

Several years ago the ethicist Daniel Wikler provocatively asked, “If we claim that relative intellectual superiority justifies restricting the liberties of the retarded, could not exceptionally gifted persons make the same claim concerning persons of normal intelligence?”¹ Wikler’s question, posed originally to raise doubts about paternalism directed at the developmentally disabled, possesses a new relevance today, as legal elites increasingly claim that “persons of normal intelligence” exhibit numerous irrational tendencies that justify restrictions on market and nonmarket transactions.² These new regulatory

[†] Jeffrey A. Stoops Professor of Law and Courtesy Professor of Economics, Florida State University, jklick@law.fsu.edu.

^{††} Sheila M. McDevitt Professor of Law and Courtesy Professor of Psychology, Florida State University, gmitchel@law.fsu.edu. Please address correspondence to either author at Florida State University College of Law, 425 W. Jefferson Street, Tallahassee, FL 32306-1601. An earlier version of this paper benefited from the comments of Amitai Aviram, Tyler Cowen, Matthias Hild, Adam Hirsch, Dave Klein, Brett McDonnell, Peter Oh, Jeff Rachlinski, Mark Seidenfeld, Alex Tabarrok, Phil Tetlock, and participants in a faculty workshop at the University of Virginia School of Law.

1. Daniel Wikler, *Paternalism and the Mildly Retarded*, 8 PHIL. & PUB. AFF. 377, 377 (1979).

2. This new paternalism represents a fairly simple and direct application of the dominant message emanating from the emerging behavioral law and economics movement, namely that people of normal, and even superior, intelligence fail to pursue their interests rationally in many important situations. See, e.g., Jon D. Hanson & Douglas A. Kysar, *Taking Behavioralism Seriously: The Problem of Market Manipulation*, 74 N.Y.U. L. REV. 630, 633 (1999) (“These cognitive illusions—sometimes referred to as ‘biases’—are not limited to the uneducated or unintelligent, and they are not readily capable of being unlearned. Instead, they affect us all with uncanny consistency and unflappable persistence.” (footnotes omitted)); Christine Jolls et al., *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471, 1541 (1998) (“In its normative orientation, conventional law and economics is often strongly anti-paternalistic. . . . [B]ounded rationality pushes toward a sort of anti-

proposals range from hard forms of paternalism, in which the government determines what is best for citizens and accordingly restricts the freedom of citizens to act otherwise, to softer forms of paternalism, in which the government regulates the form in which information and options are presented to citizens and restricts the role of laypersons in the market, legal, and political systems without completely controlling choices.³

antipaternalism—a skepticism about antipaternalism, but not an affirmative defense of paternalism.”); Russell B. Korobkin & Thomas S. Ulen, *Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics*, 88 CAL. L. REV. 1051, 1085 (2000). Courts seem to be increasingly open to the quasi-rationality message offered by behavioral law and economics scholars. See *infra* note 109.

3. For instance, Russell Korobkin uses the argument from irrationality to found two hard-paternalism proposals. First, Korobkin argues that cognitive limitations of employees in part justify government-mandated health benefits. See Russell Korobkin, *The Efficiency of Managed Care “Patient Protection” Laws: Incomplete Contracts, Bounded Rationality, and Market Failure*, 85 CORNELL L. REV. 1, 8–9 (1999) [hereinafter Korobkin, *Managed Care*]. Second, he contends that cognitive limitations of consumers justify the inclusion of government-approved mandatory terms in standard form contracts. See Russell Korobkin, *Bounded Rationality, Standard Form Contracts, and Unconscionability*, 70 U. CHI. L. REV. 1203, 1243–44, 1255 (2003) [hereinafter Korobkin, *Standard Form Contracts*]. Thaler and Sunstein’s “libertarian paternalism,” which emphasizes using default rules to enhance the well-being of irrational persons, is an example of the softer forms of paternalism being advocated. See Richard H. Thaler & Cass R. Sunstein, *Libertarian Paternalism*, 93 AM. ECON. REV. 175, 176–78 (2003); see also Cass R. Sunstein & Richard H. Thaler, *Libertarian Paternalism Is Not an Oxymoron*, 70 U. CHI. L. REV. 1159, 1167, 1202 (2003). For a discussion of why libertarian paternalism functions as a “hard” form of paternalism for some persons, see Gregory Mitchell, *Libertarian Paternalism Is an Oxymoron*, 99 NW. U. L. REV. 1245, 1254 (2005). Rachlinski provides a good review of the numerous proposals advanced on irrationality grounds to restrict individual choice in the market and the law-making process. See Jeffrey J. Rachlinski, *The Uncertain Psychological Case for Paternalism*, 97 NW. U. L. REV. 1165, 1177–1206 (2003). We place hard and soft paternalism at opposite ends of a continuum measuring degree of choice censorship or choice constraint, with the intention of keeping blurry the point at which soft paternalism crosses over into hard paternalism. Our treatment of hard and soft-paternalism differs from typical treatments of the hard/soft paternalism distinction, such as that of Gerald Dworkin, who understands hard paternalism to be sometimes justified “even if the action [to be affected] is fully voluntary” and understands soft paternalism to be justified only when the “person for whom we are acting paternalistically is in some way not competent.” Gerald Dworkin, *Paternalism: Some Second Thoughts*, in PATERNALISM 105, 107 (Rolf Sartorius ed., 1983). Many behavioral law and economics scholars implicitly adopt a version of the traditional soft paternalism position that sees irrationality as a justification for intervention, on grounds that a rational person would hypothetically consent to paternalistic measures that counter irrational decisions: “According to [one] view, ‘respect autonomy’ is still a side constraint, but one which permits interference if and only if a

These proposals promise a “benevolent hierarchy”⁴ in which central planners substitute their judgments for those of impulsive, error-prone citizens,⁵ government agencies aggressively regulate how businesses speak to the masses to prevent commercial exploitation,⁶ judges regulate the content of standard form consumer contracts,⁷ and bureaucrats rather than jurors make decisions about punitive damages to restore coherence and fairness to the civil litigation system.⁸ Yet this renewed faith in better lives through paternalistic governance seems to ignore possible unanticipated effects of such intervention.⁹ Our intention here is to draw attention to one of the perverse effects likely to follow from enactment of many of the re-

choice is irrational and the chooser would consent to the interference if he were fully rational and well informed.” Danny Scoccia, *Paternalism and Respect for Autonomy*, 100 ETHICS 318, 318 (1990).

4. Wikler, *supra* note 1, at 384.

5. See, e.g., Sunstein & Thaler, *supra* note 3, at 1162 (“[W]e argue for self-conscious efforts, by private and public institutions, to steer people’s choices in directions that will improve the choosers’ own welfare.”).

6. See, e.g., Paul Horwitz, *Free Speech as Risk Analysis: Heuristics, Biases, and Institutions in the First Amendment*, 76 TEMP. L. REV. 1, 58 (2003) (arguing that advertisers’ pervasive efforts and expertise at taking advantage of consumers’ cognitive biases justifies greater deference to government regulation of commercial speech).

7. See, e.g., Korobkin, *Standard Form Contracts*, *supra* note 3, at 1207 (“By recognizing purchasers’ bounded rationality as the most important root cause of inefficiency in form contracts, courts can modify their use of unconscionability analysis to increase both social welfare generally and buyer welfare specifically.”).

8. See Cass R. Sunstein, *What Should Be Done?*, in CASS R. SUNSTEIN ET AL., PUNITIVE DAMAGES: HOW JURIES DECIDE 242, 252–55 (2002) (discussing removal of the power to award punitive damages from the jury and moving to a system of civil fines administered by government administrators).

9. Judge Kozinski argues that one of the key lessons learned from the collectivist experiments in Eastern Europe is that well-intentioned governmental initiatives often backfire. Alex Kozinski, *The Dark Lessons of Utopia*, 58 U. CHI. L. REV. 575, 592–93 (1991) (“[O]ur ability to predict the full effects of governmental actions—much less the synergistic effects of hundreds of thousands of simultaneous government interventions—is very limited. Far too often there are unanticipated results and costs, despite the most careful efforts of government officials.”). Kozinski’s view echoes Weber’s warning “that the final result of political activity often, nay, regularly, bears very little relation to the original intention: often, indeed it is quite the opposite of what was intended.” Max Weber, *Politics as a Vocation*, in PRINCETON READINGS IN POLITICAL THOUGHT 499, 501 (Mitchell Cohen & Nicole Fermon eds., 1996), cited in Frederick W. Preston & Roger I. Roots, *Introduction: Law and Its Unintended Consequences*, 47 AM. BEHAV. SCIENTIST 1371, 1371 (2004). For examples of many laws that have arguably had serious perverse effects, see the recent issue of the *American Behavioral Scientist* devoted to the unintended consequences of law, 47 AM. BEHAV. SCIENTIST 1371 (2004).

cent paternalistic proposals, namely, inhibition of the development of the regulated parties' decision-making skills. Our ideas have distinguished precursors. Most notably, John Stuart Mill argued that restraints on behavior should be limited to prevention of harm to others, because broader restraints may adversely affect the development of individuality.¹⁰ Alexis de Tocqueville expressed similar sentiments in a more applied context when he commented on the developmental benefits that accrue to American women, relative to European women, due to their increased liberty and exposure to risks.¹¹

The imposition of a paternalistic policy presupposes an individual will act contrary to her best interests unless some

10. JOHN STUART MILL, *On Liberty* (1859), in *ON LIBERTY AND OTHER ESSAYS* 1, 70 (John Gray ed., 1991).

11. ALEXIS DE TOCQUEVILLE, *Democracy in America* (1840), in *DEMOCRACY IN AMERICA AND TWO ESSAYS ON AMERICA* 1, 684 (Gerald E. Bevan trans., 2003) ("Long before the young American woman has reached marriageable age, her emancipation from her mother's supervision has gradually started . . . [T]he great scene of society lies constantly exhibited for her to see; far from attempting to conceal this sight from her, she is daily shown more and more of it and is taught to contemplate it with a steady and calm gaze."). For a modern updating of this view, see Peter Glick & Susan T. Fiske, *An Ambivalent Alliance: Hostile and Benevolent Sexism as Complementary Justifications for Gender Inequality*, 56 *AM. PSYCHOL.* 109, 116 (2001) (noting that "some forms of sexism are, for the perpetrator, subjectively benevolent, characterizing women as pure creatures who ought to be protected, supported, and adored and whose love is necessary to make a man complete"). Within legal scholarship, Winick has consistently expressed a Millian concern about paternalism's effects on individual development and health. See Bruce J. Winick, *Coercion and Mental Health Treatment*, 74 *DENV. U. L. REV.* 1145, 1160 (1997); Bruce J. Winick, *On Autonomy: Legal and Psychological Perspectives*, 37 *VILL. L. REV.* 1705, 1764 (1992). Also, in an earlier paper one of the present authors raised the moral hazard concern that is developed more fully here. See Jonathan Klick, *The Microfoundations of Standard Form Contracts: Price Discrimination vs. Behavioral Bias*, 32 *FLA. ST. U. L. REV.* 555, 569 (2005).

With respect specifically to governance proposals generated through behavioral economics, to date scholars have primarily warned about the perverse effect that may arise from eliminating one cognitive bias that serves to temper another cognitive bias, that is, the possible upsetting of offsetting biases if government attempts to counter biases one by one without considering the interactive effects of biases. See, e.g., Jonathan Baron, *Cognitive Biases, Cognitive Limits, and Risk Communication*, 23 *J. PUB. POLY & MARKETING* 7, 11 (2004) ("[B]iases may work together to restore a kind of artificial equilibrium that works for normal situations, so that correcting one bias without correcting another one can make things worse."); see also Gregory Besharov, *Second-Best Considerations in Correcting Cognitive Biases*, 71 *SO. ECON. J.* 12, 19 (2004). We agree that this potential perverse effect is a serious concern, but it is not the only unintended consequence that may arise from government regulation of irrationality.

third party intervenes to protect those interests.¹² Such intervention may be justified on grounds that the paternalism advances efficiency, personal integrity, or sound judgment.¹³ For simplicity's sake we focus here on the goal of the new paternalism to correct inefficiencies associated with systematic psychological biases in the formation of beliefs and expression of preferences,¹⁴ but our analysis also has implications for the personal-integrity and sound-judgment rationales for paternalism.¹⁵

We question the generality of the claim that short-run inef-

12. See Gerald Dworkin, *Paternalism*, in MORALITY AND THE LAW 107, 108 (Richard A. Wasserstrom ed., 1971) (characterizing paternalism as “the interference with a person’s liberty of action justified by reasons referring exclusively to the welfare, good, happiness, needs, interests or values of the person being coerced”); Anthony T. Kronman, *Paternalism and the Law of Contracts*, 92 YALE L.J. 763, 763 (1983) (“In general, any legal rule that prohibits an action on the ground that it would be contrary to the actor’s own welfare is paternalistic.”). We recognize the difficulties associated with defining paternalism, see, e.g., Seana Valentine Shiffrin, *Paternalism, Unconscionability Doctrine, and Accommodation*, 29 PHIL. & PUB. AFF. 205, 211 (2000) (“[T]he literature on paternalism contains a variety of explicit and implicit conceptions of paternalism.”), and we recognize that some may claim that government regulation of irrationality is not paternalistic. Nevertheless, we are interested in whether government regulation of irrationality reliably achieves the goal of efficiency, whether that goal can rightly be labeled paternalistic or not.

13. See Kronman, *supra* note 12, at 765 (noting that some paternalistic regulations “are best explained by considerations of economic efficiency and distributive fairness, others by the idea of personal integrity, and a third set of limitations by the . . . notion of sound judgment”).

14. See, e.g., Dale Carpenter, *The Antipaternalism Principle in the First Amendment*, 37 CREIGHTON L. REV. 579, 645 (2004) (“[U]nder some circumstances paternalism can be efficient, especially where people’s preferences are irrational. Still others, pointing to the limits of individual human reason and the frequency of market failure, see broad areas where it is legitimate” (footnotes omitted)).

15. For instance, the notion behind the personal-integrity justification is that paternalistic intervention constitutes no invasion of personal integrity when it disrupts lowly ranked concerns to protect highly ranked concerns. See JOHN KLEINIG, PATERNALISM 68 (1984). But if paternalism interferes with the life-long developmental processes needed to distinguish lowly ranked from highly ranked values and then to pursue higher values, then the personal-integrity justification loses some of its force. The difficulty, of course, is distinguishing between those freedoms that are vital to developing personal integrity and those that will impede development. Perhaps the best case can be made for restricting minors’ access to addictive goods that pose serious health risks. Lowenstein, for instance, argues that smoking has developmental costs that surely outweigh the benefits of protecting cigarette advertisements aimed at children. Daniel Hays Lowenstein, “*Too Much Puff*”: *Persuasion, Paternalism, and Commercial Speech*, 56 U. CIN. L. REV. 1205, 1212–13 (1988).

iciencies associated with psychological biases justify paternalistic government regulations.¹⁶ In particular, we argue that there will often be long-run costs of paternalistic regulations that offset short-run gains because of the negative learning and motivational effects of paternalistic regulations. An appreciation of the role of learning and motivation in the development of rational behavior, and the necessary concomitant that individuals differ in their propensities to act rationally, suggests two broad limitations on the force and scope of irrationality-based arguments for paternalism. First, individual and situ-

16. For instance, Korobkin expressly argues that psychological biases lead to inefficiencies that can be corrected with paternalistic interventions. See Korobkin, *Managed Care*, *supra* note 3, at 88; Korobkin, *Standard Form Contracts*, *supra* note 3, at 1294 (“The design of non-salient terms is better assigned to government institutions because the market will not create pressure toward efficiency and state actors, as imperfect as they will be, at least can aim at the proper target.”).

Within legal scholarship, Zamir argues most generally that paternalistic legal rules will be economically efficient if the citizenry is assumed to be boundedly rational. See Eyal Zamir, *The Efficiency of Paternalism*, 84 VA. L. REV. 229, 252 (1998) (“Once the prevalence of systematic deviations from the rational-maximizer model is acknowledged, principled antipaternalism is no longer a tenable position of economic analysis.”). The primary difference between Zamir’s model and our analysis is that we effectively endogenize the magnitude of the cognitive bias under which an individual makes her decisions. That is, while Zamir assumes that the likelihood of an individual choosing correctly is given, we explicitly model the individual’s choice of how much cognitive effort to expend and that effort in turn determines the individual’s likelihood of choosing correctly. In terms of evaluating the ultimate welfare implications of a particular paternalistic intervention, our model implies that the relevant comparison does not just involve comparing which decision maker (individual or paternalist) is more likely to choose correctly as in Zamir’s model; it also involves comparing the cost of improving an individual’s likelihood of choosing correctly with the cost of administering the paternalistic intervention. In addition, whereas Zamir recognizes but discounts the possible long-term effects of paternalism, largely on grounds that cognitive biases are ubiquitous and persistent, *see id.* at 276–77, we believe that growing evidence of the situation- and person-dependent nature of rationality errors and the realization that the heuristic and bias research is not as robust as previously thought, counsel greater concern for the developmental and incentive effects of paternalism. See, e.g., Mandeep K. Dhami et al., *The Role of Representative Design in an Ecological Approach to Cognition*, 130 PSYCHOL. BULL. 959, 976 (2004) (“[R]esearch in the heuristics-and-biases program involves carefully setting up conditions that produce cognitive biases. The extent to which these findings generalize to conditions outside the laboratory is unclear.” (citations omitted)); David R. Shanks et al., *A Re-examination of Probability Matching and Rational Choice*, 15 J. BEHAV. DECISION MAKING 233, 248 (2002) (noting that probability matching “is heavily context-dependent and . . . can be made to disappear under appropriate conditions of task structure, training, motivation, and feedback”).

ational variation in irrational tendencies will often make debiasing interventions, or no intervention at all, more efficient than paternalistic interventions. Second, paternalistic interventions may exacerbate irrational tendencies by creating moral and cognitive hazards. Moral hazards arise because paternalistic regulations reduce an individual's motivation to act deliberately and carefully, and motivation level mediates many psychological biases. What we term "cognitive hazards" arise when paternalistic regulations interfere with information searches, educational investments, and feedback that would occur in the absence of paternalistic interventions and that are important to the individual's development of effective decision-making skills and strategies.¹⁷

Our cautionary argument regarding paternalistic interventions follows from psychological research on judgment and decision making and economic modeling of decision-making behavior under a paternalistic regime. In Part I, we provide a theoretical and empirical justification for the moral and cognitive hazards of paternalism utilizing two bodies of psychological research that have been largely ignored by behavioral law and economics scholars. First, research from developmental psychology indicates that individuals improve their decision-making skills over time through a "learning by doing" process, and that paternalistic policies threaten interference in this self-regulatory process.¹⁸ Second, research on self-fulfilling prophe-

17. Within judgment and decision-making studies it is often difficult to separate motivational from cognitive determinants of behavior, *see, e.g.*, Norbert Schwarz, *Social Judgment and Attitudes: Warmer, More Social, and Less Conscious*, 30 EUR. J. SOC. PSYCHOL. 149, 159–60 (2000), but generally motivated processing is conceived of as goal-directed thought aimed at protecting one's self-image or existing beliefs or achieving accuracy, whereas cognitive processing simply refers to the operation of information-processing mechanisms without any particular directional or self-serving goal presumed. *See, e.g.*, Ziva Kunda, *The Case for Motivated Reasoning*, 108 PSYCHOL. BULL. 480, 495 (1990) ("Although the mechanisms underlying motivated reasoning are not yet fully understood, it is now clear that directional goals do affect reasoning."). For our purposes, we need only distinguish between hazards associated with changes in motives to engage in effortful, analytic thought in the present which is more likely to lead to rational behavior (i.e., moral hazards), and hazards associated with changes in the amount of information learned (i.e., cognitive hazards). Both hazards may result from paternalistic intervention and cause long-run inefficiencies, although through different mechanisms. For discussion of the relation between cognitive and moral hazard in economic terms, *see infra* note 79.

18. *Cf.* Rachlinski, *supra* note 3, at 1214 ("The role of individual learning and adaptation . . . cannot be ignored in assessing the need for paternalism. Simple experience might, in some contexts, be a much better cure for cognitive

cies warns that regulated parties are likely to become the weak decision makers envisioned by paternalistic policy makers, as paternalistic regulations undercut personal incentives to invest in cognitive capital and the regulated parties conform to the expectancies of the paternalist.¹⁹

In Part II, we develop economic models of behavior under paternalism that further support the view that paternalism may lead to suboptimal long-run behavior. These models specify when paternalistic accommodation of irrational tendencies is warranted, when education or another debiasing approach to irrational tendencies is warranted, and when no governmental action is warranted. In Part III, using insights from our economic models of behavior under paternalism, we discuss the factors that should be considered when designing paternalistic interventions in order to limit the cognitive and moral hazards of paternalism. We also note the woefully inadequate state of empirical knowledge relevant to these factors and hence the great likelihood that many paternalistic interventions are suboptimal.

I. JUDGMENT AND DECISION MAKING UNDER PATERNALISM

A. A DEVELOPMENTAL PERSPECTIVE ON DECISION-MAKING COMPETENCE

Contrary to the static approaches to judgment and decision making that underlie most behavioral law and economics understandings of irrationality and concomitant calls for paternalism to counter irrational behaviors,²⁰ we consider how pa-

missteps than adopting a paternalist intervention.”).

19. Cf. Jon Elster, *Selfishness and Altruism*, in BEYOND SELF INTEREST 44, 47 (Jane J. Mansbridge ed., 1990) (“[T]he opportunity to choose—including the right to make the wrong choices—is a valuable, in fact, indispensable, means to self-improvement.”); Adam J. Hirsch, *Spendthrift Trusts and Public Policy: Economic and Cognitive Perspectives*, 73 WASH. U. L.Q. 1, 52 (1995) (“Many psychologists (and historians too) have noticed a tendency for individuals and groups to take on the characteristics that others, particularly those in positions of authority, ascribe to them—what is known in the psychological literature as a ‘self-fulfilling prophecy.’”).

20. Behavioral law and economics scholars often catalog the many psychological biases that have been demonstrated by psychologists and behavioral economists, typically in laboratory experiments or through classroom surveys using word problems, and then use this assortment of biases to justify a particular approach to legal regulation. See, e.g., Hanson & Kysar, *supra* note 2, at 633–34; Jolls et al., *supra* note 2, at 1476–77; Korobkin & Ulen, *supra* note

ternalistic regulations may affect cognitive behavior over time both inside and outside the regulated domain.²¹ Before introducing our economic models of behavior under paternalism, we introduce the psychological framework and empirical findings that motivate these models. Of particular importance are Byrnes's self-regulation model of decision making²² and evidence for the role of incentives and personal motivation as mediators of rational action.

2, at 1058. These menu or snapshot approaches to judgment and decision making focus on how a particular legal judgment or decision might fall prey to one or more biases. See, e.g., Korobkin & Ulen, *supra* note 2, at 1096–97. Behavioral law and economics' static approach to judgment and decision making is largely a function of behavioral decision theory's lack of an integrative theory. *Id.* at 1057. The dominant research program within behavioral decision theory, the heuristics and biases program, consists of a collection of robust empirical findings bound together by high-level concepts rather than an integrative theory that can predict how particular features of the mind and environment are likely to interact in particular cases (e.g., the heuristics and biases program predicts that accessible features of the environment and memory will exert inordinate influence on judgments, but it lacks a theory of accessibility). See, e.g., Michael R.P. Dougherty et al., *Memory as a Fundamental Heuristic for Decision Making*, in EMERGING PERSPECTIVES ON JUDGMENT AND DECISION RESEARCH 125, 128 (Sandra L. Schneider & James Shanteau eds., 2003); Daniel Kahneman, *A Perspective on Judgment and Choice: Mapping Bounded Rationality*, 58 AM. PSYCHOL. 697, 702 (2003) (“[M]uch is known about the determinants of accessibility, but there is no general theoretical account of accessibility and no prospect of one emerging soon.”). For an internal critique of psychology's emphasis on demonstrating judgment and decision-making shortcomings, see Joachim I. Krueger & David C. Funder, *Towards a Balanced Social Psychology: Causes, Consequences and Cures for the Problem-Seeking Approach to Social Behavior and Cognition*, 27 BEHAV. & BRAIN SCI. 313 (2004).

21. Brehmer notes that research limiting itself to “snapshots of judgmental processes” is

an important line of research, for there are undoubtedly situations in which momentary accuracy is important. However, to an organism behaving in time, momentary accuracy may be less important. It is only necessary that the momentary level of achievement is sufficient to point the organism in the right direction, for there are always possibilities for later corrections. It does not seem unlikely that such “cognition over time” is the natural form of cognition. If so, the levels of achievement that are found in “snapshot studies” of judgment and decision making are neither surprising nor alarming. It certainly gives an answer to the puzzle of how organisms have survived despite their seemingly inefficient cognitive equipment.

Berndt Brehmer, *Man as a Stabiliser of Systems: From Static Snapshots of Judgment Processes to Dynamic Decision Making*, 2 THINKING & REASONING 225, 226 (1996) (citations omitted).

22. JAMES P. BYRNES, THE NATURE AND DEVELOPMENT OF DECISION MAKING: A SELF-REGULATION MODEL (1998) (proposing and introducing the self-regulation model of decision making).

Byrnes's self-regulation model assumes that "[t]he key to being successful in life is knowing the difference between options that are likely to produce favorable outcomes and options that are unlikely to produce favorable outcomes."²³ Through education, experimentation, experience, and observation, individuals learn which options are most likely to produce desirable outcomes and develop competence in the ability to compile and rank-order options, and then select the option that will lead to the most favorable outcome.²⁴ The main vehicle to greater decision-making competence is alteration in existing psychological states such that later psychological states possess more reliable knowledge about what ends are most valued and how best to achieve those ends.²⁵ Outcome feedback and verbal feedback serve as the main mechanisms for change between earlier and later psychological states.²⁶

23. James P. Byrnes, *The Development of Decision-Making*, 31 J. ADOLESCENT HEALTH 208, 208 (2002); see also James P. Byrnes et al., *Learning to Make Good Decisions: A Self-Regulation Perspective*, 70 CHILD DEV. 1121, 1121 (1999) ("[T]here should be a close correspondence between effective decision-making and personal success."). Hence, Byrnes effectively employs a means-end or instrumental account of rationality that is consistent with weak microeconomic conceptions of rationality and efficiency.

24. Byrnes breaks the decision process down into four steps: (1) setting a goal, (2) compiling options to achieve the goal, (3) rank-ordering the options, and (4) selecting the highest-ranked option. Byrnes et al., *supra* note 22, at 1121. Direct involvement in a task is not the sole route to learning; observing others perform the task may lead to one's own improvement on the task or lead to adaptive avoidance of the task altogether (e.g., observing other day traders fail may wisely lead one to avoid day trading entirely). See Nigel Harvey & Ilan Fischer, *Development of Experience-Based Judgment and Decision Making: The Role of Outcome Feedback*, in THE ROUTINES OF DECISION MAKING 119, 134–35 (Tilman Betsch & Susanne Haberstroh eds., 2005).

25. See Byrnes et al., *supra* note 23, at 1122–23.

26. See *id.* at 1122. Outcome feedback refers to success or failure of a chosen option; verbal feedback refers to "commentary or advice given before or after choices are made." *Id.* The relatively few studies that have empirically examined the role of feedback in adult decision making suggest that adults can "progressively learn to make better decisions if they receive[] relatively clear feedback from outcomes." *Id.* at 1125 (citations omitted). Although verbal, or cognitive, feedback has often been shown to lead to greater learning than outcome feedback, Byrnes, Miller, and Reynolds report empirical results showing that "outcome feedback was more effective than verbal feedback" to effect positive change in choice behavior. Byrnes et al., *supra* note 23, at 1137. Compare William K. Balzer et al., *Effects of Cognitive Feedback on Performance*, 106 PSYCHOL. BULL. 410, 410 (1989) ("In contrast to [outcome feedback], [cognitive feedback] has been found to improve the accuracy of judgments in many circumstances.") with Byrnes et al., *supra* note 23, at 1137. For a review of the impact of outcome feedback on various types of judgments and decisions, see generally Harvey & Fischer, *supra* note 24.

The self-regulation model is not exclusively behaviorist in its orientation, however, as it posits an “endogenous tendency toward self-regulation” that keeps one oriented “toward increased accuracy in the face of occasional instances of success and failure that could lead it astray” (i.e., a model that controls for random disturbances in the system).²⁷ This self-regulation tendency, which is presumed to operate through the long-term memory system, leads to conservatism in the changes made between earlier and later psychological states that are relevant to decision-making success.²⁸ Thus, individual instances of success or failure, unless accompanied by verbal feedback, may result in few changes in psychological states due to this conservatism, but repeated instances of outcome feedback within a particular domain or with a particular task or goal that cuts across domains are likely to lead to changes in psychological states.²⁹

27. Byrnes et al., *supra* note 23, at 1122.

28. *See id.* at 1122–23 (“Conservative belief change is highly adaptive in an uncertain and variable environment.”). Gibson and colleagues offer a parallel process model of learning in dynamic decision environments that can explain gradual learning and generalization from feedback. *See* Faison P. Gibson et al., *Learning in Dynamic Decision Tasks: Computational Model and Empirical Evidence*, 71 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 1 (1997). One can also view the learning process as a reinforcement model, which also leads to gradual change. *See* Eric Von Magnus, *Preference, Rationality, and Risk Taking*, 94 ETHICS 637, 639–40 (1984).

29. Support for this conclusion is found in laboratory studies of markets, where choices in experimental repetitive markets often converge toward the rational choice equilibrium. *See* Vernon L. Smith, *Economics in the Laboratory*, 8 J. ECON. PERSP. 113, 118 (1994) (noting a “tendency for rational behavior to emerge in the context of a repetitive market institution” and that “[i]n many experimental markets, . . . human agents interact through the trading rules to produce social algorithms which demonstrably approximate the wealth maximizing outcomes traditionally thought to require complete information and cognitively rational actors” (footnote omitted)); *see also* Colin Camerer et al., *The Curse of Knowledge in Economic Settings: An Experimental Analysis*, 97 J. POL. ECON. 1232, 1242 (1989) (noting, with respect to the “curse of knowledge,” or the negative effects of asymmetric information, that “[m]arket experience clearly reduces bias more than individual judgment tempered by incentives and feedback”). In addition, Gervais and Odean provide a dynamic model of overconfidence in trading that supports the self-regulation model of decision making:

When a trader is successful, he attributes too much of his success to his own ability and revises his beliefs about his ability upward too much. In our model overconfidence is dynamic, changing with successes and failures. Average levels of overconfidence are greatest in those who have been trading for a short time. With more experience, people develop better self-assessments.

Simon Gervais & Terrance Odean, *Learning to Be Overconfident*, 14 REV. FIN.

It is also important to emphasize that the self-regulation model, because it incorporates cognitive components, allows for both domain-specific and domain-general learning, though the presumption is that domain-specific learning is more common.³⁰ Learning may thus take the form of increases in the accuracy of one's beliefs in a particular domain or development of domain-general theories about the relationship of higher-order goals and procedures for achieving those goals.³¹ For instance, developing effective self-control techniques in order to save for an automobile or home may generalize to effective strategies for retirement saving. Or, as demonstrated by empirical research on the endowment effect, people may learn to overcome consumer biases with greater market experience, and this learning may generalize across goods.³² Calibration of beliefs may involve improved self-knowledge, particularly with respect to likes and dislikes and consumption tendencies, and not sim-

STUD. 1, 19 (2001). Gervais and Odean also note that empirical data are consistent with this dynamic model of overconfident trading. *See id.* (“[Y]ounger investors trade more actively than older investors while earning lower returns relative to a buy-and-hold portfolio. These results are consistent with our prediction that overconfidence diminishes with greater experience.”). Bjorklund argues more generally that immaturity in one's metacognition (i.e., knowledge about the workings of one's own mind) has positive effects on physical and cognitive development, because such immaturity leads to exploratory behavior and learning that would not occur in a person with a more mature metacognitive state and a better-calibrated confidence level with respect to her abilities. *See* David F. Bjorklund, *The Role of Immaturity in Human Development*, 122 PSYCHOL. BULL. 153, 163–66 (1997). Thus, overconfidence provides adaptive benefits that, in the long run, outweigh any initial, short-run costs.

30. *See* Byrnes et al., *supra* note 23, at 1124. Thus, learning should occur more quickly in repetitive decision settings, such as with many consumer goods, but learning may also generalize across goods and decision settings.

31. The precise ways in which domain-specific and domain-general knowledge structures change are not important to our treatment. Dougherty and his colleagues offer an account of how experience and domain knowledge may progressively lead to reduced error. *See* Dougherty et al., *supra* note 20, at 144–51. Note, however, that much of the learning that takes place will be in the form of “tacit knowledge” rather than “academic intelligence.” “Tacit knowledge refers to general and domain specific skills and abilities acquired over time. It is possible that an individual may absorb these skills and abilities through experience rather than through formal education and training.” James E. Hunton & Ruth Ann McEwen, *An Assessment of the Relation Between Analysts' Earning Forecast Accuracy, Motivational Incentives and Cognitive Information Search Strategy*, 72 ACCT. REV. 497, 514 (1997) (citations omitted).

32. *See* John A. List, *Neoclassical Theory Versus Prospect Theory: Evidence from the Marketplace*, 72 ECONOMETRICA 615, 624 (2004).

ply knowledge about causal relations in the external world.³³ In addition, learning will take the form of discovering when external resources should be recruited to overcome internal resource limitations, such as knowing when to consult experts or gain explicit education in a domain.³⁴

The self-regulation model has obvious implications for hard-paternalism proposals that restrict choice and opportunities: feedback and learning cannot occur if “institutional and other structures essentially rule out the possibility of experiencing feedback that might be contrary to one’s beliefs.”³⁵ To the extent that government is more likely to intervene paternalistically on important choices, paternalistic constraints on learning are likely to be significant in such cases because more important decisions elicit more effortful evaluation processing than less important decisions.³⁶

Softer forms of paternalism may also adversely affect learning by altering the individual’s representation of the choice setting and how she encodes feedback about success or failure in a given situation.³⁷ The individual does not exert the same level of control in compiling and assessing options in the presence of soft paternalism, and verbal feedback about even a soft paternalistic situation should cause the individual to discount her own role in achieving a particular outcome.³⁸

33. See Byrnes, *supra* note 23, at 210 (“[T]here is an important, but relatively ignored, aspect of knowledge that might prove to produce consistent age differences in choices: self-knowledge.”).

34. See *id.* at 209 (“[C]ompetent decision-makers use strategies to overcome obstacles that might hinder the discovery process. For example, they might seek advice from knowledgeable people when they are not sure how to proceed”); see also Rachlinski, *supra* note 3, at 1219 (“Often, even if people employ a suboptimal strategy and cannot adapt, they can recognize their own limitations and hire others to help them make decisions.”).

35. Byrnes et al., *supra* note 23, at 1123 (citation omitted).

36. Byrnes, *supra* note 23, at 211 (“Studies show that adults are likely to alternate between effortful and less effortful evaluation strategies depending of [sic] the importance of a decision.” (citation omitted)).

37. In addition, some seemingly soft forms of paternalism may effectively operate as hard forms of paternalism. See Mitchell, *supra* note 3, at 1248 (noting that “libertarian paternalism is an oxymoron”).

38. It is possible, however, that a self-serving attributional bias will inhibit the discounting effect that should occur with respect to positive feedback in a soft-paternalism situation. See Amy H. Mezulis et al., *Is There a Universal Positivity Bias in Attributions? A Meta-Analytic Review of Individual, Developmental, and Cultural Differences in the Self-Serving Attributional Bias*, 130 PSYCHOL. BULL. 711, 738 (2004) (“We found strong support for the existence of a robust self-serving bias in attributions.”).

From this consideration of self-regulatory processes in decision making, three propositions may be extracted: (1) paternalistic policies that restrict choice options restrict learning opportunities; (2) the noisier the learning environment, the more difficult to learn, and paternalistic policies introduce noise into, or mute feedback signals in, the learning environment;³⁹ (3) the more extensive the paternalism imposed on citizens, the greater the cognitive hazard, due to restricted learning opportunities and more noise in learning environments.

This dynamic approach to decision making and the effects of paternalism finds further support in the renewed appreciation of the role that incentives play in the rationality of judgment and decision making. Contrary to some suggestions otherwise within behavioral law and economics, material incentives do improve the quality of choice under certain conditions:

Incentives improve performance in easy tasks that are effort responsive, like judgment, prediction, problem-solving, recalling items from memory, or clerical tasks. Incentives sometimes hurt when problems are too difficult or when simple intuition or habit provides an optimal answer and thinking harder makes things worse. In games, auctions, and risky choices the most typical result is that incentives do not affect mean performance, but incentives often reduce variance in responses. In situations where there is no clear standard of performance, incentives often cause subjects to move away from favorable “self-presentation” behavior toward more realistic choices.⁴⁰

39. As the feedback structure becomes more complex, individuals have a harder time making effective use of feedback. See Brehmer, *supra* note 21, at 234–36. However, to the extent that feedback about task information, or information about the relation between environmental cues and the true state of the object or criterion to be judged, is available, some learning may occur even in more complex environments. See William K. Balzer et al., *Effects of Cognitive Feedback Components, Display Format, and Elaboration on Performance*, 58 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 369, 382 (1994) (“[I]t is the [task information] component of [cognitive feedback] (i.e., telling judges the correct strategy for weighting and integrating information to optimize their judgments) that is essential for improving judgment performance.”).

40. Colin F. Camerer & Robin M. Hogarth, *The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework*, 19 J. RISK & UNCERTAINTY 7, 34 (1999); see also Vernon L. Smith, *Method in Experiment: Rhetoric and Reality*, 5 EXPERIMENTAL ECON. 91, 101–02 (2002) (“Anyone who doubts that payoffs can and do matter has not looked at the evidence. What is not predictable by any theory is what situations will be sensitive at what payoff levels and what situations will not be sensitive at the levels commonly used.”). In addition, cognitive behavior under conditions of low material payoffs or hypothetical payoffs may differ significantly from behavior in response to economic situations with significant real consequences, such that it becomes risky to assume bias in low-cost situations will manifest itself

Incentives play an important role in the quality of judgment and choice even if we take the most restrictive view of the positive effects of incentives—namely, that incentives move behavior toward the rational response only in cases where a dominant or clear normative response exists or where irrational behavior occurs due to lack of attention or interest.⁴¹ Incentives remain important under even this restrictive view because several psychological biases arise from inattention or insufficient motivation to engage in information search, and thus the incidence and severity of these biases are conditional on incentive levels.⁴²

identically in high-cost situations. See Charles A. Holt & Susan K. Laury, *Varying the Scale of Financial Incentives Under Real and Hypothetical Conditions*, 24 BEHAV. & BRAIN SCI. 417, 418 (2001); see also James F. Smith & Thomas Kida, *Heuristics and Biases: Expertise and Task Realism in Auditing*, 109 PSYCHOL. BULL. 472, 486 (1991) (“For many audit judgments, the costs associated with certain risks are sufficiently large that they seem to significantly influence the nature of audit training and formalized audit procedures.”).

41. See Dan N. Stone & David A. Ziebart, *A Model of Financial Incentive Effects in Decision Making*, 61 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 250, 259 (1995) (“[I]ncentives appear to increase the extent of attention given to a task, but also to increase potentially distracting emotions . . .”). A slightly less-restrictive view is that for some biases that involve more than lack of interest or inattentiveness incentives will be effective only for those people who possess the cognitive capacity or ability to compute the rational solution. See, e.g., Vidya Awasthi & Jamie Pratt, *The Effects of Monetary Incentives on Effort and Decision Performance: The Role of Cognitive Characteristics*, 65 ACCT. REV. 797, 808 (1990) (reporting that monetary incentives improved the performance of subjects with higher perceptual differentiation ability (i.e., the ability to abstract familiar concepts or relationships from a complex setting) with respect to application of the conjunction probability and sample size rules). An even less restrictive view of incentives would argue for the positive effects of incentives on decision avoidance, use of decision aids, and recruitment of expert assistance. This view remains untested within judgment and decision-making research because this research rarely presents these options to subjects, but instead examines how subjects’ own, unassisted performance on rational-thinking tests is affected by incentives.

42. Frey and Eichenberger note that positive incentive effects have been found with respect to numerous anomalies, including preference reversals, the Allais paradox, the certainty effect, ambiguity aversion, deviations between willingness-to-pay prices and willingness-to-accept prices, base rate neglect, overoptimism, anchoring, the hindsight bias, temporal inconsistencies, and framing effects. See Bruno S. Frey & Reiner Eichenberger, *Economic Incentives Transform Psychological Anomalies*, 23 J. ECON. BEHAV. & ORG. 215, 225 n.16 (1994); see also Ralph Hertwig & Andreas Ortmann, *Experimental Practices in Economics: A Methodological Challenge for Psychologists?*, 24 BEHAV. & BRAIN SCI. 383, 391–96 (2001) (surveying a variety of studies in which incentives positively affected performance). Engelmann and Strobel, in a study of the false consensus effect, show how incentives can improve information

In addition to direct material incentives, holding people accountable for their judgments and decisions can likewise move behavior toward the rational norm.⁴³ “Predecisional accountability to an unknown audience will attenuate biases that arise from lack of self-critical attention to one’s decision processes and failure to use all relevant cues.”⁴⁴ Thus, expecting to have to account for a choice may have positive effects on decision-making quality.⁴⁵

A direct linkage exists between incentives and self-regulation: incentives often motivate a decision maker to invest cognitive effort and other resources to achieve a goal, with the positive by-product of increased procedural and self-knowledge *even in cases of outcome failure*. That is, the self-regulation

search and belief formation. See Dirk Engelmann & Martin Strobel, *The False Consensus Effect Disappears if Representative Information and Monetary Incentives Are Given*, 3 EXPERIMENTAL ECON. 241, 253 (2000) (“[G]iven both these incentives and representative information, although subjects show a consensus effect, they show no false consensus effect.”). Based on these results, Engelmann and Strobel note that “the false consensus effect might not be very relevant for economic applications.” *Id.* Incentive effects may also be understood within a dual-process model of cognition: to the extent incentives encourage shifts from intuitive to deliberative thinking, those biases that can be reduced by more effortful, deliberative thought should be positively affected by increasing incentives. Cf. Keith E. Stanovich & Richard F. West, *Who Uses Base Rates and P(D/~H)? An Analysis of Individual Differences*, 26 MEMORY & COGNITION 161, 166–69, 171, 175–76 (1998) (reporting significant correlations between propensity to engage in more deliberative thought and positive performance tests of cognitive ability, deductive reasoning ability, and inductive reasoning ability). Combining market forces with incentives provides even more potent debiasing force. See, e.g., David M. Grether, *Individual Behavior and Market Performance*, 76 AM. J. AGRIC. ECON. 1079, 1079–82 (1994) (discussing the role of incentives and markets in reducing the incidence and effects of the representativeness heuristic, preference reversals, and the endowment effect).

43. The particular contours of the accountability constraint are important: Self-critical and effortful thinking is most likely to be activated when decision makers learn prior to forming any opinions that they will be accountable to an audience (a) whose views are unknown, (b) who is interested in accuracy, (c) who is interested in processes rather than specific outcomes, (d) who is reasonably well-informed, and (e) who has a legitimate reason for inquiring into the reasons behind participants’ judgments.

Jennifer S. Lerner & Philip E. Tetlock, *Accounting for the Effects of Accountability*, 125 PSYCHOL. BULL. 255, 259 (1999).

44. *Id.* at 265.

45. Accountability may, however, exacerbate bias “to the extent that (a) a given judgment bias results from using normatively (but not obviously) proscribed information or (b) a given choice bias results from the fact that the option appears easiest to justify also happens to be the biased option.” *Id.*

model of decision making emphasizes that the decision maker may benefit from even bad or unlucky choices. Indeed, learning may be greatest in response to negative or unfavorable outcomes.⁴⁶ Thus, removing incentives to make good decisions may negatively impact activity levels and the amount of cognitive resources invested in activities, causing reductions in the amount and kinds of procedural- and self-knowledge gained.

In considering the motivational effects of paternalism, it is useful to distinguish between paternalism imposed before and after a choice is made (i.e., *ex ante* versus *ex post* paternalism).⁴⁷ *Ex ante* paternalism reduces the incentive to search for information, carefully evaluate decision options, or develop good decision-making strategies. *Ex post* paternalism reduces the risk of thoughtless action, because the government will insulate the decision maker from the consequences of the thoughtless choice. Thus, *ex post* paternalism operates as a form of social insurance for irrational behavior.⁴⁸

46. See, e.g., Peter H. Ditto et al., *Motivated Sensitivity to Preference-Inconsistent Information*, 75 J. PERSONALITY & SOC. PSYCHOL. 53, 64 (1998) (reporting the results of three studies showing that preference-inconsistent feedback motivated more effortful processing and a greater sensitivity to information quality than preference-consistent feedback); *id.* at 54 (“A large body of research in social cognition suggests that negative information and negative affective states produce more systematic, detail-oriented cognitive processing than positive information and positive affective states.” (citations omitted)); Dan Zakay et al., *Outcome Value and Early Warning Indications as Determinants of Willingness to Learn from Experience*, 51 EXPERIMENTAL PSYCHOL. 150, 155 (2004).

47. *Ex ante* paternalism eliminates or reduces contractual freedom before a transaction occurs, including the imposition of cooling-off periods and information-disclosure requirements. Bankruptcy protection and unconscionability challenges to contractual validity are the classic forms of *ex post* paternalism (recognizing, of course, that nonpaternalistic justifications may also be offered for both bankruptcy and the unconscionability doctrine), but the behavioral law and economics literature has given rise to new claims for judicial relief. See *Honorable v. Easy Life Real Estate Sys.*, 100 F. Supp. 2d 885, 888 (N.D. Ill. 2000) (noting that plaintiffs filed a claim for race discrimination in housing sales that relied on a theory of market manipulation via exploitation of psychological bias as advanced by Hanson and Kysar).

48. Under some accounts of paternalism, if a party seeks state assistance for relief from a contract, as through an assertion of the unconscionability doctrine, then a court’s invalidation of the contract would not be a paternalistic act. See Shiffrin, *supra* note 12, at 210–11. Although it is not necessary to believe that an act of paternalism must be against the will of the assisted party, we need not resolve this issue here (for a discussion of the various paternalism issues raised by the unconscionability doctrine, see Shiffrin, *supra* note 12). For our purposes, *ex post* paternalism is simply shorthand for government assistance available to protect a party from an earlier, supposedly irrational act.

We may therefore add two propositions that reflect the possibility that decision competence will be endogenous to the incentives created by a paternalistic policy: (4) *ex ante* paternalism provides a negative incentive to invest in cognitive capital and exert cognitive effort, which may have adverse effects both inside and outside the regulated domain; (5) *ex post* paternalism provides a positive incentive to reduce cognitive effort and care in many domains. These propositions emphasize that paternalistic policies, on the margin, reduce an individual's incentive to cultivate her cognitive capacity. That is, while a social planner may still determine that paternalism raises social welfare by some measure, one of the costs that must be considered is that biases themselves may be worsened or prolonged by the paternalistic policies.

To be clear, we do not contend that Byrnes's self-regulation model removes all concern about individual instances of poor judgment or choice,⁴⁹ or that the self-regulation model provides a compelling argument against all paternalistic proposals. Rather, this empirically derived model highlights key processes in the development of decision-making competence that may be adversely affected by paternalistic policies, and it directs attention away from a static or piecemeal approach to judgment and decision making.⁵⁰ Byrnes's model stands out because it is one of the few attempts to integrate developmental research into the literature on judgment and decision making, despite the long-standing recognition of the importance of learning to cognitive competence.⁵¹

49. Nor does Byrnes. Byrnes treats processing biases, including those arising from heuristics, as moderating factors that prevent optimal use of resources or distract from goal attainment. See Byrnes, *supra* note 23, at 212.

50. Evidence that market experience reduces the incidence of some irrational behaviors further supports this dynamic approach to judgment and decision making. See, e.g., John A. List, *Does Market Experience Eliminate Market Anomalies?*, 118 Q.J. ECON. 41, 70 (2003) (finding "strong evidence that individual behavior converges to the neoclassical prediction as trading experience intensifies").

51. Einhorn noted long ago the crucial relation of outcome feedback to decision quality:

A major variable in understanding heuristics is outcome feedback. Since outcome feedback is the main source of information for evaluating the quality of our decision/judgment rules, knowledge of how task variables both affect outcomes and influence the way outcomes are coded and stored in memory becomes critical in explaining how heuristics are learned and used.

Hillel J. Einhorn, *Learning From Experience and Suboptimal Rules in Decision Making*, in JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES 268,

Also, we do not assume that all individuals exhibit the same potential for self-regulation of their decision making nor exhibit the same sensitivity to motivational effects on their rational behavior.⁵² As Byrnes notes, several personality traits serve to amplify or moderate the development of self-regulated decision making,⁵³ and individuals learn at different rates.⁵⁴ However, in light of the growing empirical evidence that individuals differ in their ability to achieve high levels of decision competence,⁵⁵ we believe it defensible to assume that individuals differ in their potential for self-regulation and to argue that, at the margin, paternalistic policies will interfere with the development of decision-making competence.

B. THE AUTOGENETIC EFFECTS OF PATERNALISM

Acting paternalistically toward a particular group, on

269 (Daniel Kahneman et al. eds., 1982). Reyna and Brainerd's fuzzy-trace theory offers another important developmental perspective on judgment and decision making. See Valerie F. Reyna et al., *Memory, Development, and Rationality: An Integrative Theory of Judgment and Decision Making*, in EMERGING PERSPECTIVES ON JUDGMENT AND DECISION RESEARCH, *supra* note 20, at 201, 201; see also Judite V. Kokis et al., *Heuristic and Analytic Processing: Age Trends and Associations With Cognitive Ability and Cognitive Styles*, 83 J. EXPERIMENTAL CHILD PSYCHOL. 26 (2002) (describing the alternative dual-process approach to cognitive development).

52. A particularly important mediating variable is likely to be the individual's cognitive disposition, which encompasses the more flexible and malleable aspects of cognition (e.g., the disposition to weigh the opinions of others before forming a decision and the amount of time and effort expended to resolve decision-making problems) and which may be contrasted with cognitive capacity and the fairly stable and mechanical aspects of cognition (e.g., working memory capacity, perceptual speed). See, e.g., KEITH E. STANOVICH, WHO IS RATIONAL? STUDIES OF INDIVIDUAL DIFFERENCES IN REASONING 157 (1999) (describing cognitive disposition as the "psychological mechanisms and strategies that tend to generate characteristic behavioral tendencies and tactics" (citation omitted)). A person with low cognitive capacity but who is disposed to exert greater time and effort to the resolution of decisional problems may achieve results equal to or better than those achieved by persons with greater cognitive capacity.

53. See, e.g., Byrnes et al., *supra* note 23, at 1123 (emphasizing the moderating effects of dogmatism and impulsivity).

54. See BYRNES, *supra* note 22, at 116 ("[S]elf-regulation is arrayed along a continuum ranging from complete dysregulation to complete self-regulation. Most of us fall somewhere in the middle of the continuum.")

55. See Andrew M. Parker & Baruch Fischhoff, *Decision-Making Competence: External Validation Through an Individual-Differences Approach*, 18 J. BEHAV. DECISION MAKING 1, 21 (2005) (validating a measure of individual differences in decision-making competence). See generally STANOVICH, *supra* note 52 (demonstrating that individual differences implicate rationality).

grounds that the group is at risk of making irrational and inefficient decisions, is likely to have autogenetic consequences.⁵⁶ First, the paternalist takes steps to restrict the contractual freedom of the regulated parties in the belief that the regulated parties lack full decision-making competence. Yet such contractual freedom is needed to develop competence in decision making, as suggested above in our discussion of the self-regulation model of decision making. Thus, restriction of freedom to contract is likely to reinforce the need for paternalistic oversight in the regulated domain and other domains.⁵⁷ The perception of irrationality in the general public likewise leads to self-fulfilling behaviors within the paternalist himself, such as interpreting ambiguous evidence as evidence of irrational consumer behavior or engaging in strict review of disconfirming data and lax review of confirmatory data.⁵⁸ If the paternalist invests significant political capital to advance paternalistic policies, the pressure to find evidence confirming the need for the paternalism is likely to be significant as well, making the likelihood of a self-fulfilling prophecy quite high.

Second, labeling a contractual domain an area of consumer exploitation is likely to lead to ex post facto requests for paternalistic intervention.⁵⁹ Thus, in addition to paternalism's interference with individual learning, the willingness of courts to

56. Kukla uses the term "autogenetic" as shorthand for beliefs that lead to self-fulfilling or self-negating prophecies. See Andre Kukla, *The Structure of Self-Fulfilling and Self-Negating Prophecies*, 4 THEORY & PSYCHOL. 5, 5 (1994).

57. Cf. Robert E. Scott, *Error and Rationality in Individual Decisionmaking: An Essay on the Relationship Between Cognitive Illusions and the Management of Choices*, 59 S. CAL. L. REV. 329, 362 (1986) ("[I]ntervention to correct a market failure may have far more powerful secondary effects on consumer satisfaction than has been commonly acknowledged [T]he social engineer must be sensitive to the damage that is likely to be caused to the mechanics by which individuals regulate their choices.").

58. See, e.g., Dale T. Miller & William Turnbull, *Expectancies and Interpersonal Processes*, 37 ANN. REV. PSYCHOL. 233, 244 (1986) ("The influence of erroneous expectancies is not only manifest in the behavior of targets. In spite of objective evidence to the contrary, perceivers may conclude that their expectancies have been confirmed. To the extent that confirmation is 'in the eye of the beholder,' stereotypes and other false expectancies will persist even in the face of objective disconfirmation."). For a discussion of such confirmatory biases within behavioral law and economics, see Gregory Mitchell, *Taking Behavioralism Too Seriously? The Unwarranted Pessimism of the New Behavioral Analysis of Law*, 43 WM. & MARY L. REV. 1907, 1911, 2019 (2002).

59. See *supra* note 47. For additional cases in which plaintiffs have requested paternalistic protection from the courts for supposed manipulation of cognitive or motivational biases, see Rachlinski, *supra* note 3, at 1166 n.12.

engage in paternalistic oversight reduces the risk of personal liability for poor choices and increases the likelihood that even rational actors will seek relief from choices that turn out badly. In other words, the motivational effects on litigants and lawyers of a paternalistic attitude in the courts add to paternalism's autogenetic prospects.

Third, changing the market by restricting the profit exploitation opportunities of firms should cause rational firms to alter their exploitation strategies to preserve their profits.⁶⁰ Paternalistic intervention thus breeds demand for more paternalistic interventions, as the paternalist tries to keep pace with the arbitrage efforts of firms seeking to exploit cognitive biases.⁶¹ Such interference with competitive market forces may adversely affect the development of rational behavior, because highly competitive free markets tend to foster rational choice better than less competitive and nontraditional markets.⁶² One reason that rational choice theory fares best in highly competitive markets seems to be that these markets provide "external scaffolding" for individual choice that channels behavior in utility-maximizing directions, while the environment in less competitive markets allows suboptimal behavior to survive and, in some cases, even prosper.⁶³ Thus, to the extent economic effi-

60. *Cf.* Frey & Eichenberger, *supra* note 42, at 219 ("Firms not only have an incentive to exploit the given stock of anomalies but also an interest in expanding the existing capacity of anomalies (which allows them to raise exploitation). New anomalies can be detected by investing resources in appropriate research, and known anomalies can be combined in such a way that they are magnified."); Hanson & Kysar, *supra* note 2, at 1424–25 ("[I]ndividuals' irrationality makes them] susceptible to manipulation by those actors in a position to influence the decisionmaking context. Moreover, the actors in the dominant position *must* capitalize on this manipulation or eventually be displaced from the market.").

61. Alternatively, the paternalistic intervention may drive some firms from the market, with possibly adverse effects on pricing in the market.

62. ANDY CLARK, BEING THERE: PUTTING BRAIN, BODY, AND WORLD TOGETHER AGAIN 183 (1997) ("[T]raditional economic theory nicely models choice in competitive posted price markets and in certain restricted experimental studies.").

63. Clark explains how markets can foster rational choice:

[T]he crucial factor distinguishing the successful and unsuccessful cases (of the use of neoclassical, substantive-rationality-assuming theory) is the availability of a structurally determined *theory of interests*. In cases where the overall structuring environment acts so as to select in favor of actions which are restricted so as to conform to a specific model of preferences, neoclassical theory works. And it works because individual psychology no longer matters: the "preferences" are imposed by the wider situation and need not be echoed in individ-

ciency is the primary goal, improving competition rather than protecting individuals from their inefficient irrational tendencies is likely the better long-term strategy.

II. SIMPLE ECONOMIC MODELS OF BEHAVIOR UNDER PATERNALISM

In order to make our somewhat vague claims about the psychological and social effects of paternalism more concrete and testable, we now promulgate economic models of decision-making behavior under paternalism. These models isolate the variables likely to be important to cost-benefit calculations regarding alternative courses of action. Although there is some evidence that individuals do engage in such cost-benefit calculations,⁶⁴ it is not assumed that individuals consciously engage in such cost-benefit calculations in all instances. Rather, empirical testing may reveal that in some situations individuals engage in explicit cost-benefit reasoning along the lines we suggest but that individuals also develop conditioned responses to particular choice situations that reflect implicit or prior explicit cost-benefit analyses consistent with our hypotheses.⁶⁵

ual psychology.

Id. Markets are thus just a special case of what Clark calls “external scaffolding” of thought: “Language and culture, in particular, emerge as advanced species of external scaffolding ‘designed’ to squeeze maximum coherence and utility from fundamentally short-sighted, special-purpose, internally fragmented minds.” *Id.* at 33. According to Clark, “we are masters at structuring our physical and social worlds so as to press complex coherent behaviors from . . . unruly resources. . . . Our brains make the world smart so that we can be dumb in peace!” *Id.* at 180.

64. First, some individuals, particularly those with training in economics, employ cost-benefit reasoning in their day-to-day lives. *See, e.g.*, Richard P. Larrick et al., *Teaching the Use of Cost-Benefit Reasoning in Everyday Life*, 1 PSYCHOL. SCI. 362 (1990); Richard P. Larrick et al., *Who Uses Cost-Benefit Rules of Choice? Implications for the Normative Status of Microeconomic Theory*, 56 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 331 (1993). Second, all individuals seem to possess some ability to shift from intuitive thought into a more deliberative mode of thought in which the costs and benefits of alternative courses of action are more carefully considered. *See, e.g.*, ZIVA KUNDA, SOCIAL COGNITION: MAKING SENSE OF PEOPLE 106 (1999).

65. An intriguing possibility is that the brain’s dopaminergic system operates on principles that approximate economic norms for evaluation, with subconscious, neuron-level valuations of potential future rewards becoming associated with different courses of action and activities causing changes in a subset of dopamine neurons to improve the system’s ability to predict error regarding future behaviors. *See* P. Read Montague & Gregory S. Berns, *Neural Economics and the Biological Substrates of Valuation*, 36 NEURON 265, 281 (2002); *see also* Peter Dayan & Bernard W. Balleine, *Reward, Motivation, and*

A. EX ANTE PATERNALISM

In a world of ex ante paternalism, choices are restricted such that options disfavored by policy makers are removed from an individual's choice set or are taxed at a relatively high level to discourage individuals from choosing them.⁶⁶ While this kind of paternalism has the potential to stop individuals from making "mistakes" in their choices, it is not clear that this mistake avoidance is entirely salutary, as discussed above. Specifically, in a situation in which there are no artificial constraints on the choice set, individuals have the incentive to invest effort in searching for the choice that will maximize their utility.

Operationally, the individual faces the following maximization problem in which her choice of effort (e), which involves some expenditure C , determines the probability (p) that she will choose the good that provides her with high utility (U_H) as opposed to the good providing low utility (U_L):

$$\text{Max}_e p(e) \cdot U_H + [1 - p(e)] U_L - C \cdot e$$

which yields the solution

$$\left(\frac{\partial p}{\partial e} \right) [U_H - U_L] = C$$

The first order condition implies that an individual's choice of search effort will be determined by: (1) the productivity of search $\left(\frac{\partial p}{\partial e} \right)$ or how much each unit of effort increases the

Reinforcement Learning, 36 NEURON 285, 293 (2002); Paul W. Glimcher, *Decisions, Decisions, Decisions: Choosing a Biological Science of Choice*, 36 NEURON 323, 328–29 (2002). But see Colin Camerer et al., *Neuroeconomics: How Neuroscience Can Inform Economics*, 43 J. ECON. LITERATURE 9, 55 (2005) ("Our view is that establishing a neural basis for some rational choice principles will not necessarily vindicate the approach as widely applied to humans.").

66. See, e.g., Jonathan Gruber, *Smoking's Internalities*, 25 REGULATION 52 (2002/2003); Jonathan Gruber & Botond Koszegi, *Is Addiction "Rational"? Theory and Evidence*, 116 Q.J. ECON. 1261 (2001). Note that in these articles, the taxation of cigarettes is advocated not because of the ability of taxation to internalize the cost of external effects on third parties or even to account for the costs that smoking imposes on the public health system (so-called fiscal externalities), but rather such taxation is advocated to remedy "internalities" (i.e., effects on a future self). For a more direct exposition of this position, see Jonathan Gruber & Sendhil Mullainathan, *Do Cigarette Taxes Make Smokers Happier?* (Nat'l Bureau of Econ. Research, Working Paper No. 8872, 2002), available at <http://nber.org/papers/w8872.pdf>.

likelihood that the “right” good will be chosen;⁶⁷ (2) the value of choosing the correct as opposed to the incorrect good $[U_H - U_L]$ which is represented by the utility differential generated by the two goods; and (3) the cost of effort.

As long as search effort is costly, an individual will likely accept an expected outcome in which she sometimes chooses incorrectly because the marginal value of increased search, after some point, will not exceed the marginal cost of the effort required by the search.⁶⁸ However, in some ways, this one-time static model of behavior undervalues the benefit provided by search effort. That is, in many cases, effort in the current period provides benefits beyond the choice at hand. In the future, if faced with the same choice, the individual’s cost of search effort will be reduced if she invested effort earlier.⁶⁹ Effectively, search effort is tantamount to a learning process and the stock of information gathered during that process is accessible when the decision must be made again in the future. This search effort may even pay dividends in situations in which an individual must make decisions in different, though related contexts.⁷⁰ An individual’s cognitive capital can, in an abstract sense, be divided into specific capital which is helpful in making the same decision in the future, and general capital, which will be helpful in making decisions that have elements in common with decisions made in the past. Viewed in this way, any current search effort contributes to durable cognitive capital which generates lower search costs in the future across a wide array of decisions.⁷¹

67. Note that since we can normalize effort units in an arbitrary way, this derivative term can also be interpreted as an elasticity.

68. This assumes an interior solution exists for the problem.

69. Alternately, one could frame this as improved productivity of search for the same cost.

70. It is also possible that the search process will lead to domain-general knowledge and strategies that pay dividends in even seemingly unrelated contexts. For instance, learning to think in an active, open-minded way, such that one considers opposing positions and counterarguments to one’s initial position, has been shown effective at debiasing across several domains. See, e.g., Linda Babcock et al., *Creating Convergence: Debiasing Biased Litigants*, 22 LAW & SOC. INQUIRY 913, 920 (1997); Asher Koriat et al., *Reasons for Confidence*, 6 J. EXPERIMENTAL PSYCHOL.: HUM. LEARNING & MEMORY 107, 116 (1980); Charles G. Lord et al., *Considering the Opposite: A Corrective Strategy for Social Judgment*, 47 J. PERSONALITY & SOC. PSYCHOL. 1231, 1239 (1984); Thomas Mussweiler et al., *Overcoming the Inevitable Anchoring Effect: Considering the Opposite Compensates for Selective Accessibility*, 26 PERSONALITY & SOC. PSYCHOL. BULL. 1142, 1149 (2000).

71. In terms of the self-regulation model of decision making, such a search

In the choice-constrained world envisioned by some behavioral law and economics scholars, this search process is avoided or diminished to the extent that the benefits of search effort are significantly reduced. In the maximization problem laid out above, it would be as if p were exogenously set by the paternalistic policy maker. If p is exogenous, and C is greater than zero, the individual will not undertake any search effort.⁷² Her choice will instead be made for her. This result is an application of the concept of moral hazard drawn from the economics of insurance literature. Specifically, a moral hazard arises when an individual receives no (or an attenuated) personal benefit from her productive activity (generally, in this context, the benefit comes in the form of a reduced cost); therefore, her incentive to engage in the productive activity is diminished. Her welfare, from this particular consumption decision, will primarily be a function of the probability that the policy maker has chosen correctly for her (Ψ). Formally, restricting attention to the one-time decision case, the individual will be better off in the choice-constrained world if

$$\Psi \cdot U_H + (1 - \Psi)U_L > p(e^*)U_H + [1 - p(e^*)]U_L - C \cdot e^*$$

where e^* is the individual's optimal level of search effort in the nonpaternalistic setting.

Simplified, this condition requires that the expected benefit of the paternalistic system, $\Psi(U_H - U_L)$, must be greater than the expected outcome of the search model net of search costs, $p(e^*) \cdot (U_H - U_L) - C \cdot e^*$, if paternalism is to benefit an individual. Such a condition will hold if the policy maker is almost as likely to choose correctly as the individual is, where the allowable gap in likelihoods is determined by the cost of search effort relative to the utility gap between the correct and incorrect out-

can lead to gains in crystallized intelligence (acquisition of facts) or fluid intelligence (acquisition of problem-solving strategies) of a domain-specific or domain-general nature. See Margaret E. Beier & Phillip L. Ackerman, *Age, Ability, and the Role of Prior Knowledge on the Acquisition of New Domain Knowledge: Promising Results in a Real-World Learning Environment*, 20 PSYCHOL. & AGING 341, 341 (2005) ("*Fluid intelligence* (Gf . . . , as the processing and reasoning components of intelligence) has been identified as an aptitude for learning, whereas *crystallized intelligence* (Gc) is generally defined as the knowledge acquired through education and experience.>").

72. In the case where the choice set is reduced to some lower number of options but still allows some choice, incentives to invest effort in search will be reduced, as will be the case where the policy maker uses tax policy to induce a certain choice rather than by constraining the choice set directly.

2006] *REGULATION OF IRRATIONALITY* 1645

comes, $\frac{C \cdot e^*}{U_H - U_L}$. That is, if the policy maker has a lower likelihood of choosing correctly for the individual, we might still favor paternalism if the cost of search effort is very high or if the gap between the correct and incorrect outcomes is very small (e.g., perhaps the policy maker has realized economies of scale in information acquisition that the individual cannot, or the choice alternatives differ little, as in certain insurance markets).⁷³

Note that in the above analysis we ignore the cost to the policy maker of determining which choice to impose. This omission seems reasonable if the population is generally homogenous in preferences, or if, for the choice at hand, there is a great deal of agreement regarding what constitutes the correct choice. In these cases, the policy maker enjoys economies of scale in its own decision costs since, effectively, one choice is made for the entire population, whereas in the nonpaternalistic world, each individual would undertake her own search costs to arrive at what would end up being a common choice.

However, the reasonable scope of paternalism shrinks considerably when there is more heterogeneity of preferences among the population. As heterogeneity grows, if the policy maker does not have the ability to offer different choice sets to different people based upon their underlying preferences, Ψ will decrease relative to $p(e^*)$. Also, as heterogeneity grows, if the policy maker attempts to offer different constrained choice sets to individuals based on their preferences, the economies of scale in search costs for the policy maker break down.

In homogenous populations, however, even if the formal case for paternalism strengthens, the practical case is weakened. Individuals with similar preferences will be able to economize on information by simply aping those around them in their choices. Further, in such situations, market pressures will induce firms to provide goods and services that satisfy

73. For instance, Baker and colleagues find that more actively managed funds, in terms of turnover, exhibit significantly better performance than funds with low turnover. However, funds in the highest turnover quintile only outperform those in the lowest quintile by thirty-four basis points on an annualized basis. Thus, if paternalistic regulations prohibited individuals from investing in high-turnover funds, some people would be made worse off, but the effect would be relatively small. See Malcolm Baker et al., *Can Mutual Fund Managers Pick Stocks? Evidence from Their Trades Prior to Earnings Announcements* 26 (Nat'l Bureau of Econ. Research, Working Paper No. 10685, 2004), available at <http://papers.nber.org/papers/w10685.pdf>.

those common preferences, weeding out the goods and services that individuals do not want.

The strongest case for paternalism centers around the case where an individual's optimal likelihood of choosing correctly is low relative to the policy maker's ability to choose correctly for the individual because the relationship between her effort and likelihood of choosing correctly is low (or, isomorphically, because her cost of effort is very high). That is, we would view paternalism most favorably when it is relatively costly for an individual to improve her decision making (e.g., in high-stakes decisions where choices are irrevocable and there is little prior chance for learning). Again, as indicated in the simple model above, in a one-shot transaction, the efficiency of this paternalism depends on the policy maker choosing correctly. However, this welfare analysis ignores the capital stock aspect of individuals investing in cognitive effort. That is, if individuals are effectively precluded from engaging in search at time t because of the policy maker's decision to constrain the choice set, the individual will have a relatively smaller stock of cognitive capital to draw upon in a similar choice at time $t+1$.

In terms of specific capital, this diminution of the capital stock might not be very costly if the paternalist continues to constrain choice in this area for the foreseeable future. However, unless the paternalist constrains choice in very many contexts, the decrease in general cognitive capital has the potential to harm the individual. That is, while in the unconstrained world an individual's cost of cognitive effort will decrease over time because she has contributed to her capital stock through prior decision-making processes and learning, her decision-making capacities will not be exercised in the paternalistic domain. This unintended cost of paternalistic policies reflects a sort of "cognitive hazard."

Assuming that policy makers cannot constrain choices in every context (e.g., due to excessive policy-making and enforcement costs) or do not wish to do so because of the existence of sufficiently diverse populations or concerns about personal rights in some contexts, then an optimally designed paternalistic system needs to consider the cost of this cognitive hazard. Practically speaking, a social planner would need to examine the benefits of constraining choice by comparing the expected gain from choice constraint relative to the individual's choice in the unconstrained world in which she optimally chooses her cognitive effort level, recognizing possible reductions in cogni-

tive capital that will be occasioned by constraining the choice set. Presumably, adding this element to the calculation would not reduce the set of welfare-enhancing paternalistic interventions to the null set, but it would raise the bar for a policy of ex ante paternalism.

B. EX POST PATERNALISM

If behavioral biases are not completely exogenous, then they will be sensitive to changes in the net costs generated by the bias as well.⁷⁴ To the extent that paternalistic interventions shield an individual from the costs generated by her deviations from rationality,⁷⁵ on the margin, she will have greater incentive to exhibit the bias, or, perhaps more intuitively, she will have less of an incentive to surmount her biases. The insurance generated by ex post paternalist protections has the potential to generate the kind of moral hazard that is commonly associated with insurance.⁷⁶

If cognitive biases are responsive to incentives, the efficiency arguments advanced by supporters of ex post paternalism are drawn into question. That is, if an individual recognizes that she will be protected from the consequences of her

74. For a discussion of evidence in support of the view that individuals are sensitive to the costs of psychological bias, and that they often engage in self-corrective actions in attempts to counter the effects of bias, sometimes with success and sometimes without success, see Kahneman, *supra* note 20, at 710–12; see also Diego Fernandez-Duque et al., *Executive Attention and Metacognitive Regulation*, 9 CONSCIOUSNESS & COGNITION 288, 293 (2000) (noting that “[n]ormal subjects have the ability to internally evaluate their own performance, detecting errors even in the absence of any external feedback,” although the authors do not specifically discuss nationality errors); Frey & Eichenberger, *supra* note 42, at 223 (“Individuals falling prey to an anomaly are often aware that they act in a non-rational way and that they could improve their utility or profit by adjusting their behavior. Perception is the more likely, the less costly is a comparison with non-anomalous behavior.”).

75. See, e.g., Korobkin, *Standard Form Contracts*, *supra* note 3, at 1207 (“By recognizing purchasers’ bounded rationality as the most important root cause of inefficiency in form contracts, courts can modify their use of unconscionability analysis to increase both social welfare generally and buyer welfare specifically.”).

76. Arguably, moral hazard as it is invoked in the economics literature is simply an application of the law of demand, which states that as the price of an activity decreases, an individual will increase her consumption of the activity. See Jonathan Klick & Thomas Stratmann, *Subsidizing Addiction: Do State Health Insurance Mandates Increase Alcohol Consumption*, 35 J. LEG. STUD. 175, 177 (2006); Jonathan Klick & Thomas Stratmann, *The Effect of Abortion Legalization on Sexual Behavior: Evidence from Sexually Transmitted Diseases*, 32 J. LEG. STUD. 407, 412 (2003).

decisions either because of statutory proclamation or because of her knowledge that courts have acted paternalistically toward others in similar situations, she will not invest an optimal level of cognitive effort. Thus, the short-term efficiency gain posited by the paternalists may not offset the efficiency loss that will apply to parties in the future.

As a simple illustration of this relationship, take an individual who can achieve either of two utility levels in a given transaction: the outcome that would occur if there were no cognitive limitation present (U_H) and the outcome resulting in the presence of the cognitive limitation (U_L). By assumption, U_H is strictly greater than U_L . The individual can exert greater cognitive effort (e) to increase her probability (p) of attaining U_H , but each additional unit of effort reduces her utility by C . Further, she has some expectation that in the event she realizes U_L as the end state of the transaction, the courts will compensate her with some award that increases her utility by I .

In deciding on the level of cognitive effort she wishes to expend, the individual faces the following maximization problem:

$$\text{Max}_e p(e)U_H + [1 - p(e)]U_L + [1 - p(e)]I - C \cdot e$$

Effectively, the individual is maximizing her expected utility from the transaction, including the insurance provided to her by the paternalistic institution, net of the disutility she experiences by exerting increased cognitive effort.⁷⁷

The result of this maximization problem can be characterized by the function's first order condition:

$$\frac{\partial p}{\partial e}(U_H - U_L - I) = C$$

The condition presented above has the intuitive implication that an individual will continue to expend effort up to the point where doing so is cost justified. That is, given that each additional unit of effort decreases her utility by C , she will only expend additional effort if her expected gain in utility of doing so is greater than C .

It is immediately apparent that if the difference between

77. This cognitive effort cost could take many forms. The most direct would simply be the time cost of considering the transaction at greater length in order to determine whether the individual is receiving the greatest gain possible from the bargain.

U_H and U_L is entirely offset by the “insurance” payment provided by the paternalistic institution, the individual will expend no cognitive effort.⁷⁸ This stands in contrast to the result that exists in a world with no paternalism. If there is no insurance in the model, the individual will exert some positive level of cognitive effort as long as $U_H > U_L$.

The situation becomes more acute if a given cognitive limitation worsens through reinforcement. That is, if an individual is protected from the costs of her limitation in period 1, the magnitude of the limitation is larger in period 2 because of conditioning or adaptation. Another way to understand the effect is that the cost of cognitive effort in period t decreases as cognitive capital is built up through exercising cognitive faculties in periods prior to t . Effectively then, paternalist protection not only generates a moral hazard in which an individual underinvests in cognitive effort during the current period, but it also reduces an individual’s cognitive abilities in future periods relative to the situation in which no protection is provided. In the no-protection case, an individual will have an incentive to consider the effect of current expenditures on cognitive effort on the cost of future cognitive effort. The presence of insurance against the costs of cognitive mistakes currently and the expectation of insurance in the future both reduce cognitive investments. This phenomenon is the manifestation of our cognitive hazard concept in the context of ex post paternalistic interventions.⁷⁹

78. In fact, in such a situation, the individual will be at a corner solution in which she would like to expend negative amounts of effort if she were compensated for doing so (i.e., if she could face a negative cost of effort).

79. Although the concept of cognitive hazard, in principle, can be seen as an extension of the moral hazard idea, it is analytically useful to keep the two phenomena distinct in any welfare analysis given the different time frames that are relevant for each. That is, because the costs of cognitive hazards are only realized in the future, they should be discounted appropriately by a social planner, while the costs associated with moral hazards will be realized at the time the current decision takes place. For example, assume an individual can invest \$1 worth of effort to guarantee that she will achieve the high utility outcome (which exceeds the low utility outcome by \$5) now and whenever she is in a similar situation in the future, but she has no incentive to make that investment due to the existence of insurance provided by the paternalist which involves a transfer of \$5 to the individual and administrative costs of \$1. The social cost of the moral hazard in this situation is zero because the individual’s cost of effort equals the social cost of administering the paternalistic insurance. However, the social cost of her cognitive hazard (i.e., the failure to improve cognitive capital to draw on in future situations) will be the present value of \$1 for every instance in the future in which this transaction arises

C. A ROLE FOR EDUCATION

The cost of cognitive effort, and hence the magnitude of a revealed limitation, will not only be dependent upon an individual's own internal investments, but may be influenced by external investments in the form of education. That is, in some circumstances, individuals can learn debiasing techniques that increase the likelihood of rational choice.⁸⁰

One of the simplest ways to add education to our primary model is to assume that the probability of choosing the correct option is a function of education inputs (E) and cognitive effort inputs (e).⁸¹ Thus, if an individual were free to improve her de-

again using the relevant social discount rate. Had the individual invested in effort during the first transaction, by assumption, there would be no need for the paternalist to remedy her mistake through the insurance mechanism in the future. If, instead, the insurance program entails administrative costs that exceed the individual's cost of effort, say by \$2, the moral hazard cost of the original intervention would be \$1 which represents the differential resources consumed by the paternalistic remedy relative to what the individual would have needed to expend to reach the high utility. This is a current cost of the system and so it is not discounted, while the cognitive hazard costs will be discounted as before. Thus, as long as the social planner assumes a positive discount rate, moral hazard costs (i.e., costs that are currently realized) will be weighted relatively more heavily than cognitive hazard costs (i.e., costs that will be realized in the future).

80. Given the relative neglect of statistical and economic training at the compulsory education levels, it should not be particularly surprising that people often evince poor comprehension of statistical and economics concepts. See PETER SEDLMEIER, IMPROVING STATISTICAL REASONING: THEORETICAL MODELS AND PRACTICAL IMPLICATIONS 196 (1999) (discussing the low levels of statistical education at primary and secondary school levels); William B. Walstad, *Economic Education in U.S. High Schools*, 15 J. ECON. PERSP. 195, 208 (2001) ("Less than half of all high school graduates take a well-defined high school economics course at present."). Nevertheless, a number of educational programs have been shown effective at debiasing. See Gregory Mitchell, *Why Law and Economics' Perfect Rationality Should Not Be Traded for Behavioral Law and Economics' Equal Incompetence*, 91 GEO. L.J. 67, 87-94 (2002) (summarizing evidence on the relationship between education and rationality); *id.* at 132-35 (summarizing evidence on procedural debiasing techniques). See generally Richard P. Larrick, *Debiasing*, in BLACKWELL HANDBOOK ON JUDGMENT AND DECISION MAKING 316 (Derek J. Koehler & Nigel Harvey eds., 2004) (describing debiasing strategies for individuals). For a recent study suggesting that education can have positive effects on the expression of psychological biases, see Daniel J. Benjamin & Jesse Shapiro, *Does Cognitive Ability Reduce Psychological Bias?* 25-26 (Feb. 25, 2005), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=675264 (reporting that although the primary finding of the study was "a robust positive relationship between cognitive ability and normative decision-making," there is some evidence that "seems to suggest that human capital policy may indeed be able to improve economic decision-making." (footnote omitted)).

81. Operationally, this could be modeled in a number of ways. We choose

cision-making process by investing in education at the cost of G per educational unit and cognitive effort at the cost of C per effort unit, she would face the following maximization problem:

$$\text{Max}_{e,E} p(e, E) \cdot U_H + [1 - p(e, E)] U_L - C \cdot e - G \cdot E$$

generating the following first order conditions:

$$\left(\frac{\partial p}{\partial e} \right) [U_H - U_L] = C$$

$$\left(\frac{\partial p}{\partial E} \right) [U_H - U_L] = G$$

which can be combined into the following relationship:

$$\frac{\partial p / \partial e}{\partial p / \partial E} = \frac{C}{G}$$

This result holds that the individual will set her marginal rate of substitution⁸² between cognitive effort and education equal to the cost ratio of the two inputs. Thus, as long as there is some gain from improving p ,⁸³ as cognitive effort grows more expensive relative to education, the individual will invest in more education and vice versa.

However, if the individual is making her decision within a paternalistic institution such that her choice set is censored or such that she will be saved from making a bad decision, she will invest less in both education and cognitive effort, generating the same short-term moral hazards discussed above. This will also generate the longer-term cognitive hazard discussed above, as the individual has less incentive to invest in her cognitive capital through effort or education.

to present a model in which education and effort are substitute inputs in the probability function, but the intuition would not change if the two inputs were modeled as complements such that p was a function of effort times education or such that education reduced the per unit cost of cognitive effort.

82. See ANDREU MAS-COLELL ET AL., MICROECONOMIC THEORY 54 (1995).

83. This will be the case as long as the gap between U_H and U_L (net of the implicit insurance payment available if this is a case of ex ante paternalism) is smaller than the minimum of C and G . That is, as long as there is an interior solution to the optimization problem.

The addition of education to the model, however, suggests an alternate route through which paternalist policy makers can affect an individual's chances of choosing the correct option from a set, namely, the paternalist may subsidize the individual's educational inputs. This option, however, is not a perfect one. While under most circumstances such a subsidy will increase aggregate cognitive inputs (i.e., effort plus education), there will also be some substitution effect whereby, on the margin, the individual will decrease her own cognitive effort and education inputs. That is, it will generally not be the case that the paternalist can increase aggregate inputs one for one. If the paternalist provides a subsidy paying for one unit of education, the individual's aggregate input units will increase by less than one unit.

The provision of paternalistic education subsidies could be justified, however, under several conditions. Most obviously, if the paternalist can provide education at a lower per unit cost, even taking the individual's substitution into account, paternalist education provision would be efficient. This should be the case when there are economies of scale involved in cognitive education. Educational subsidies could also be justified in cases where individuals are assumed to save (i.e., invest in cognitive capital) too little due to peculiar discounting functions,⁸⁴ so long as this limited savings tendency generates some kind of relevant externality.⁸⁵

This alternative paternalistic policy tool, and the recognition that individuals will react in complex ways to paternalistic policy interventions, implies that optimal paternalism will typically involve more than simply identifying a bias and counteracting it directly through legal and policy institutions. Instead, optimal paternalism represents a mechanism design problem in which a social planner must consider using more than just the sledgehammer of constraining choice sets *ex ante* or providing implicit social insurance through some form of *ex*

84. For example, if an individual exhibits hyperbolic discounting as hypothesized by Laibson and others. See, e.g., David Laibson, *Golden Eggs and Hyperbolic Discounting*, 62 Q.J. ECON. 443 (1997).

85. As in Gruber's "internality" argument, in which the present self does not take the well-being of the future self into account, or when there are so-called fiscal externalities, where an individual's poor decisions affect others through governmental programs. See Edgar K. Browning, *The Myth of Fiscal Externalities*, 27 PUB. FIN. REV. 3 (1999). For example, given the U.S. welfare system, it may be the case that we want to induce individuals to save more so that they do not end up drawing resources from public assistance programs.

post paternalism. In addition to those tools, the paternalist should consider the indirect tool of education. Further, as implied above, the paternalist should recognize that short-term gains from interventions may be mitigated by long-term losses in cognitive capital, at least in some contexts for some individuals.

III. OPTIMAL DESIGN OF PATERNALISTIC INTERVENTIONS

In this Part, we consider the factors likely to affect optimal institutional design undertaken to counter supposed inefficiencies associated with irrational behavior. A social planner charged with designing an optimal system would presumably have as control variables the degree to which choices are constrained or the level of insurance provided to biased individuals and the level of debiasing education available. As noted above, choice sets may be constrained directly or indirectly through taxes, and social insurance may take the form of judicial decisions (including bankruptcy decrees) to transfer wealth from an individual who benefits from another individual's bias to the biased individual herself. Imposing no choice-set constraint or offering no insurance would represent a *laissez-faire* stance in which there is no paternalistic oversight of the biased behavior. The education variable would involve providing resources for decision-making education or debiasing in the existing school system or through subsidies to market providers of debiasing education.⁸⁶

Our psychological and economic accounts of behavior under paternalism suggest that the optimal mix of paternalistic and educational interventions will depend upon the following factors: (1) the efficiency loss due to current underinvestment in cognitive effort and education; (2) the capitalized loss of the future return due to current underinvestment in cognitive effort

86. In addition to making educational investments to improve one's own decision-making competence, a common strategy to overcome personal limitations is to rely on the expertise of others. *See* BYRNES, *supra* note 22, at 52 ("In summary, the data suggests that certain forms of advice seeking increase with age, others decrease, and others stay the same. From the standpoint of [the self-regulation model], self-regulated decision-makers seek advice only when it is necessary and seek it in appropriate ways."). A social planner could be designated to choose the appropriate level of education and the most efficient mix of educational providers. If it seems unlikely that individuals will seek out education on their own, it might be useful to provide it through existing compulsory schooling.

and education; (3) the costs of private effort and education and the costs of public education subsidies; (4) the efficiency gains from limiting decision-making mistakes; (5) the efficiency costs of the policy maker choosing incorrectly for some individuals;⁸⁷ (6) the welfare gains of discouraging the exploitation of cognitive biases; and (7) the nondistortionary costs of developing and enforcing the constrained choice set and paternalistic insurance system (i.e., administrative costs).

Unfortunately, very little empirical information exists to help us place parameters on these factors with respect to the consequences of particular anomalies (i.e., with respect to factors four and six in particular),⁸⁸ because most paternalistic proposals within behavioral law and economics proceed from little more than an identification of a cognitive bias that may explain a pattern of seemingly inefficient choices.⁸⁹ In many instances it is extremely difficult to tie apparent real-world inefficiencies to a market failure based on irrational behavior as opposed to some other defect in the marketplace,⁹⁰ and in some

87. This cost is most obvious in the constrained choice set case. However, costs of this nature also accrue to ex post paternalistic interventions to the extent that they inject uncertainty in transactions (e.g., reliance on contract terms).

88. Cf. James A. Henderson, Jr. & Jeffrey J. Rachlinski, *Product-Related Risk and Cognitive Biases: The Shortcomings of Enterprise Liability*, 6 ROGER WILLIAMS U. L. REV. 213, 255 (2001) (asserting that proponents arguing for enterprise liability within products liability law on grounds that markets exploit psychological biases to manipulate consumer perceptions of product risks “fail[] to provide a *serious* quantitative assessment of the extent of market manipulation”).

89. This practice follows from the practice within psychology to focus on demonstrating deviations from rationality rather than examining the costs, adaptivity, or boundaries of such deviations. See Krueger & Funder, *supra* note 20, at 313 (noting that mainstream social psychology’s “prevalent research strategy has been to propose a prescriptive norm for social behavior or cognition and then to demonstrate that human performance falls short of it.”); Ralph Hertwig & Annika Wallin, *Out of the Theoretical Cul-de-Sac*, 27 BEHAV. & BRAIN SCI. 342, 343 (2004) (“Social psychology and related fields have oversold violations of behavioral and cognitive norms.”). There appears, however, to be some movement toward a more balanced view of human cognition within psychology, if not behavioral economics. See Andreas Ortmann & Michal Ostadnický, *Proper Experimental Design and Implementation Are Necessary Conditions For a Balanced Social Psychology*, 27 BEHAV. & BRAIN SCI. 352, 352 (2004) (“[A]lthough the unbalanced view of humans as cognitive misers seems slowly but surely on its way out in social psychology and judgment and decision-making, the heuristics-and-biases program, which seems mostly responsible for the unbalanced view, has during the past decade invaded economics with little resistance . . .” (citations omitted)).

90. For an illustration of the difficulty in quantifying welfare costs attrib-

cases even designating the optimal behavior is difficult.⁹¹ Given

utable to irrational behavior, see Avishalom Tor, *The Fable of Entry: Bounded Rationality, Market Discipline, and Legal Policy*, 101 MICH. L. REV. 482 (2002). Tor does an admirable job considering possible explanations for excessive business-startup behavior from both rational choice and bounded-rationality perspectives and considering the positive and negative effects of excessive-startup behavior. Ultimately, however, he notes the difficulty of reaching firm conclusions about the efficiency effects of irrational behavior in this context:

In principle, an effective detection and prevention of some negative net present value entry at the margin could be beneficial. The problem is, however, that there is no easy means of quantifying with any certainty the benefits and costs of entry. Measuring the direct deadweight losses of negative expected value entry is the least difficult, since—at least in principle—one could calculate the number of failed entrants per industry and the average costs they have sunk into their ventures. The calculation of net losses would be more complex, though, for those entrants who obtain some profits before exiting the market.

It is even less clear, moreover, how to measure the externalities of entry

An examination of the quantities that must be measured for any quantification of the positive externalities of negative expected value entry reveals another impenetrable maze.

Id. at 546. Tor concludes that greater governmental regulation of market entry cannot be justified. *See id.* at 567 (“A comparison of the social costs and benefits of negative expected value entry found, however, that while overconfident entrants generate negative externalities, they also bring about significant social benefits Upon further analysis it also became apparent that the regulation of entry would not only be mostly undesirable, but also costly, impractical, and largely ineffective.”). We suspect that if other behavioral law and economics scholars undertook similar attempts to quantify the social costs and benefits of seemingly irrational behavior they would reach similar conclusions, or at least be less sanguine about the benefits of intervention.

91. Identifying optimal behavior is particularly difficult with respect to intertemporal choices, for it is not clear how one should optimize between present and future aims. *See* Shane Frederick et al., *Time Discounting and Time Preference: A Critical Review*, in *TIME AND DECISION* 13, 19 (George Loewenstein et al. eds., 2003) (noting that, although the discounted utility (DU) model has become the normative model of intertemporal choice within economics, “Samuelson did not endorse the DU model as a normative model of intertemporal choice”); *id.* at 30 (“[T]he patterns of preferences that are regarded as ‘anomalies’ in the context of the DU model do not necessarily violate any standard or principle that people believe they should uphold. Even when the choice pattern is pointed out to people, they do not regard themselves as having made a mistake (and probably have not made one!)”); *see also* Jay Bhattacharya & Darius Lakdawalla, *Time-Inconsistency and Welfare* 1–2 (Nat’l Bureau of Econ. Research, Working Paper No. 10345, 2004), available at <http://www.nber.org/papers/w10345>.

Also, it can be very difficult to predict exactly how a bias will evidence itself in complex situations. Yet the efficiency potential of various interventions may depend quite heavily on knowing the functional form of the underlying bias to be remedied. As demonstrated in Eric Posner’s discussion of optimism

the difficulties of quantifying aggregate social costs of irrational behavior, a fallback position would be to estimate the strength or prevalence of the biased behavior itself. Yet in most cases the proposals do not even contain information about the base rate frequency of the biased behavior within the target population, much less an estimate of the economic losses to individuals of the biased behavior.⁹²

Despite the dearth of empirical data relevant to many of these factors, we may nevertheless offer some guidance on the issues raised by these factors. First, many anomalies exhibit considerable elasticity in response to education, material incentives, and self-initiated effortful thought.⁹³ The greater the elasticity, the less compelling the argument for hard paternalism absent data on the welfare costs of an anomaly. The common use of experimental studies that focus on demonstrating anomalies rather than testing the limits of the anomalies often masks both individual and situational differences in biased behavior.⁹⁴

bias, for instance, plausible assumptions about the risk-limiting behavior of an irrational optimist may completely reverse the optimal policy recommendation. See Eric A. Posner, *Probability Errors: Some Positive and Some Normative Implications for Tort and Contract Law*, 11 SUP. CT. ECON. REV. 125, 127–28 (2004).

92. This failure to consider base rates of cognitive bias is a function of the conflation of statistically significant experimental findings with findings of practical significance outside the laboratory. That is, if a study finds that a statistically significant percentage of the participants exhibited irrational behavior, then the finding is couched as evidence of some irrational tendency. Yet in many cases only a minority of participants needs to exhibit the biased behavior to achieve statistical significance (i.e., for the researcher to conclude that the bias is not simply random error). For a more detailed discussion of the importance of distinguishing between statistical and practical significance, see Mitchell, *supra* note 58, at 1954–60. Failure to consider base rates may contribute to what Sunstein calls “probability neglect” driving policy: “the demand for legal intervention can be greatly affected by probability neglect, so that government may end up engaging in extensive regulation precisely because intense emotional reactions are making people relatively insensitive to the (low) probability that the relevant dangers will ever come to fruition.” Cass R. Sunstein, *Probability Neglect: Emotions, Worst Cases, and Law*, 112 YALE L.J. 61, 68 (2002).

93. For a detailed discussion of the considerable individual differences and situational variation in rational and irrational behavior, see Mitchell, *supra* note 80, at 83–119, 139–67.

94. See, e.g., Jay J.J. Christensen-Szalanski & Cynthia Fobian Willham, *The Hindsight Bias: A Meta-analysis*, 48 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 147, 149 (1991) (“Many studies in the judgment literature merely indicate whether a bias exists according to a particular statistical level of probability. This knowledge, however, is not adequate information for

Second, some very simple—and presumably cheap-to-implement—educational and procedural interventions prove quite effective at debiasing. For instance, short courses in statistical reasoning that could easily be incorporated into an educational curriculum show great promise.⁹⁵

Third, we should expect greater error in the paternalist's specification of a restricted choice set (the fifth factor) in markets with traditionally large numbers of options or many heterogeneous goods. This is particularly true for highly competitive markets, where such heterogeneity of goods can be seen as reflective of true heterogeneous demand. The need for governmental intervention should be less in such markets assuming that it is difficult for firms to distinguish between rational and irrational consumers, because bias-exploiting contractual terms should be less likely to persist in such markets.⁹⁶ Thus, a persistent supply of heterogeneous goods in a market populated by many firms may serve as a rough indicator that rational consumer behavior predominates in this market.

a practitioner deciding whether to be concerned about a bias.”); Chris Swoyer, *Judgment and Decision-Making: Extrapolations and Applications*, in JUDGMENTS, DECISIONS, AND PUBLIC POLICY 9, 11 (Rajeev Gowda & Jeffrey C. Fox eds., 2002) (“Although little work has been done on individual differences in judgment and choice, these differences are often substantial, and researchers are beginning to ask: *Who* reasons in which ways?”).

95. See SEDLMEIER, *supra* note 80, at 140–41 (summarizing the results of a series of studies on training in statistical reasoning which found large immediate training effects, high transfer to tasks not used in the training, and generally high long-term gains from the training); Peter Sedlmeier & Gerd Gigerenzer, *Teaching Bayesian Reasoning in Less Than Two Hours*, 130 J. EXPERIMENTAL PSYCHOL.: GEN. 380, 396–97 (2001) (presenting results from training studies indicating that Bayesian reasoning may be improved dramatically through various types of short training programs). Sedlmeier and Gigerenzer focused primarily on the training of college students. For an example of successful training in probabilistic thinking among third-graders, see Graham A. Jones et al., *Students' Probabilistic Thinking in Instruction*, 30 J. RES. IN MATHEMATICS EDUC. 487 (1999). For an example of successful training of novice probation officers in the use of statistical principles to predict behavior, see Geoffrey T. Fong et al., *Improving Probation Decisions Through Statistical Training*, 17 CRIM. J. & BEHAV. 370 (1990).

96. See Alan Schwartz, *How Much Irrationality Does the Market Permit?* 29–30 (Am. L. & Econ. Ass'n Ann. Meetings, Working Paper No. 29, 2004), available at <http://law.bepress.com/cgi/viewcontent.cgi?article=1158&context=alea> (“If some consumers are sophisticated while others are naive, but firms cannot tell into which class a consumer falls, and if all consumer types will shop for low prices and preferred contracts, then competition among firms for the marginal consumer will lower the price of every contract type. Further, if there are enough sophisticated consumers, and if the other consumers are not very naive . . . only good contracts will exist in equilibrium.”).

Fourth, allowing individuals to exploit the cognitive biases of their trading partners creates incentives for rent seeking on the part of the unbiased individual. That is, investments to exploit cognitive biases may increase the profits of the exploiter but decrease total welfare.⁹⁷ To the extent that courts undo such transactions, the insurance simply takes the form of a wealth transfer from the exploiter to the exploited. However, this wealth transfer improves efficiency because it discourages exploiter investments *ex ante*. But note that if the elasticities referenced in the first two factors are nonzero, investments in exploitation will actually have a debiasing effect on biased individuals. Thus, it is necessary to net out these benefits when considering the costs of this rent-seeking behavior. In effect, education and exploitation are substitute avenues for tempering cognitive biases. Each carries a cost in that education will consume resources and exploitation will generate transaction elements that lower total welfare. In areas where cognitive bias elasticities are high, the optimal solution will be to limit paternalistic interventions and to provide some educational subsidy. When elasticities are low, optimality will likely demand protection of the biased. The chosen mix of partial protection and education will depend on the relative elasticities and the relative costs of cognitive effort, education, and the paternalistic intervention.⁹⁸

Finally, the administrative costs of paternalism (the sev-

97. This claim is similar but stronger than that made by Korobkin in his form contracts paper. There Korobkin states that firms will compete on salient attributes of form contracts and ignore nonsalient attributes, yielding a transaction that lowers total welfare. *See* Korobkin, *Standard Form Contracts, supra* note 3, at 1234 (“Although market forces should ensure that sellers will offer efficient salient contract terms, non-salient attributes are subject to inefficiencies driven by the strategic behavior of sellers attempting to increase their profits at the expense of unknowing buyers.”). If we add to this calculation investments in exploitation tools, the welfare loss grows even larger.

98. Other, nonefficiency concerns may exist as well. For example, judicially provided insurance will involve wealth transfers from the exploiter to the exploited, which implies that the exploited held a property right not to be exploited. Such an assignment of property rights is not normatively required. In fact, one could argue that switching to a legal regime that accommodates biases through these kinds of wealth transfers violates the property rights of the exploiters. *Cf.* Frey & Eichenberger, *supra* note 42, at 228 (“In today’s welfare-oriented societies, both the existing law as well as government in pursuit of votes tend to work against the economic elimination of actors falling prey to anomalies. Actors, whether individuals or firms, who behave ‘irrationally’ are prevented from bankruptcy or even major losses, while ‘rational’ actors evading anomalies are ‘punished’ by being taxed.”).

enth factor) should take into account the transaction costs of judicial policing of transactions as well as the dampening effect on trade that will occur as individuals effectively lose some freedom to contract. That is, to the extent that courts police transactions, uncertainty about the enforceability of a contract increases, which should lead to less contracting or higher prices to offset enforcement cost increases.

In this discussion of optimal institutional design, we exhibit a bias for quantification of the social costs and benefits of paternalistic versus laissez-faire responses to irrational behavior, despite the realization that quantifying these costs and benefits will often be very difficult. We feel that such a bias is nevertheless appropriate because of the risks of “anecdotal policy making.” Just as some complain about the use of “just so” rational choice stories to explain away anomalous behavior,⁹⁹ a corresponding complaint can be made about the use of anomaly stories to drive public policy. As demonstrated by the proliferation of bounded rationality accounts of legally significant behavior, it is quite easy to generate stories that explain legal phenomena in terms of one of the many behavioral anomalies that have been identified within psychology and behavioral economics.¹⁰⁰ Yet these expert opinions are much more prone to diagnostic and prediction errors than a reliance on actuarial or statistical data,¹⁰¹ and hence seeking to quantify the costs and benefits of irrationality regulation may serve a useful disciplining effect that leads to more rational policy making.¹⁰²

99. See Jeanne L. Schroeder, *Just So Stories: Posnerian Methodology*, 22 CARDOZO L. REV. 351, 352 (2001) (“I explore at length the degeneration of Posner’s conception of rationality from the elegant, if simplistic, model drawn from neoclassical economics to its current ad hoc state.”).

100. See John Conlisk, *Why Bounded Rationality?*, 34 J. ECON. LIT. 669, 670 (1996) (“The sheer number of experiments reporting biases is so great that a sizable number of books and long survey papers have been written just to review the evidence.”); Krueger & Funder, *supra* note 20, at 317 (“Just as God has been said to have an inordinate fondness for beetles, having made so many, social psychologists may have an inordinate fondness for errors, having found so many.” (citation omitted)). In many cases, however, one bias story can be countered with another bias story given the proliferation of experimentally found biases. See Ortmann & Ostadnick, *supra* note 89, at 352 n.1 (noting that in many cases psychologists have identified a bias and a “contradictory sibling” bias).

101. See generally MICHAEL A. BISHOP & J.D. TROUT, *EPISTEMOLOGY AND THE PSYCHOLOGY OF HUMAN JUDGMENT* (2004) (discussing the benefits of statistical prediction rules).

102. Cf. William Meadow & Cass R. Sunstein, *Statistics, Not Experts*, 51 DUKE L.J. 629, 631 (2001) (“In many settings, the fallible opinions of isolated

Indeed, one of the key messages of behavioral law and economics—that intuitive judgments will often lead one astray—brings into question the use of plausibility as a guide to action. That is, contrary to the position recently advocated by Russell Korobkin,¹⁰³ telling a plausible story about irrational behavior to support a paternalistic intervention, without more, counts as little justification for the intervention given the malleability of plausibility impressions.¹⁰⁴ Put in the language of scientific testing, telling a story that seems plausible in light of the data provides a fairly weak test of a policy proposal, while the quantification of costs and benefits serves as a much more severe

experts should be supplemented or replaced by statistical data. Those opinions should be seen as a kind of crude second-best, far inferior to the data that it approximates.”).

103. See Russell Korobkin, *Possibility and Plausibility in Law and Economics*, 32 FLA. ST. U. L. REV. 781, 783 (2005) (“I propose that the choice between using [a Rational Choice Theory]-based behavioral assumption and a [Behavioral Decision Theory]-based behavioral assumption in law and economics analysis should turn on the relative plausibility of competing accounts in light of existing knowledge, which is often incomplete and indeterminate.”). To the extent Korobkin can be understood to argue that behavioral decision theory accounts should initially be treated as more plausible than rational choice accounts if the former derive from empirical studies and the latter from intuitive, a priori assumptions, we are more sympathetic to Korobkin’s argument. And we certainly agree with the spirit of Korobkin’s argument that models with arguably implausible assumptions should be subjected to empirical testing. *See id.*

104. The relative plausibility standard raises serious problems with respect to what Dawes calls “pseudodiagnosticity” if all we have to guide us is a qualitative assessment of which hypothesis seems better supported by the evidence, evidence which may be selectively gathered by each competing camp:

The main problem here is that hypotheses are not compared: instead, single hypotheses are evaluated in terms of the degree to which evidence is “consistent with” them; in addition, evidence is often *sought* in terms of its consistency with or inconsistency with “favorite hypotheses”—rather than in terms of its ability to distinguish *between* hypotheses.

Robyn M. Dawes, *Behavioral Decision Making and Judgment*, in 1 THE HANDBOOK OF SOCIAL PSYCHOLOGY 497, 533 (Daniel T. Gilbert et al. eds., 4th ed. 1998); *see also* ROBYN M. DAWES, EVERYDAY IRRATIONALITY: HOW PSEUDOSCIENTISTS, LUNATICS, AND THE REST OF US SYSTEMATICALLY FAIL TO THINK RATIONALLY 114 (2001). In short, without some special definition, plausibility seems to be a trait that a hypothesis either has or does not have, rather than a continuum onto which hypotheses can be placed and compared, and in any event, it is unclear how one uncontroversially tests for the existence of plausibility or quantifies a hypothesis’s “degree” of plausibility. *See* Joseph Agassi, *Criteria for Plausible Arguments*, 83 MIND 406, 410–11 (1974) (discussing problems with a relative plausibility standard for judging the scientific acceptability of theses).

test of a proposal.¹⁰⁵ Given the restrictions on liberty associated with paternalism and the attendant cognitive and moral hazards identified here, we believe that the paternalist should bear the burden of demonstrating that the benefits of a paternalistic intervention will likely outweigh the costs.

This is not to say that irrational behavior does not cause significant inefficiencies in some domains, nor that legal regulation will never be needed to ameliorate the situation. Indeed, our framework leads us to believe that there may be a significant role for the government to play with respect to compelling informational disclosures in certain forms that make the information more user-friendly,¹⁰⁶ encouraging competition, subsidizing or compelling some forms of education, and even restricting choices when doing so leads to net efficiency gains. But a failure to take into account the direct and indirect costs of paternalism may lead to perverse results, and, as a behavioral law and economics approach would caution, once a new regulation is in place it may be difficult to dislodge due to a status quo bias.¹⁰⁷

CONCLUSION

One of the key insights of behavioral law and economics is that context influences behavior in many unanticipated ways. We have taken seriously this contextualist insight to consider how government regulation intended to counter irrational tendencies may actually exacerbate the problem (i.e., we have examined irrational behavior as if it were endogenous to the regulatory context). Rather than advance any firm conclusions about the hazards of paternalism, however, our primary goal has been to raise a set of concerns that have received little attention within the behavioral law and economics literature that

105. See, e.g., DEBORAH G. MAYO, *ERROR AND THE GROWTH OF EXPERIMENTAL KNOWLEDGE* 187 (1996) (“If hypotheses that fit the data equally well were equally well supported . . . by the data, then this objection would have considerable weight. But the very *raison d’être* of the severity demand is to show that this is not so.”).

106. See, e.g., Baruch Fischhoff, *Need to Know: Analytical and Psychological Criteria*, 6 *ROGER WILLIAMS U. L. REV.* 55 (2000) (illustrating how to combine analyses of informational settings with psychological research to formulate effective informational disclosures).

107. See Jeffrey Stempel, *Not-So-Peaceful Coexistence: Inherent Tensions in Addressing Tort Reform*, 4 *NEV. L.J.* 337, 351 (2003/2004) (“Applied to public policy, legislation, and elections, the status quo bias probably operates to increase the burden on those seeking a change in laws or government.”).

has become dominated by discussions of paternalistic solutions to irrational tendencies. Of course, only further empirical research can determine whether our concerns are valid and, if so, which paternalistic policies in which domains pose the greatest moral and cognitive hazards.

An important benefit of this analytical exercise was an exposition of factors that should inform institutional design decisions aimed at addressing anomalous or irrational behavior. Simply enumerating these factors illustrated the complexity of paternalism, and, as David Shapiro argued several years ago, the many interests implicated by paternalistic interventions raise serious questions about the wisdom of courts engaging in ex post paternalism.¹⁰⁸ Yet courts seem increasingly willing to endorse the descriptions and prescriptions of behavioral law and economics.¹⁰⁹ If we are right that cognitive and moral hazards accompany at least some instances of ex post paternalism, then the courts would be well-advised to proceed much more cautiously and perhaps even to leave decisions about paternalistic intervention to legislatures, which should be better able to

108. Shapiro argues for the primacy of legislatures in matters of paternalism. David L. Shapiro, *Courts, Legislatures, and Paternalism*, 74 VA. L. REV. 519, 521 (1988).

109. See *O Centro Espirita Beneficente Uniao Do Vegetal v. Ashcroft*, 389 F.3d 973, 1016 (10th Cir. 2004) (McConnell, J., concurring) (“Notwithstanding the tendency of those trained in economics to view opportunity costs as equivalent to actual expenditures, modern social science research has confirmed the reality of ‘loss aversion’ . . . and the closely related ‘endowment effect.’”); *United States v. Lipscomb*, 299 F.3d 303, 335 (5th Cir. 2002) (“As behavioral law and economics warns us, inadequate information, biases, and heuristics often prevent individuals from acting rationally.”); *Abrahamson v. Bd. of Educ.*, 2002 WL 1354711, at *8 (S.D.N.Y. June 21, 2002) (“This phenomenon, called loss aversion or the endowment effect, might sway a teacher to retire at the end of the Option period, even though he might not have chosen early retirement at the beginning of the three years.”); Catherine M. Sharkey, *Punitive Damages: Should Juries Decide?*, 82 TEX. L. REV. 381, 382 n.5 (2003) (collecting cases citing the studies reported in CASS R. SUNSTEIN ET AL., *PUNITIVE DAMAGES: HOW JURIES DECIDE* (2002)); Neil Vidmar, *Experimental Simulations and Tort Reform: Avoidance, Error, and Overreaching in Sunstein et al.’s Punitive Damages*, 53 EMORY L.J. 1359, 1361 (2004) (noting that Judge Weinstein recently cited behavioral research on punitive damage decision making as social authority); see also *supra* notes 47, 59. Of course, courts may simply cite behavioral studies to support conclusions driven by other considerations, but in some cases the arguments of the behavioral law and economics scholars may possess some motivational force. Moreover, we do not mean to suggest that all of these invocations of behavioral economics raise the same moral and cognitive hazard concerns, but they do suggest a growing willingness on the part of at least some courts to incorporate behavioral economics into their opinions and perhaps even into their decision-making processes.

2006] *REGULATION OF IRRATIONALITY* 1663

consider competing values and marshal the evidence relevant to optimal institutional design.