

Libertarian Paternalism Is an Oxymoron

Gregory Mitchell*

In “Libertarian Paternalism Is Not an Oxymoron,” Professors Sunstein and Thaler set out to show that state control over the structure of choice sets can improve the welfare of citizens without reducing personal autonomy.¹ A public or private institution that adopts the perspective of the “libertarian paternalist” will “steer people’s choice in directions that will improve the choosers’ own welfare” but will not prescribe or proscribe any particular choices.² This limited regulation of choice behavior should be acceptable even to the committed libertarian, Sunstein and Thaler argue, once he understands the paternalism inherent in many choice settings: because

* Associate Professor, Florida State University College of Law, and Visiting Associate Professor, Vanderbilt School of Law. The author may be contacted by e-mail at gmitchel@law.fsu.edu or by regular mail at Florida State University College of Law, 425 West Jefferson Street, Tallahassee, Florida 32306-1601. I appreciate the helpful comments of Amitai Aviram, Paul Edelman, Chris Guthrie, Steve Hetcher, Adam Hirsch, Jon Klick, Mae Kuykendall, Peter Oh, Jim Rossi, Phil Tetlock, Mike Vandenberg, and particularly Richard Thaler on a draft.

¹ See Cass R. Sunstein & Richard H. Thaler, *Libertarian Paternalism Is Not an Oxymoron*, 70 U. CHI. L. REV. 1159, 1160 (2003) (“We propose a form of paternalism, libertarian in spirit, that should be acceptable to those who are firmly committed to freedom of choice on grounds of either autonomy or welfare.” (footnote omitted)). Despite its newness, the libertarian paternalism concept has already been cited to discount libertarian concerns about paternalistic proposals generated through behavioral economic analysis. See Michael Pereira, Book Review, 9 FORD. J. CORP. FIN. L. 715, 758 (2003) (“An emerging strain of scholarship offers a means of reconciling intervention in individual choices with a concern for autonomy. The idea, known variously as ‘asymmetric paternalism’ or ‘libertarian paternalism,’ suggests that planners (government or private) can increase social welfare by setting default rules that create benefits for those who make errors but impose little or no harm on those who are fully rational.” (footnote omitted)).

Thaler and Sunstein also published a shorter essay on libertarian paternalism for an economic audience. See Richard H. Thaler & Cass R. Sunstein, *Libertarian Paternalism*, 93 AM. ECON. REV. 175, 175 (2003). My comments focus on the more detailed discussion of libertarian paternalism found in the *University of Chicago Law Review* article.

² See Sunstein & Thaler, *supra* note 1, at 1162 (“The paternalistic aspect consists in the claim that it is legitimate for private and public institutions to attempt to influence people’s behavior even when third-party effects are absent. In other words, we argue for self-conscious efforts, by public and private institutions, to steer people’s choices in directions that will improve the choosers’ own welfare.”); *id.* at 1162 (“Libertarian paternalism is a relatively weak and nonintrusive type of paternalism, because choices are not blocked or fenced off.”).

Sunstein and Thaler encourage “planners” within both public and private institutions to design choice settings in ways that steer people in welfare-improving directions. See *id.* at 1162 & n. 14 (noting that their arguments apply with equal force to private and public planners). I will focus mostly on their arguments as applied to public planners because, as Sunstein and Thaler note, the libertarian’s “chief objection is to paternalistic law and government.” *Id.* at 1162.

preferences are unstable and sensitive to the way in which choices are framed, public and private institutions that control choice frames will inevitably shape the preferences of choosers.³

Sunstein and Thaler’s prototypical example of a libertarian paternalist policy is a 401(k) plan with the default option set to automatic enrollment to encourage participation but that permits employees to opt-out of this default enrollment.⁴ The plan capitalizes on the “stickiness” of default rules to move individuals in a direction that they would likely choose if they did not suffer from problems of low self-control and impulsive consumption tendencies.⁵ However, because individuals can easily opt out of the default option, the paternalism of the plan does not overwhelm the liberty of the individual who strongly prefers a different retirement savings plan or current consumption over future consumption. As Sunstein and Thaler emphasize, because some default must be chosen, and because many individuals are likely to remain irrationally with the default option, it is better to set the default to the welfare-enhancing choice than to be blind to the power of the default, and so long as individuals remain free to deviate from the default option, the libertarian should not be troubled by this weak form of paternalism.⁶ With this and

³ See Sunstein & Thaler, *supra* note 1, at 1166 (“Once it is understood that some organizational decisions are inevitable, that a form of paternalism cannot be avoided, and that the alternatives to paternalism (such as choosing options to make people worse off) are unattractive, we can abandon the less interesting question of whether to be paternalistic or not, and turn to the more constructive question of how to choose among the possible choice-influencing options.”). I intentionally use the masculine pronoun here because Richard Epstein serves as Sunstein and Thaler’s example of the committed libertarian. *See id.* at 1163 n.15.

⁴ *See id.* at 1159-60.

⁵ Or move them more quickly in the direction they are likely to move eventually. *See id.* at 1191 (“most employees do join the plan eventually, and very few drop out if automatically enrolled” (footnote omitted)). On the “stickiness” of defaults, see *id.* at 1174-77.

⁶ Or as Sunstein and Thaler put it:

As the savings problem illustrates, the design features of both legal and organizational rules have surprisingly powerful influences on people’s choices. We urge that such rules should be chosen with the explicit goal of improving the welfare of the people affected by them. The libertarian aspect of our strategies lies in the straightforward insistence that, in general people should be free to opt out of specified arrangements if they choose to do so. To borrow a phrase,

numerous other examples of how libertarianism may be reconciled with paternalism, Sunstein and Thaler hope to overthrow the “dogmatic anti-paternalism of numerous analysts of the law” that they see causing an “inept neglect” of many problematic areas of the law.⁷

In this essay, I discuss three defects present in the argument for libertarian paternalism: (1) a logical error and empirical oversight in the claim that paternalism is inevitable in situations where preferences exhibit irrational sensitivity to the choice frame; (2) a failure to justify the choice of welfare over liberty as the value guiding the paternalistic side of libertarian paternalism, and (3) a neglect of the redistributive effects of libertarian paternalism.⁸

libertarian paternalists urge that people should be “free to choose.” Hence, we do not aim to defend any approach that blocks individual choices.

See id. at 1161 (footnote omitted).

⁷ Sunstein & Thaler, *supra* note 1, at 1163, 1202.

⁸ The main empirical premise of libertarian paternalism, regarding the inevitable sensitivity of preferences to choice frames, is controversial, but even if we assume the correctness of this premise, problems remain. In fact, *only if* we assume the full or partial truth of this premise does the concept of libertarian paternalism become oxymoronic. That is, rather paradoxically, only if this empirical premise is completely wrong can Sunstein and Thaler validly claim that libertarian paternalism is not an oxymoron, for then the policies of the libertarian paternalist would not interfere with the personal autonomy of the decision-maker who does not exhibit the preference instability that Sunstein and Thaler posit nor would the policies have redistributive effects (i.e., if choice framing had no impact on choice, then the default options chosen by the central planner would be irrelevant). To be sure, the likelihood that a policy will do no harm should its behavioral assumptions be wrong is an attractive policy trait, but it is not the trait that Sunstein and Thaler emphasize. Sunstein and Thaler contend, of course, that their proposal will do no harm even if their behavioral assumptions are correct: central planners can reconcile libertarian and paternalist values by a careful framing of choice options that takes advantage of preference endogeneity to advance individual welfare without compelling individuals to make, or remain with, any particular choice.

Sunstein and Thaler correctly assume that normatively irrelevant features of a choice frame affect choices under some circumstances. However, there is now considerable evidence bringing into question the robustness of these effects. For a summary and discussion of this evidence with respect to framing effects, see Gregory Mitchell, *Taking Behavioralism Too Seriously? The Unwarranted Pessimism of the New Behavioral Analysis of Law*, 43 WM. & MARY L. REV. 1907, 1961-63, 2005-11 (2002); *see also* James N. Druckman, *The Implications of Framing Effects for Citizen Competence*, 23 POL. BEHAV. 225, 237 (2001) (“These results reveal the fragility and heterogeneity of equivalency framing effects: they occur less frequently than many believe, and the overall evidence for them is ‘mixed.’ This contradicts many portrayals that treat the effects as ‘pervasive’ and robust across people and issues.” (citations omitted)). For instance, simply adding information about different political parties advocating the different policy options used in the famous Asian disease problem (see Amos Tversky & Daniel Kahneman, *The Framing of Decisions and the Psychology of Choice*, 211 SCIENCE 453, 453 (1981)) reduces the classic framing effect in the problem of emphasizing lives saved versus lives lost (i.e., political party information swamps the usual framing effect). *See* James N. Druckman, *Using Credible Advice to Overcome Framing Effects*, 17 J.L. ECON. & ORG. 62, 73 (2001) (“party advice clearly reduced framing effects—in some cases, causing them to completely disappear” (footnote omitted)). And feelings of disgust may eliminate the endowment effect, while a feeling of

Consideration of the first two defects reveals that Sunstein and Thaler’s libertarian paternalism surrenders too much libertarian ground to the paternalist. Consideration of the third defect reveals that any form of libertarian paternalism, even the more truly *libertarian* paternalism proposed in Parts I and II below, may lead to a redistribution of resources from rational to irrational persons that cannot be reconciled with the libertarian prohibition on state-based takings for any purpose other than remedying involuntary exchanges. I conclude that, despite the ambitious efforts of Sunstein and Thaler to show otherwise, they offer no new regulatory path that will permit paternalistic efforts at welfare improvement without intruding on personal autonomy: in short, libertarian paternalism is an oxymoron.

This critical exercise serves two constructive purposes. First, the essay brings out some of the difficult issues that must be confronted when the government considers whether and how to regulate irrational behavior, issues that are becoming much more prominent with the rise of behavioral law and economics and its many paternalistic prescriptions.⁹ Second, the essay shows

sadness can induce a “reverse endowment effect.” See Jennifer S. Lerner, Deborah A. Small & George Loewenstein, *Heart Strings and Purse Strings: Carryover Effects of Emotions on Economic Decisions*, 15 *PSYCHOL. SCI.* 337, 340 (2004) (“Overall, the pattern of results supports the hypotheses that disgust triggers goals to expel, reducing buying and selling prices, whereas sadness triggers the goal of changing one’s circumstances, increasing buying prices but reducing selling prices. The effects are sufficiently strong that in one case (disgust) they eliminate the endowment effect, and in the other case (sadness) they actually reverse it.”). Most importantly for present purposes, methods exist for diminishing the effects of choice frames and, even without debiasing efforts, there are significant individual differences in framing effects, and it is this superability and variation in the irrational effects of choice frames that turn out to be most significant to the libertarian paternalism argument. I return to these points below.

In addition, I set to the side debate about the normative status of seemingly arbitrary features of the choice frame. See e.g., Craig McKenzie, *Judgment and Decision Making*, in *HANDBOOK OF COGNITION* (K. Lamberts & R.L. Goldstone eds., forthcoming) (“[R]ather than indicating deep irrationality, framing effects (or at least attribute framing effects) appear to be the result of both speakers and listeners exploiting regularities in language in an adaptive way.”) (manuscript at 17, on file with author). Therefore, I assume that it is irrational to allow features of the choice frame considered arbitrary within rational choice theory (e.g., whether options are phrased as losses or gains relative to a reference point, the order of presentation of options, or the presence or absence of dominated options) to affect choices.

⁹ For instance, Jeff Rachlinski recently noted that “virtually every scholar who has written on the application of psychological research on judgment and choice to law has concluded that cognitive psychology supports institutional constraint on individual choice.” Jeffrey J. Rachlinski, *The Uncertain Psychological Case for*

that, while fidelity to libertarian principles leaves little room for the government to regulate irrational behavior, there are some forms of irrationality regulation more congenial to libertarian principles than Sunstein and Thaler’s version of libertarian paternalism.¹⁰

I. The Evitability of Choice-Framing Paternalism

Three premises form the core of Sunstein and Thaler’s argument for libertarian paternalism: (1) in many situations, individuals lack stable or well-defined preferences¹¹; (2) in these situations, the manner in which choice sets are presented to individuals can greatly affect the choices made (i.e., preferences are endogenous to the way in which choices are presented or framed)¹²; (3) as a result of this preference endogeneity, (a) public and private institutions that control the structuring of choice sets possess an unavoidable power to shape individuals’ labile

Paternalism, 97 NW. U.L. REV. 1165, 1166 (2003). Rachlinski also catalogs the numerous paternalistic proposals found within the behavioral law and economics scholarship. *See id.* at 1177-1206.

¹⁰ Sunstein and Thaler’s libertarian paternalism may be contrasted with the concept of “asymmetric paternalism” advanced by Camerer and his colleagues and my own “do no harm” approach to irrationality regulation. *See* Colin Camerer et al., *Regulation for Conservatives: Behavioral Economics and the Case for “Asymmetric Paternalism,”* 151 U. PA. L. REV. 1211 (2003); Gregory Mitchell, *Why Law and Economics’ Perfect Rationality Should Not be Traded for Behavioral Law and Economics’ Equal Incompetence*, 91 GEO. L.J. 67 (2002). Camerer and colleagues discuss several forms of irrationality regulation similar to the reforms advocated by Sunstein and Thaler, including the design of default rules, *see* Camerer et al., *supra*, at 1224-30, but make no strong claim about the libertarian “friendliness” of asymmetric paternalism: while some forms of asymmetric paternalism will not greatly intrude on personal autonomy, other forms will be fairly intrusive. *See id.* at 1224 (noting that the four types of regulation discussed—default rules, provision or reframing of information, cooling-off periods, and limiting consumer choices—were “ranked roughly in increasing order of departure from pure asymmetric paternalism—i.e., the increasing ‘heavy-handedness’ of the policy”). The main virtue of asymmetric paternalism is that it is a focused form of paternalism that “helps those whose rationality is bounded from making a costly mistake and harms more rational folk very little.” *Id.* at 1254. The “do no harm” approach that I advocate focuses on procedural measures that may be implemented without altering the substantive rights of the parties to help people overcome their cognitive biases. *See* Mitchell, *supra*, at 132-37. In contrast, and as discussed in Part I below, Sunstein and Thaler treat irrational choice behaviors as if they are insuperable features of many choice situations that must be accepted and worked around to move people in positive directions, even if this means altering the substantive rights and interests of the affected individuals.

¹¹ *See* Sunstein & Thaler, *supra* note 1, at 1161 (“in many domains, people lack clear, stable, or well-ordered preferences”).

¹² *See id.* (“What [people] choose is strongly influenced by details of the context in which they make their choice, for example default rules, framing effects (that is, the wording of possible options) and starting points. These contextual influences render the very meaning of the term ‘preferences’ unclear.”); *id.* at 1164 (“If the arrangement of alternatives has a significant effect on the selections the customers make, then their true ‘preferences’ do not formally exist.”).

preferences¹³ and (b) individual choices will not reliably advance individual welfare unless choice frames push individuals to select the welfare-advancing option (call this “choice-framing paternalism”).¹⁴ If true, these propositions undercut the key libertarian assumption that individuals are the best judges and protectors of their own welfare.¹⁵

¹³ See *id.* at 1164 (“The first misconception is that there are viable alternatives to paternalism. In many situations, some organization or agent *must* make a choice that will affect the behavior of some other people. There is, in those situations, no alternative to a kind of paternalism—at least in the form of an intervention that affects what people choose.”); *id.* at 1166 (“We suggest that because of the likely effects of default rules, framing effects, and starting points on choices and preferences, paternalism, at least in a weak sense, is impossible to avoid.”).

¹⁴ See *id.* at 1162 (“In other words, we argue for self-conscious efforts, by private and public institutions, to steer people’s choices in directions that will improve the chooser’s own welfare. In our understanding, a policy therefore counts as ‘paternalistic’ if it attempts to influence the choices of affected parties in a way that will make choosers better off.” (footnote omitted)).

¹⁵ See *id.* at 1163 (“The false assumption is that almost all people, almost all of the time, make choices that are in the best interest or at the very least are better, by their own lights, than the choices that would be made by third parties.”); *id.* at 1167 (“The presumption that individual choices should be respected is usually based on the claim that people do an excellent job of making choices, or at least that they do a far better job than third parties could possibly do.”).

For an article discussing the foundations of libertarianism, Sunstein and Thaler spend remarkably little time discussing those foundations or providing support for their views of those foundations. The reason may be, as James Child notes, “there is no complete libertarian system in the form that Rawls, Gauthier, or Raz have produced systems of political philosophy Thus, anyone who wishes to analyze the first principles of a libertarian political philosophy is left to reconstruct it him- or herself.” James W. Child, *Can Libertarianism Sustain a Fraud Standard?*, 104 *ETHICS* 722, 724 (1994). Child designates self-ownership and private ownership of property as the basic postulates of libertarianism, and lists rights of defense of oneself and one’s property against forceful and coercive attacks as second-order rights flowing from the first principles. See *id.* at 725-28. In support of Sunstein and Thaler’s view on the importance of rationality to the libertarian system, Child notes that libertarians often implicitly assume that citizens possess a “general competence,” or “have sufficient capacities to engage in practical reasoning and to be moral agents.” *Id.* at 729. Child decomposes “general competence” into deliberative and volitional capacities, which includes a requirement of stable preferences:

By relying on much recent work on the notion of general competence and the related notion of a capacity for autonomy, we can formulate the constituents of the competence to enter market transactions. They would include the following deliberative capacities: (1) to acquire, understand and appraise information, which includes considering its probability of truth or falsity and its relevance; (2) to entertain a stable set of preferences by which choice among various options with various payoffs can be made; and (3) in light of this information and these values, to consider choices and weigh the possible risks, costs, and benefits of those choices. This includes the risk and cost of acting on false information. These would be combined with the broadly volitional capacities to arrive at decisions on the basis of those deliberations and act in accord with those decisions.

Id. at 729-30 (footnote omitted).

If the libertarian accepts this syllogism for the inevitability of choice-framing paternalism, the question turns from whether institutions should ever be paternalistic to what forms of paternalism will do the least damage to libertarian values.¹⁶ Sunstein and Thaler argue that the best solution for the libertarian, in light of the inevitability of choice-framing paternalism, is to direct central planners to frame choices in ways that favor the welfare-advancing choice, but not requiring that people make this “better” choice and allowing them to opt out of the planner-chosen option in situations where a default position must be chosen for people.¹⁷ It is this blending of welfare management by planners with the preservation of an individual’s freedom to opt out of the planner’s utopian vision that results in the concept of libertarian paternalism.¹⁸

Fortunately for the libertarian, Sunstein and Thaler’s syllogism for the inevitability of choice-framing paternalism contains an error. The logical implication of the claim that normatively irrelevant features of the choice setting influence preferences is not the inevitability

¹⁶ See Sunstein & Thaler, *supra* note 1, at 1166 (“Once it is understood that some organizational decisions are inevitable, that a form of paternalism cannot be avoided, and that the alternatives to paternalism (such as choosing options to make people worse off) are unattractive, we can abandon the less interesting question of whether to be paternalistic or not, and turn to the more constructive question of how to choose among the possible choice-influencing options.”).

¹⁷ See *id.* at 1161 (“[T]he design features of both legal and organizational rules have surprisingly powerful influences on people’s choices. We urge that such rules should be chosen with the explicit goal of improving the welfare of the people affected by them. The libertarian aspect of our strategies lies in the straightforward insistence that, in general, people should be free to choose to opt out of specified arrangements if they choose to do so.”); *id.* at 1162 (“Libertarian paternalism is a relatively weak and nonintrusive type of paternalism, because choices are not blocked or fenced off.”); *id.* at 1175 (“So long as people can contract around the default rule, it is fair to say that the legal system is protecting freedom of choice, and in that sense complying with libertarian goals.”).

¹⁸ Sunstein and Thaler’s argument for libertarian paternalism is a variant of the general argument used to justify “soft” paternalism. See Thaddeus Mason Pope, *Counting the Dragon’s Teeth and Claws: The Definition of Hard Paternalism*, 20 GA. ST. U. L. REV. 659, 667 (2004) (“The sort of soft paternalism that is justified on the basis that the subject lacks the requisite decision-making capacity to engage in the restricted conduct is described as either soft or weak paternalism.” (footnotes omitted)). The unique features of the Sunstein and Thaler argument involve the claim that soft paternalism is inevitable in a wide range of settings and the prescriptive direction to the central planner to design choice settings to enhance welfare but also permit individuals to opt out of this welfare-enhancing scheme.

of paternalism but rather the inevitability of manipulation of choices by central planners *so long as individuals remain subject to these irrational influences*. Thus, it is only if there is no way for individuals to overcome these irrational influences that state control over citizen choice is inevitable.¹⁹ To restore the logic of their syllogism, Sunstein and Thaler would need to add a premise stating that choice framing effects are insuperable, but closer inspection of Sunstein and Thaler’s argument and a consideration of additional empirical evidence reveal that such a premise is not warranted.

First, Sunstein and Thaler acknowledge that “true” preferences control choices in some settings, despite ever-present choice framing effects, but they fail to delimit clearly those situations where preferences are endogenous to the choice frame from those where they are not.²⁰ Certainly Sunstein and Thaler paint a picture that highlights the pervasiveness of choice framing effects: they write that “in many domains” what people choose “is strongly influenced by details of the context in which they make their choice,”²¹ and tell us that experience, expertise, and

¹⁹ There exists another logical error in Sunstein and Thaler’s argument—not all forms of choice manipulation will be paternalistic in nature (i.e., the more powerful party will not necessarily seek to manipulate choices in ways that benefit the dominated party)—but Sunstein and Thaler recognize that a central planner could manipulate choices in non-paternalistic ways. See Sunstein & Thaler, *supra* note 1, at 1164 (discussing the range of choice manipulations available). The paternalism they speak of has more to do with the way they counsel central planners to act in light of this manipulative power: “the approach we recommend nonetheless counts as paternalistic, because public and private planners are not trying to track people’s anticipated choices, but are self-consciously attempting to move people in welfare-promoting directions.” *Id.* at 1162 (footnote omitted); see also *id.* at 1188 (“Minimal paternalism is the form of paternalism that occurs whenever a planner (private or public) constructs a default rule or starting point with the goal of influencing behavior. So long as it is costless or nearly costless to depart from the default plan, minimal paternalism is maximally libertarian. This is the form of paternalism that we have described as inevitable.”). This choice in labeling by Sunstein and Thaler is not a serious error once it is understood that the paternalism Sunstein and Thaler refer to is the paternalism of a planner who follows their recommendations; otherwise, their premises only support the conclusion about the inevitability of choice manipulation by central planners.

²⁰ Sunstein and Thaler equate true preferences with stable preferences. See *id.* at 1164 (“Across a certain domain of possibilities, consumers will often lack well-formed preferences, in the sense of preferences that are firmly held and preexist the director’s own choices about how to order the relevant items. If the arrangement of the alternatives has a significant effect on the selections the customers make, then their true “preferences” do not formally exist.”).

²¹ Sunstein & Thaler, *supra* note 1, at 1161; see also *id.* (indicating that problems arising from a lack of well-defined preferences occur in “many situations involving the behavior of workers and consumers”).

incentives cannot be counted on to protect decision-makers from the influence of the choice frame.²² But they acknowledge that “[i]n some domains, consumers and workers are highly informed—so much so that they will not even be influenced by default rules,”²³ and they suggest that reflection about one’s choices can align choice behavior with true preferences.²⁴ Indeed, were Sunstein and Thaler to take a more radical position on the endogeneity of preferences, then

²² See *id.* (“Repeated experiences with such problems might be expected to eliminate this framing effect, but doctors too are vulnerable to it.” (footnote omitted)); *id.* at 1168-69 (“It is possible to raise questions about some of these findings and to think that people may do a better job of choosing in the real world than they do in the laboratory. But studies of actual choices reveal many of the same problems, even when the stakes are high.” (footnotes omitted)); *id.* at 1170 (“It is also true that learning frequently enables people to overcome their own limitations. But many of the most important decisions (for example, buying a home or choosing a spouse) are made infrequently and typically without the aid of impartial experts.”). For contrary evidence and discussion of the debates on the role of education, experience, and incentives on rational behavior, see Gregory Mitchell, *Tendencies Versus Boundaries: Levels of Generality in Behavioral Law and Economics*, 56 VAND. L. REV. 1781, 1787-90 (2003); Mitchell, *supra* note 10, at 114-19.

²³ *Id.* at 1197. Also, when discussing research by Waldfogel finding that individuals value their own purchases more highly than gifts from third parties (i.e., evidence that even if individuals are not always good judges of their welfare, they were better judges of their own welfare than third parties, at least in this study), Sunstein and Thaler write that “Waldfogel is studying the context of ordinary consumer purchases, in which people are in an especially good position to know what they like. We are focusing on less familiar situations, which present special puzzles.” *Id.* at 1169 n.31 (emphasis added).

²⁴ Sunstein and Thaler analogize to the reflective benefits of cooling-off periods to justify the inference of “true” preferences from a failure to opt out of the default set by a libertarian paternalist. See *id.* at 1191 (“[I]t is reasonable to think that if, on reflection, workers realized that they had been ‘tricked’ into saving too much, they might take the effort to opt out. The fact that very few participants choose to opt out supports (though it does not prove) the claim that they are helped by a system that makes joining easy.”).

Sunstein and Thaler offer this argument about the benefits of reflection in connection with their suggestion that actual choice behavior might be some guide to what choices promote welfare. See *id.* at 1191 (“It is also possible to make inferences from actual choices about welfare. For example, most employees do join the plan eventually, and very few drop out if automatically enrolled. These facts suggest, that, at least on average, defaulting people into the plan will mostly hasten the rate at which people join the plan, and that the vast majority of those who are so nudged will be grateful.”). This approach to welfare measurement conflicts with their premise that choice behavior does not reliably reflect true preferences or promote welfare. See *id.* at 1174 (“[W]hat if many or most employees do not have stable or well-formed preferences, and what if employee choices are inevitably a product of the default rule? In such cases, it is meaningless to ask what most employees would do.”); *id.* at 1193 (“What people choose often depends on the starting point, and hence the starting point cannot be selected by asking what people choose.”). Recognizing this, Sunstein and Thaler offer the ultimately self-defeating argument that persons “‘tricked’” into making the libertarian paternalistic choice will, “on reflection,” realize that these choices were good choices and remain with the default option. *Id.* Of course, if reflection is effective at revealing true preferences and causing people to make good choices, then Sunstein and Thaler’s concerns about the inevitability of choice-framing paternalism largely disappear. That is, Sunstein and Thaler offer no reason to believe that reflection only works when individuals have first been put onto the libertarian paternalist path toward goodness but not some other path. If, given some time for reflection, individuals will make a “true” choice regardless of default option, then default options are not very sticky or troublesome unless serious harm can occur during this unspecified period of time in which reflection will work to reveal true preferences.

there would be no room left in their argument for libertarian principles, with their emphasis on self-ownership,²⁵ because radical endogeneity of preferences leaves no core values or stable preferences to comprise the self.²⁶

What is missing from Sunstein and Thaler's argument, and from behavioral law and economics in general,²⁷ is a theory of when choice frames will control choice and when they will not. Sunstein and Thaler do suggest some choice settings that favor the expression of stable over transient preferences: choices made frequently,²⁸ choices involving less emotion,²⁹ choices on

²⁵ See, e.g., Child, *supra* note 15, at 725 ("Almost all libertarians begin with some claim of self-ownership.").

²⁶ By suggesting that some preferences are "true," Sunstein and Thaler seem to distance themselves from the more radical implications of the research on preference lability:

The lability of preferences implied by the demonstration of framing and elicitation effects raises difficult questions concerning the assessment of preferences and values. In the classical analysis, the relation of preference is inferred from observed responses (e.g., choice matching) and is assumed to reflect the decision maker's underlying utility or value. But if different elicitation procedures produce different ordering of options, how can preferences and values be defined? And in what sense do they exist? To be sure, people make choices, set prices, rate options and even explain their decisions to others. Preferences, therefore, exist as observed data. However, if these data do not satisfy the elementary requirements of invariance, it is unclear how to define a relation of preference that can serve as a basis for the measurement of value. In the absence of well-defined preferences, the foundations of choice theory and decision analysis are called into question.

Amos Tversky, Shmuel Sattah & Paul Slovic, *Contingent Weighting in Judgment and Choice*, in CHOICES, VALUES, AND FRAMES 503, 517 (Daniel Kahneman & Amos Tversky eds., 2000).

²⁷ Elsewhere I discuss the heavy reliance on *ceteris paribus* clauses within judgment and decision-making research and the mischief these clauses can play when trying to translate behavioral economic findings into policy prescriptions. See Mitchell, *supra* note 22, at 1797-1811. Cf. Edward L. Glaeser, *Psychology and the Market*, 94 AM. ECON. REV. 408, 408 (2004) ("To understand real-world outcomes, one must understand the supply of situational factors. Economics, not psychology, has the tools to understand how market outcomes will reflect the interaction of interested suppliers of influence and consumers who then respond to that influence."); *id.* at 409 ("[S]ituationalism creates just as much of a problem for psychology as for economics. Real-world situations are man-made. Economics, not psychology, provides a framework for understanding the supply of framing, influence, and situation."); Chris Swoyer, *Judgment and Decision-Making: Extrapolations and Applications*, in JUDGMENTS, DECISIONS, AND PUBLIC POLICY 9, 11-12 (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 starting to ask: *Who* reasons in which ways?").

²⁸ Although Sunstein and Thaler suggest at one point that repetition is not a good check on choice framing effects, see *supra* note 22 and accompanying text, later they suggest that repetition may be helpful, see Sunstein & Thaler, *supra* note 1, at 1163 ("it seems reasonable to say that people make better choices in which they have experience and good information (say, choosing ice cream flavors) than in contexts in which they are inexperienced and poorly

which one has time to reflect,³⁰ choices involving options about which one is highly informed,³¹ and choices involving a small number of options.³² Presumably, then, the central planner who subscribes to libertarian paternalism should steer clear of choice settings with these features (or rank them low in priority for intervention), but these are very loose guidelines for the libertarian paternalist. The problem here, however, is not that the lack of a theory of dominant choice frames makes a slippery slope even more slippery, for, as Sunstein and Thaler note, the over-eager libertarian paternalist should do no great harm to persons with manipulation-proof, true preferences so long as the planner builds in an escape hatch of some sort (though a potential redistribution problem remains, as discussed in Part III).³³

Rather, the real problem with the lack of a theory about when choice frames control choice is that we lose information that might educate us on a direct approach to improving decision-making, by which we might foster individual freedom of choice rather than have the central planner make choices for people.³⁴ In fact, to persons who lack stable preferences and

informed (say, choosing among medical treatments or investment options.”); *see also id.* at 1189 (suggesting that choices made in unfamiliar settings, with a lack of experience, are worse than otherwise).

²⁹ *See id.* at 1188 (discussing benefits of “cooling off” periods); *id.* at 1189 (indicating that choices presenting a “risk of impulsiveness” are likely to be worse than otherwise).

³⁰ *See supra* note 24.

³¹ *See id.* at 1197-98 (discussing information effects);

³² *See id.* at 1198 (discussing number of options). In addition, Hoch and Loewenstein discuss the role of certain environmental factors in impulsive consumer choices. *See* Stephen J. Hoch & George F. Loewenstein, *Time-Inconsistent Preferences and Consumer Self-Control*, 17 J. CONSUMER RES. 492, 496-98 (1991) (discussing physical proximity, temporal proximity, and social comparisons).

³³ *See* Sunstein & Thaler, *supra* note 1, at 1199 (“the libertarian condition [on libertarian paternalist policies], requiring opt-out rights, greatly limits the steepness of the slope”).

³⁴ The cognitive psychologist Keith Stanovich has begun to develop a theory of the situations that pose the greatest difficulty for rational thought in light of what he calls “fundamental computational biases,” or processing defaults, of human cognition (which Stanovich lists as the tendency to contextualize problems with accessible information, tendency to “socialize” problems, tendency to infer deliberative design and pattern, and tendency toward a narrative mode of thought). *See* KEITH E. STANOVICH, *THE ROBOT’S REBELLION* 110-13 (2004). In particular, Stanovich

the means to decide for themselves what their true preferences are in a particular choice setting, the opt-out escape hatch that Sunstein and Thaler trumpet as the libertarian saving grace of their proposal will be meaningless. That is, if choice is as sticky as Sunstein and Thaler claim,³⁵ then people who lack the means to determine their true preferences in a given setting cannot make any real use of the opt-out provision and will simply stick with the default set by the libertarian paternalist.³⁶ For these people, libertarian paternalism is just paternalism.³⁷

theorizes that situations requiring radical decontextualization or abstraction to reach a correct solution pose considerable problems in light of the tendency to reinterpret new problems in terms of familiar problems:

[M]odern technological societies continually spawn situations where humans must decontextualize information—where they must deal abstractly and in a personalized manner with information. Such situations require the active suppression of the social, narrative, and contextualizing styles that characterize the operation of [the autonomous set of systems]. These situations may not be numerous, but they tend to be in particularly important domains of modern life—indeed, they in part *define* modern life in postindustrial knowledge-based societies.

Id. at 122. Some of the choice situations that Sunstein and Thaler describe as problematic would seem to fit Stanovich’s theory (planning for retirement, perhaps), but others would not (dieting choices are quite familiar and, though we may want to engage in self-deception about the effects of eating on our health, it does not require much abstract or decontextualized thought to anticipate the consequences of overeating). Stanovich’s work suggests the utility of techniques that shift thought into a more analytical mode. *See, e.g., id.* at 173 (“Humans thus escape the clutches of the selfish genes by getting their analytic systems in control, by developing override capabilities in instances where the responses primed by [the autonomous set of systems] and the analytic system conflict.”). I discuss a few of these techniques in the text. For a differing view on the importance of decontextualization, see Craig R.M. McKenzie, *Rational models as theories—not standards—of behavior*, 7 *TRENDS IN COGNITIVE SCI.* 403, 405 (2003) (“depending on a purely logical analysis will not get you very far in the real world, where context, meaning and relevance, rather than pure structure, are crucial” (citations omitted)).

³⁵ *See, e.g., id.* at 1171 (“Throughout we shall be drawing attention to the effects of default plans on choice. Often those plans will be remarkably ‘sticky.’”).

³⁶ Curiously, as noted above, Sunstein and Thaler suggest that we should consider the failure of persons to opt out after a chance for reflection as some evidence that the persons find the libertarian paternalist default favorable. *See supra* note 24. Yet if choice frame effects are so weak and temporary that “true” preferences will reveal themselves after some passage of time or some chance for (unprompted) reflection, then the importance of framing effects and the power of Sunstein and Thaler’s argument are greatly reduced. As noted previously, it may in fact be the case that framing effects are very weak, *see supra* note 8, but that is certainly not the position that Sunstein and Thaler advocate. Indeed, elsewhere Thaler makes clear that the libertarian paternalist Save More Tomorrow plan for increasing retirement contributions (the “Smart” program) counts on cognitive biases to work. *See* Richard H. Thaler & Shlomo Benartzi, *Save More Tomorrow™: Using Behavioral Economics to Increase Employee Saving*, 112 *J. POL. ECON.* S164, S170-71 (2004) (“In this way, inertia and status quo bias work toward keeping people in the plan.”).

³⁷ To be sure, it is a paternalism that seeks to make them better off, but then that is the very definition of paternalism and thus should not be very comforting to the libertarian.

Second, empirical work that Sunstein and Thaler do not discuss does provide some guidance on how we might enhance freedom of choice by making persons less susceptible to choice set manipulations.³⁸ In considering these debiasing techniques, we need to distinguish between errors that arise from the choice context itself (“contextualization problems”) and errors that arise from the decision-maker’s weakness of will or myopic focus on the present (“self-control problems”). The first category of problems contains unwanted influences on choices due to framing effects, anchoring effects, the status quo bias, and the endowment effect, and the second category of problems involves self-control problems that arise from an overweighting of short-term interests relative to long-term interests or an inability or unwillingness to delay gratification.³⁹

Two non-exclusive approaches may be taken to combat contextualization problems. The first approach encourages people to reframe choice problems into more personally meaningful frames. For instance, asking decision-makers to imagine ways that a chosen option might fail or an ignored option might succeed can remove or reduce the effects of several biases (although asking individuals to generate many alternative futures may actually exacerbate the underlying biases if individuals have difficulty generating counterfactuals).⁴⁰ The second approach

³⁸ For a general discussion of individual-level debiasing measures, as opposed to measures that debias the environment and measures that take advantage of existing biases to direct people in certain directions (i.e., the libertarian paternalism approach, or what Larrick calls “rebiasing”), see generally Richard P. Larrick, *Debiasing*, in BLACKWELL HANDBOOK ON JUDGMENT AND DECISION MAKING (Derek J. Koehler & Nigel Harvey eds., forthcoming 2004) (manuscript on file with author). Consistent with the view that an emphasis on individual-level debiasing should be more congenial to the libertarian focused on personal autonomy and free choice, Larrick notes that debiasing at the individual level “tends to increase [individuals’] decision skills and their ability to apply those skills to new decision domains” *Id.* at 3.

³⁹ Elsewhere, Sunstein and Thaler, along with Jolls, discuss effects falling within the first category as examples of bounded rationality and effects falling within the second category as examples of bounded willpower. See Christine Jolls, Cass R. Sunstein & Richard H. Thaler, *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471, 1477, 1479 (1998).

⁴⁰ See Lawrence J. Sanna & Norbert Schwarz, *Integrating Temporal Biases: The Interplay of Focal Thoughts and Accessibility Experiences*, 15 PSYCHOL. SCI. 474, 480 (2004) (discussing how generating thoughts about success and

encourages greater reflection and deliberation, with the aim of shifting persons from an intuitive into a more analytical mindset that will cause people to examine choices more closely and apply more formal rules for solving problems. For instance, counseling decision-makers to consider counter-arguments and opposing viewpoints is one of the most successful debiasing measures.⁴¹

failure may eliminate or reduce temporal biases, such as overconfidence, the planning fallacy, the impact bias, and the hindsight bias but noting that “debiasing may fail when people try to generate more alternatives than they can easily accomplish”); Lawrence J. Sanna et al., *When Debiasing Backfires: Accessible Content and Accessibility Experiences in Debiasing Hindsight*, 28 J. EXP’L PSYCHOL.: LEARNING MEMORY & COGNITION 497 (2002) (finding that generating two alternative outcomes slightly attenuated the hindsight bias, but generating ten alternatives amplified the bias).

⁴¹ Asking or directing experimental subjects to consider alternative or opposing arguments, positions, or evidence has been found to ameliorate the adverse effects of several biases, including the primacy or anchoring effect, biased assimilation of new evidence, biased hypothesis testing, the overconfidence phenomenon, the explanation bias, the self-serving bias, and the hindsight bias. See Hal R. Arkes, David Faust, Thomas J. Guilmette & Kathleen Hart, *Eliminating the Hindsight Bias*, 73 J. APPLIED PSYCHOL. 305, 307 (1988) (“The debiasing procedure previously used by Koriat et al. (1980) to reduce overconfidence in answers to general questions significantly reduced the number of subjects manifesting the hindsight bias.”); Linda Babcock, George Loewenstein, & Samuel Issacharoff, *Creating Convergence: Debiasing Biased Litigants*, 22 LAW & SOC. INQUIRY 913, 920 (1998) (“When subjects consciously considered the weaknesses in their case or reasons that the judge might rule against them, their judgments exhibited no self-serving bias.”); Gretchen Chapman & Eric J. Johnson, *Anchoring, Activation, and the Construction of Values*, 79 ORG’L BEHAV. & HUMAN DECISION PROCESSES 115, 144 (1999) (“Experiments 1, 2, and 4 demonstrated that a prompt to consider reasons or features different from the anchor reduced or eliminated the anchoring effect.”); *id.* (“The present experiments demonstrate a method to reduce or eliminate the anchoring bias. Prompting subjects to identify target features different from the anchor (Experiment 1) or to consider reasons at odd with anchor (Experiment 2 and 4) greatly reduced the anchoring bias. This debiasing result is important because previous efforts to debias anchoring have often met with little success.”); Adam D. Galinsky & Thomas Mussweiler, *First Offers as Anchors: The Role of Perspective-Taking and Negotiator Focus*, 81 J. PERS. & SOC. PSYCHOL. 657, 665-66 (2001) (discussing advantages of making first offer in negotiations and perspective-taking and focusing measures that opposing party can employ to eliminate or prevent the anchoring effect from the first offer); Asher Koriat, Sarah Lichtenstein & Baruch Fischhoff, *Reasons for Confidence*, 6 J. EXP’L PSYCH.: HUMAN LEARNING & MEMORY 107, 116 (1980) (“We have shown that calibration can be significantly improved by requiring people to explicate all considerations that seem pertinent to their decision.”); Charles G. Lord, Mark R. Lepper & Elizabeth Preston, *Considering the Opposite: A Corrective Strategy for Social Judgment*, 47 J. PERS. & SOC. PSYCHOL. 1231, 1239 (1984) (“In two different domains of social judgment, biased assimilation of evidence and biased hypothesis testing, and with two different inducements, direct instructions and indirect manipulation of accessibility through stimulus salience, the cognitive strategy of considering opposite possibilities promoted impartiality.” (footnote omitted)); Thomas Mussweiler, Fritz Strack & Tim Pfeiffer, *Overcoming the Inevitable Anchoring Effect: Considering the Opposite Compensates for Selective Accessibility*, 26 PERS. & SOC. PSYCHOL. BULL. 1142, 1149 (2000) (“The Selective Accessibility Model . . . suggests that increasing the accessibility of anchor-inconsistent knowledge reduces the magnitude of the anchoring bias. Considering the opposite appears to be one way to achieve this objective and consequently proved to be an effective corrective strategy.”); Edward R. Hirt & Keith D. Markman, *Multiple Explanation: A Consider-an-Alternative Strategy for Debiasing Judgments*, 69 J. PERS. & SOC. PSYCHOL. 1069 (1995) (addressing the explanation bias, which refers to the tendency for subjective likelihood estimates of potential events occurring to increase after participants provide an explanation for why such potential events might occur); Gregory H. Mumma & Steven B. Wilson, *Procedural Debiasing of Primacy/Anchoring Effects in Clinical-Like Judgments*, 51 J. CLIN. PSYCHOL. 841, 850 (1995) (“A consider-the-opposite debiasing strategy that involved sorting the cues according to diagnosticity while reviewing the cues was highly successful in debiasing single-cue anchoring.”); see also Edward R. Hirt et al., *Activating a Mental Simulation Mind-Set Through Generation of*

Likewise, simply asking people to give reasons for their choices can reduce the influence of gain/loss framing effects.⁴²

With respect to temporal focus and self-control problems, measures that induce greater self-monitoring and encourage individuals to consider future prospects more carefully can result in more rational decisions.⁴³ Thus, including a mandatory budgeting exercise in retirement planning decisions may broaden the temporal focus and increase self-monitoring, and hence self-control. Likewise, making available self-guided decision aids may help individuals identify and weigh criteria relevant to an important decision and clarify internal conflicts.⁴⁴

Alternatives: Implications for Debiasing in Related and Unrelated Domains, 40 J. EXPERIMENTAL SOC. PSYCHOL. 374, 381-82 (discussing importance of individual differences in need for structure in the role of alternatives generation and use of the strategy across domains); Merrie Jo Stallard & Debra L. Worthington, *Reducing the Hindsight Bias Utilizing Attorney Closing Arguments*, 22 LAW & HUMAN BEHAV. 671, 680 (1998) (“Having the defense attorney argue to focus on the preoutcome time frame acted as an effective reduction tool, providing participants with an alternative story upon which they could draw.”). *But see* Mumma & Wilson, *supra*, at 850 (“sequence anchoring was resistant to debiasing by all three interventions,” including the consider-the-opposite strategy). To the extent that encouraging decision-makers to consider opposing viewpoints suggests that they consider ways their choices may fail or omitted choices may succeed, this “consider-the-alternatives” approach includes the “imagine failure and success approach.

⁴²See, e.g., Paul M. Miller & N.S. Fagley, *The Effects of Framing, Problem Variations, and Providing Rationale on Choice*, 17 PERS. & SOC. PSYCHOL. BULL. 517, 521 (1991) (“Asking subjects to ‘briefly explain your rationale’ was believed to be a milder intervention than say, asking them to provide a justification for their choice. Yet it had a great impact on the occurrence of framing effects.”); Winston Sieck & J. Frank Yates, *Exposition Effects on Decision Making: Choice and Confidence in Choice*, 70 ORG’L BEHAV. & HUMAN DECISION PROCESSES 207, 217 (1997) (“In all three experiments, we observed two basic consequences of exposition, a choice effect and a confidence effect. In the choice effect, exposition reduced the influence of alternative frames. The confidence effect was such that exposition bolstered subjects’ beliefs that their choices were indeed the best.”).

⁴³ See, e.g., Roy F. Baumeister, *Yielding to Temptation: Self-Control Failure, Impulsive Purchasing, and Consumer Behavior*, 28 J. CONSUMER RES. 670, 672 (2002) (“When people lose track of their behavior, self-control breaks down.”); *id.* (“When people keep careful track of their money and expenditures, impulsive purchases are less likely.”); Daniel T. Gilbert et al., *The Future Is Now: Temporal Correction in Affective Forecasting*, 88 ORG’L BEHAV. & HUMAN DECISION PROCESSES 430, 441 (2002) (“One way to combat shortsightedness, then may be to encourage people to represent future events as though they were happening in the present (‘Imagine learning you have emphysema’) and to then adjust for the event’s actual temporal location (‘How would your reaction change if you learned it in 50 years?’). If people were to consider the temporal location of a far future event only after reacting to its atemporal representation, shortsightedness might be ameliorated or reduced.”).

⁴⁴ See, e.g., Daniel T. Gilbert et al. As Bazerman and colleagues discuss, one of the main benefits of formal decision aids may be in clarifying the nature of internal conflict so that clearer choices can be made. See Max H. Bazerman et al., *Negotiating with Yourself and Losing: Making Decisions with Competing Internal Preferences*, 23 ACAD. MGMT. REV. 225, 236 (1998). The psychologist Jonathan Baron provides a collection of useful decision aids on his website: <http://www.psych.upenn.edu/~baron/daids.html>.

In addition, some very simple changes in the choice setting may enhance self-control. For instance, self-control seems to be greater earlier rather than later in the day, before the reserve of self-control has been depleted by numerous temptations and decisions.⁴⁵ Thus, asking employees to make important decisions, such as retirement savings decisions, in the morning may result in choices exhibiting greater self-control. We may also hypothesize that the libertarian yet paternalistic cafeteria director discussed by Sunstein and Thaler, in addition to arranging food selections to encourage healthful choices,⁴⁶ may add several mirrors to the cafeteria to induce a self-awareness in diners that leads to less impulsive dietary choices.⁴⁷

In sum, the conclusion that choice-framing paternalism is inevitable does not follow from Sunstein and Thaler's premises. Choice-framing paternalism would be inevitable only if normatively irrelevant features of the choice setting could not be ignored or overcome, but that is not the case. Therefore, for the committed libertarian, a better alternative to libertarian

Bazerman and colleagues also note, however, that the "want self" may not give way to demands of the "should self." Bazerman et al., *supra*, at 236. Accordingly, they also recommend committing oneself to "rational negotiation strategy" to reconcile internal conflicts. *See id.* at 236-37.

⁴⁵ *See* Baumeister, *supra* note 43, at 673 ("The self's resources are restored during sleep and then become progressively depleted during the day, especially insofar as the day makes demands for choices and self-control. Hence failures of self-control are rare in the morning and become progressively more likely as the day wears on.").

⁴⁶ *See* Sunstein & Thaler, *supra* note 1, at 1164-65, 1184.

⁴⁷ *See* Stacey M. Sentryz & Brad J. Bushman, *Mirror, Mirror on the Wall, Who's the Thinnest One of All? Effects of Self-Awareness on Consumption of Full-Fat, Reduced-Fat, and No-Fat Products*, 83 J. APPLIED PSYCHOL. 944,946-47, 948 (1998) (discussing results of studies showing that the presence of large mirrors in room or grocery store reduced consumption of unhealthy foods but not less healthy, high-fat foods compared to reduced- and no-fat, healthier foods); *see also* Baumeister, *supra* note 43, at 672 ("If monitoring was restored by focusing attention on the self's eating, the dieters kept their eating under control."); Patricia Pliner & Gerard Iuppa, *Effects of Increasing Awareness on Food Consumption in Obese and Normal Weight Subjects*, 3 ADDICTIVE BEHAV. 19 (1978) (reporting that both obese and normal weight persons ate less in the presence of an uncovered mirror than in the presence of a covered mirror). Interestingly, the positive effects of mirrors, with their ability to trigger self-awareness, may extend beyond the domain of diet and self-control. *See* Jerald Greenberg, *Overcoming Egocentric Bias in Perceived Fairness Through Self-Awareness*, 46 SOC. PSYCHOL. Q. 152 (1983) (using mirror-induced self-awareness to eliminate egocentric bias). To the extent that the cafeteria director rearranges the order of food selections is meant to encourage reflection and open-mindedness about dietary choices, then the cafeteria director simply rearranges rather than control other people's weight by controlling their food selections,

paternalism would be to help decision-makers resist or learn how to overcome irrational influences on their choices, in order to enhance the freedom and quality of choice and to overcome the manipulation of their preferences by central planners.⁴⁸

My brief discussion of some of the evidence supporting a debiasing approach is not meant to encourage a Panglossian attitude toward the tractability of all self-control and contextualization problems, but it is meant to suggest viable, more libertarian-friendly alternatives to libertarian paternalism for some of the choice settings that Sunstein and Thaler discuss.⁴⁹ In other settings, debiasing mechanisms might usefully be joined with the libertarian

⁴⁸ Of course, simply trying to debias or educate individuals to achieve greater decision-making competence has an element of paternalism in that it presumes that more competence is better for the individual, but such approaches importantly do not necessarily seek to manipulate choices in a particular direction or make choices for the individual. To the extent educational paternalism directed at improving freedom of choice is acceptable to the libertarian, then more programmatic educational measures might be quite helpful. For example, facility with statistical reasoning can be greatly improved with specialized training. See PETER SEDLMEIER, IMPROVING STATISTICAL REASONING: THEORETICAL MODELS AND PRACTICAL IMPLICATIONS 140-41 (1999) (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 training, and generally high long-term effectiveness from training); see also Peter Wright, *Marketplace Metacognition and Social Intelligence*, 28 J. CONSUMER RES. 677, 680 (2002) (discussing PREEMPT programs, or Preadult Education on Marketplace Persuasion Tactics). For a discussion of the benefits a more programmatic approach to teaching critical thinking skills within the educational system, see Jonathan Baron, *Why Teach Thinking? –An Essay*, 42 APPLIED PSYCHOL.: INT’L REV. 91 (1993).

Sunstein and Thaler note the paternalistic nature of education designed to shape preferences in a certain way, see Sunstein & Thaler, *supra* note 1, at 1183 (“And to the extent that those who design informational strategies are taking account of how people think and are attempting to steer people in desirable directions, their efforts will inevitably have a paternalistic dimension.”), but not all forms of education require that the educator indoctrinate a particular substantive goal. Some educational programs are “geared toward improving general decision-making skills (applicable to any content),” whereas other programs are “geared toward decision-making within a specific content area (e.g., drug education)” and may seek to steer people toward a particular outcome. JAMES P. BYRNES, THE NATURE AND DEVELOPMENT OF DECISION MAKING 175 (1998). The debiasing measures discussed in the text are meant to improve general competence rather than push individuals toward any particular choice. For instance, helping individuals to be active, open-minded thinkers or more self-aware choosers presumes no specific end-state, just as teaching individuals to better understand and apply the rules of logic and probability can help regardless of choice options. A conservative or liberal, libertarian or paternalist, can benefit from such skills regardless of the starting and ending points in an argument or plan of action.

⁴⁹ Although there is good evidence to support the view offered here to counter that offered by Sunstein and Thaler, one could also reframe my argument as a proposal for competitive empirical testing. My contention is that only a weak form of the preference endogeneity premise is empirically justifiable, as opposed to the strong form of the premise used by Sunstein and Thaler, and that this weaker form of choice-frame sensitivity permits the use of debiasing mechanisms that tread more lightly on libertarian values than the preference manipulation of libertarian paternalism. I hypothesize, based on existing evidence, that, at a minimum, a number of choice situations that Sunstein and Thaler label as problematic and subject to “inevitable paternalism” do not pose insuperable preference

paternalist approach to produce a truer form of libertarian paternalism (e.g., a libertarian should not object to use of a pre-commitment device like the “Save More Tomorrow” plan to combat willpower problems so long as the plan was chosen under conditions that favor the expression of true preferences).⁵⁰

Notwithstanding the evitability of choice-framing paternalism, it is inevitable that some default rules be set for all contractual arrangements, as Charny has pointed out,⁵¹ and indeed

determination problems for the individual who is given the tools to help debias his or her own decision-making process.

⁵⁰ Although I have largely set aside questions about the proper interpretation of results from behavioral economics seeming to show irrational behavior, *see supra* note 8, Hirschman’s point about the necessity of first-order preference (or taste) instability, if we believe that humans hold transcendent values (or second-order preferences or metapreferences) and have the ability to reflect on what ends are better and worse, should be kept in mind when attempting to decide how serious a problem preference instability really is:

men and women have the ability to step back from their ‘revealed’ wants, volitions, and preferences to ask themselves whether they really want these wants and prefer these preferences, and consequently to form metapreferences that may differ from their preferences. . . . [C]ertainty about the existence of metapreferences can only be gained through *changes* in preferences, that is, through changes in actual actual behavior, at some instability in transient preferences simply implies the existence of values or metapreferences.

Albert O. Hirschman, *Against Parsimony: Three Easy Ways of Complicating Some Categories of Economic Discourse*, 74 AMER. ECON. REV. 89, 89 (1984).

⁵¹ Charny discusses epistemological and practical reasons why default rules are inevitable:

The law must supply a set of background conditions to interpretation and enforcement of contracts—commonly referred to as “default rules.” Without default rules, no contract could have legal effect. This is the case for two reasons. Most fundamentally, no text can completely specify its own means of interpretation. A contractual statement that purported to be such a complete specification would itself have to be interpreted by some set of rules of interpretation. If the text purported to supply those rules, then those rules would have to be interpreted, and so on, ad infinitum. This, the default rules must, at a minimum, contain a set of rules about how the language of contract is to be interpreted.

Second, aside from this basic epistemological constraint, an important set of practical constraints limits the completeness of contracts. In almost all transactions, it would be extremely costly to draft a contract that purported explicitly to address the obligations of the parties for all conceivable future contingencies. As a practical matter, then, most contracts are quite incomplete. The law supplies these missing terms. For example, the doctrines of mistake and impracticability add terms to address low-probability contingencies; so do doctrines that impose general duties, such as the duty of good faith. Correlatively, once the parties know that the law will supply the term, they take that into account when calculating the benefits of drafting an express term. Parties will not incur the costs of specifying the term if they suspect that courts will supply the appropriate term in any event.

more generally for all choice settings, as Sunstein and Thaler point out.⁵² But not all default rules are equally paternalistic or paternalistic in the same way, and, as discussed in the next Part, it does not follow from libertarian principles that the central planner should choose the default rule that enhances the welfare of affected individuals. Rather, the least objectionable form of default-rule paternalism from a libertarian perspective requires that the default rule be set to enhance liberty rather than welfare.

II. Using Paternalism to Maximize Liberty, Not Welfare

Libertarianism involves a commitment to individual liberty regardless of whether individuals use their liberty wisely, and requires that individuals bear the consequences of their mistakes as the price paid for the freedom to make their own mistakes. Accordingly, a libertarian system of rights “include[s] no rights to be given positive assistance, aid, or nurturance by others.”⁵³ “One’s right to act as one chooses in concert with others so long as one does not wrongfully harm others who do not consent to bear the costs includes the right to suffer any purported injury to which one voluntarily consents.”⁵⁴

Therefore, the proper evaluative view of choice behavior from the libertarian perspective is not an objective consequentialist view, but rather a view that examines only the quality of individual consent. This means that the committed libertarian will find Sunstein and Thaler’s

David Charny, *Hypothetical Bargains: The Normative Structure of Contract Interpretation*, 89 MICH. L. REV. 1815, 1819-20 (1991) (footnotes omitted).

⁵² See, e.g., Sunstein & Thaler, *supra* note 1, at 1165 (“governments . . . have to provide starting points of one or another kind; this is not avoidable”). Even the requirement that choosers make an “active” choice among a menu of options, rather than setting one option as the automatic but mutable choice, is a form of default rule. See *id.* at 1173 (“[R]equired active choosing is simply another option among many that the employer can elect. In fact the very requirement that employees make a choice has a strong paternalistic element.”).

⁵³ Richard J. Arneson, *Perfectionism and Politics*, 111 ETHICS 37, 41 (2000).

⁵⁴ *Id.* at 40.

evidence on irrational choice behavior relevant to his understanding of the constraints on voluntary consent, but he will draw a very different prescriptive lesson from this evidence than that drawn by Sunstein and Thaler. Whereas Sunstein and Thaler treat this evidence as granting libertarian license to impose some conception of objectively good preferences on irrational persons,⁵⁵ the committed libertarian will treat this evidence as justifying only the imposition of paternalistic policies aimed at maximizing the liberty and autonomy of these individuals, without regard to objective outcomes.

To the committed libertarian, the evidence of irrational choice behavior could justify only two types of paternalistic interference with individual liberty: (1) interference to improve decision-making competence (or “general competence” as Child labels it⁵⁶), which would take the form of the debiasing measures discussed in Part I, and, if these debiasing measures fail or are not available, (2) interference to prevent liberty-restricting irrational choices.⁵⁷ Sunstein and Thaler assume away the very real possibility of inoculating individuals against irrational influences in choice settings, as already discussed in Part I. The second type of paternalism—protection against liberty-restricting choices—raises complex issues, some of which are addressed below, but for present purposes the most important point is that a truly *libertarian*

⁵⁵ See *infra* notes 56-64 and accompanying text.

⁵⁶ See *supra* note 15.

⁵⁷ In discussing the interplay of perfectionism and libertarianism, Arneson offers a similar view:

Restriction of choice now even to the extent of forcing a single choice upon the individual can pave the way to autonomous choice of that good or others later. This long-run fostering of autonomous choice of the good is easier to bring about if the coercion does not force upon the agent a single option thought to be good but instead prohibits some tempting bad options while leaving many other options open.

See *id.* at 44. But cf. David Gordon, *Miller on Market Neutrality, Co-Operatives and Libertarianism*, 13 BRIT. J. POL. SCI. 125, 127 n.4 (1983) (“In the libertarian view, coercing someone in order to secure the realization of that person’s long-run preferences counts as an unacceptable interference with personal liberty.”).

paternalism of this second type will look very different from the libertarian paternalism offered by Sunstein and Thaler.

Sunstein and Thaler's prescriptions focus on institutional and contractual designs aimed at improving the welfare of irrational persons while leaving rational persons the right to choose as they see fit.⁵⁸ To this end, Sunstein and Thaler counsel the central planner to consider the welfare effects of different default rules, using either a cost-benefit analysis or one of three suggested rules of thumb,⁵⁹ and to then design the choice setting to "influence the choices of affected parties in a way that will make choosers better off."⁶⁰ Although Sunstein and Thaler embrace welfare improvement as the overarching principle guiding the central planner's designs, Sunstein and Thaler do not endorse any particular measure of welfare to guide the planner,⁶¹ nor do they direct the planner to try to "track people's anticipated choices."⁶² Thus, Sunstein and Thaler do not choose one particular conception of the good that the planner should try to achieve, but most of their illustrative examples involve making people healthier or wealthier,⁶³ suggesting that some form of material consequentialism may at times be a reasonable goal for the central planner to pursue.⁶⁴ For instance, Sunstein and Thaler direct the central planner designing a

⁵⁸ As already discussed, the libertarian opt-out option Sunstein and Thaler emphasize is meaningful only to rational persons (i.e., persons who can resist the central planner's manipulative strategies). *See supra* notes 35-37 and accompanying text.

⁵⁹ *See* Sunstein & Thaler, *supra* note 1, at 1190-95.

⁶⁰ *Id.* (footnote omitted).

⁶¹ *See* Sunstein & Thaler, *supra* note 1, at 1163 n.17 ("We are not attempting to say anything controversial about welfare, or to take sides in reasonable disputes about how to understand that term.").

⁶² *Id.* at 1162.

⁶³ *See, e.g., id.* at 1167-68, 1184 (equating welfare with improved health); *id.* at 1169-70 (equating welfare with greater wealth).

⁶⁴ Once one recognizes a substantial risk of error in individual beliefs associated with due to judgmental biases and errors, or an unreliability in revealed preferences (i.e., choice behavior) as a measure of subjective utility due to choice-framing effects, the case for moving from a subjective to a more objective measure of welfare becomes

retirement plan to balance the wealth effects of different default rules.⁶⁵ Most significantly for this discussion, none of Sunstein and Thaler’s examples or suggestions for the central planner emphasizes the maximization of future liberty as the goal.⁶⁶

A libertarian approach to choice-framing paternalism would direct the central planner to frame choices in ways that push irrational persons in directions that maximize their liberty or help them retain the greatest degree of future freedom to contract. This is not an order to give individuals as many options as possible or force them to make active choices—two design approaches that Sunstein and Thaler caution against.⁶⁷ Instead, this direction tells the central

clearer, as Hausman and McPherson discuss. See Daniel Hausman & Michael McPherson, *Preference, Belief, and Welfare*, 84 AMER. ECON. REV. 396, 396 (1994) (“[T]he dependence of preferences on unreliable beliefs is a pervasive feature of social life which makes a preference-based standard of welfare generally problematic.”); *id.* at 398-99 (“[A] more objective standpoint would be less dependent on idiosyncratic and unreliable personal beliefs about the likelihood of various outcomes and better suited to reasoned deliberation.” (footnote omitted)); see also Matthew D. Adler & Eric A. Posner, *Implementing Cost-Benefit Analysis When Preferences Are Distorted*, 29 J. LEGAL STUD. 1105, 1113-14 (2000) (arguing for a “restricted preference-based account of well-being,” which corrects for distortions to actual preferences, for purposes of cost-benefit analysis). Hausman and McPherson also recognize that choosing the proper objective measure of welfare raises many difficult issues, including that using an objective measure of welfare “may raise worries about liberty, about government overriding persons’ own judgments and desires.” Hausman & McPherson, *supra*, at 399.

⁶⁵ See *id.* at 1190 (“Under automatic enrollment, some employees, who otherwise would not join the plan, will now do so. Presumably, some are made better off (especially if there is an employer match), but some may be made worse off (for example, those who are highly liquidity-constrained and do not exercise their right to opt out). A cost-benefit analysis would attempt to evaluate these gains and losses.”). And they tell us that, “[i]f the issue were only enrollment, we think it highly likely that the gains would exceed the losses.” *Id.*; see also *id.* at 1166 (“programs should be designed using a type of welfare analysis, one in which a serious attempt is made to measure the costs and benefits of outcomes (rather than relying on estimates of willingness to pay)”).

⁶⁶ That is, the libertarian aspect of Sunstein and Thaler’s libertarian paternalism is confined to its opt-out component: “So long as it is costless or nearly costless to depart from the default plan, minimal paternalism is maximally libertarian.” *Id.* at 1188; see also *id.* at 1175 (“So long as people can contract around the default rule, it is fair to say that the legal system is protecting freedom of choice, and in that sense complying with libertarian goals.”).

Sunstein and Thaler do provide that occasionally liberty should be considered a part of welfare for planning purposes. See *id.* at 1167 n.22 (“Our principal concern here is with welfare and consequences, though as we suggest below, freedom of choice is sometimes an ingredient in welfare.”); *id.* at 1198-99 (“Freedom of choice is itself an ingredient in welfare. In some situations people derive welfare from the very act of choosing. . . . But much of the time, especially in technical areas, people do not particularly enjoy the process of choice, and a large number of options becomes a burden.”).

⁶⁷ See Sunstein & Thaler, *supra* note 1, at 1196 (“Generally, the more complex the decision, the less attractive it will be to force people to choose for themselves, as opposed to having the option of choosing not to choose, and receiving a default option that has been selected with some care.”); *id.* (“Libertarian paternalists want to promote

planner to set the default to that option which, if chosen mindlessly, will be least restrictive of individual liberty, while leaving mindful individuals to opt out of the default option and enter into greater entanglements if they so choose.

The at-will employment rule is a good example of a libertarian default rule. On the one hand, it maximizes the liberty of irrational individuals by preventing them from entering mindlessly into binding employment arrangements, but on the other hand, since it can be overridden, rational individuals may contract around the at-will default to enter into more permanent employment arrangements. Through a libertarian default rule, the state protects irrational persons from involuntary encumbrances but allows rational persons to enter into whatever additional encumbrances they may freely choose. Most importantly, the state does not impose any substantive notion of the good life but helps individuals retain freedom to determine their own preferred ends. The fact that some irrational individuals might achieve better objective outcomes under a for-cause default rule is of no consequence to the libertarian.⁶⁸

Just as Sunstein and Thaler's welfarist version of libertarian paternalism requires many difficult judgments about material welfare by the central planner,⁶⁹ the libertarian version of libertarian paternalism discussed here would also require many difficult judgments about liberty. One of the most difficult issues would be how to deal with conflicts between immediate and

freedom of choice, but they need not seek to provide bad options, and among the set of reasonable ones, they need not argue that more is necessarily better. Indeed, that argument is quite implausible in many contexts.”). *But see* Russell Korobkin, *The Endowment Effect and Legal Analysis*, 97 NW. U.L. REV. 1227, 1273-74 (2003) (“Although default terms cannot truly be neutral, the government can avoid shaping preferences with default rules if it wishes to do so. One approach is to require contracting parties to select from among a menu of terms, rather than anointing one the default term. . . . A similar approach is to use a ‘non-enforcement’ default, in which courts will not enforce the contract at all unless the parties provide a term to deal with a particular contingency.” (footnote omitted)).

⁶⁸ For a discussion of the behavioral law and economics argument for switching the at-will default to a for-cause default, see generally Cass R. Sunstein, *Switching the Default Rule*, 77 N.Y.U.L. REV. 106 (2002).

⁶⁹ *See, e.g.*, Sunstein & Thaler, *supra* note 1, at 1190 (noting the difficulty of performing the necessary cost-benefit analyses and suggesting the use of experiments to test different approaches).

future liberty. For instance, on one view, the Save More Tomorrow plan interferes with individual liberty because it commits the irrational individual to a savings plan that interferes with the present self's current resources and hence freedom to contract. On a counter view, however, this plan advances the liberty of a future potential self by conserving resources for future use. Also, because it may be necessary in certain domains to possess a minimal level of resources and capabilities before one can exercise one's liberties effectively, the libertarian central planner, like Sunstein and Thaler's paternalistic planner, must confront the problem of when redistribution is justified, if ever.⁷⁰

At times there will be a confluence of the paternalist and libertarian approaches to institutional design. For instance, the libertarian approach to designing a buffet for irrational diners would place the most tempting foods at the end of the line so that these diners first observe other options before reaching the choices that appeal most to those with low self-control.⁷¹ While this design may look much like Sunstein and Thaler's design when the tempting foods are also the less healthy foods (as will typically be the case),⁷² the motivations behind the interventions differ: rather than imposing a paternalistic view of the proper diet on irrational diners, the libertarian designer aims to keep more choices open for the irrational diner.

Lest one see this libertarian approach to institutional design as just another manifestation of what Sunstein and Thaler label the dogmatic libertarian's inept neglect of the welfare effects

⁷⁰ See *infra* Pt. III. Given the general prohibition on redistribution under the libertarian view, as discussed in the next Part, it may be defensible to adopt the rule of thumb that the libertarian central planner should intervene only when one option clearly dominates another with respect to its effect on personal liberty.

⁷¹ This second-best libertarian approach assumes that a less intrusive debiasing measure, such as placing mirrors in the libertarian paternalist cafeteria, was not successful or was infeasible. See *supra* note 47 & accompanying text.

⁷² See Sunstein & Thaler, *supra* note 1, at 1195-99

of different legal regimes, two points bear mentioning.⁷³ First, keep in mind that the first approach of the libertarian central planner would be to debias individuals so that they can make their own rational decisions about which choices best promote their own welfare. The libertarian default rule serves only as a failsafe in the event that individuals fail to be able to act rationally. Second, one should not conflate an irrational choice with an objectively bad choice nor assume that the paternalist can make choices for the irrational individual that will lead to net gains in happiness. This second point requires some explication.

Recall that Sunstein and Thaler begin their argument with an attack on the libertarian assumption that individuals are good judges of their own welfare.⁷⁴ The basis for this attack is empirical research from psychology and behavioral economics testing the descriptive accuracy of rational choice theory's axioms, such as the dominance (if option A dominates option B in at least one set and is no worse than option B in all other sets, then option A should always be chosen over option B), transitivity (if option A dominates option B, and option B dominates option C, then option A should dominate option C), and invariance (alterations in the manner of presentation of options that do not affect the outcomes associated with options should be irrelevant to choices) axioms.⁷⁵ For expressed preferences to comply with the axioms of rational choice theory, choices must be stable and consistent with one another (or the desires and beliefs

⁷³ See *supra* note 7 and accompanying text.

⁷⁴ See Sunstein & Thaler, *supra* note 1, at 1163 (“We believe that [dogmatic anti-paternalism] is based on a combination of a false assumption and two misconceptions. The false assumption is that almost all people, almost all of the time, make choices that are in the best interest or at the very least are better, by their own lights, than the choices that would be made by third parties. This claim is either tautological, and therefore uninteresting, or testable. We claim that it is testable and false, indeed obviously false.”); see also *supra* note 15 and accompanying text.

⁷⁵ See Amos Tversky & Daniel Kahneman, *Rational Choice and the Framing of Decisions*, 59 J. BUS. S251, S252-54 (1986) (describing the three axioms mentioned in the text, along with the cancellation axiom).

that give rise to expressed preferences must be stable and logically coherent⁷⁶) (e.g., the dominance axiom requires a stability in beliefs and desires about the greater attractiveness of option A over option B across a wide range of situations). As Sunstein and Thaler emphasize, however, evidence from behavioral decision research shows that which desires and beliefs will control choice behavior (or the expression preferences) depends crucially on the particular context in which choices are elicited (i.e., rather than consulting a prepared menu of rank-ordered preferences when presented with choice, people will often construct a preference ordering on the spot, which can lead to a level of inconsistency across situations).⁷⁷ Accordingly, changing the way that choice options are presented to choosers, without changing the material payoff of the options, can affect preferences, in violation of the invariance axiom.⁷⁸ Or adding a new option to a choice set can cause individuals to reverse their preferences between options A and B, in violation of the invariance and dominance axioms.⁷⁹ In short, behavioral

⁷⁶ I am taking some psychological license with rational choice theory in making the connection between choice and belief/desire. For instance, Savage, one of the founders of rational choice theory, understood preferences in terms of choices rather than underlying beliefs or desires. See Robert Sugden, *Rational Choice: A Survey of Contributions from Economics and Philosophy*, 101 *ECON. J.* 751, 758 (1991) (“Savage insists that preference must be interpreted in terms of choice.”); *id.* at 759 (“Savage defines preference in terms of choice, not desire.”); *id.* at 760 (“Notice that for Savage, rationality is understood in terms of the consistency of choices *with one another*, and not in terms of their consistency with any given system of desires and beliefs.”).

⁷⁷ See, e.g., Sunstein & Thaler, *supra* note 1, at 1201 (“Our central empirical claim has been that in many domains, people’s preferences are labile and ill-formed, and hence starting points and default rules are likely to be quite sticky.”); see also Paul Slovic, *The Construction of Preference*, 50 *AMER. PSYCHOL.* 364, 365 (1995) (“preferences appear to be remarkably labile, sensitive to the way a choice problem is described or ‘framed’ and to the mode of response used to express the preference” (citations omitted)); John W. Payne, James R. Bettman & Eric J. Johnson, *Behavioral Decision Research: A Constructive Processing Perspective*, 43 *ANN. REV. PSYCHOL.* 87, 89 (1992) (“preferences for and beliefs about objects or events of any complexity are often constructed—not merely revealed—in the generation of a response to a judgment or choice task” (citation omitted)).

⁷⁸ See, e.g., Tversky & Kahneman, *supra* note 75, at S256-57 (“[V]ariations in the framing of decision problems produce systematic violations of invariance and dominance that cannot be defended on normative grounds. . . . Normative models of choice, which assume invariance, therefore cannot provide an adequate descriptive account of choice behavior. . . . Failures of invariance are explained by framing effects that control the representation of options, in conjunction with nonlinearities of value and belief.”).

⁷⁹ See, e.g., Chris Guthrie, *Panacea or Pandora’s Box?: The Cost of Options in Negotiation*, 88 *IOWA L. REV.* 601, 614 (2003) (“Rational models of choice generally assume that the presence of an additional option will not alter a person’s relative valuation of options already under consideration unless the new option conveys relevant

decision research reveals that the axioms of rational choice theory often fail to describe actual behavior, or, stated alternatively, human choice behavior often deviates from the dictates of rational choice theory.⁸⁰

What conclusion about individual welfare can fairly be drawn from this evidence of irrational choice behavior? Only that the irrational individual has failed to do what he or she most prefers, for rational choice theory employs an ordinal definition of utility that does not permit the kinds of external normative evaluations or interpersonal welfare comparisons that an objective measure of welfare, such as monetary wealth or healthiness, would permit⁸¹:

When economists say that individuals maximize utility, they are only saying that people do not rank any feasible option above the option they choose. . . . Good economists sometimes speak misleadingly of individuals as *aiming to* maximize utility or *seeking* more utility, but they do not or should not mean that utility is an object of choice, some ultimately good thing that people want in addition to healthy children or better television. The theory of rational preference or choice specifies no distinctive aims that all people must embrace. Utility is just an index of preferences. An individual who is a utility maximizer just does what he or she most prefers. To say that individuals are utility maximizers says nothing about the nature of their preferences. *All it does is to connect preference and choice (or action in a particularly simple way). Rational individuals rank available alternatives and choose what they most prefer.*⁸²

information about those other options. . . . Psychologists have discovered, however, that people's assessments of initially considered options are often systematically influenced by the emergence of an additional, irrelevant option." (footnotes omitted)).

⁸⁰ See Sunstein & Thaler, *supra* note 1, at 1168 ("On a more scientific level, research by psychologists and economists over the past three decades has raised questions about the rationality of many judgments and decisions that individuals make."); see also Craig R.M. McKenzie, *Rational Models as Theories—Not Standards—of Behavior*, 7 TRENDS COGNITIVE SCI. 403, 403 (2003) ("Cognitive scientists who study reasoning and decision-making often compare people's behavior with a rational model of a given task, and the typical published result is that behavior deviates from the model.").

⁸¹ For discussion of the "ordinalist revolution," which "refers to the rejection of cardinal notions of utility and to the general acceptance of the position that utility was not comparable across individuals," see Robert Cooter & Peter Rapport, *Were the Ordinalists Wrong About Welfare Economics?*, 22 J. ECON. LIT. 507, 507 (1984); see also Yew-Kwang Ng, *A Case for Happiness, Cardinalism, and Interpersonal Comparability*, 107 ECON. J. 1848, 1848 (1997) ("After the indifference-curve or ordinalism revolution in the 1930s, modern economists are very adverse to the more subjective concepts and very hostile to cardinal utility and interpersonal comparisons of utility.").

⁸² DANIEL M. HAUSMAN, *THE INEXACT AND SEPARATE SCIENCE OF ECONOMICS* 18 (1992); see also Henry J. Aaron, *Public Policy, Values, and Consciousness*, 8 J. ECON. PERSPECTIVES 3, 14 (1994) ("I believe that it is important in analyzing the effects of public (or private) policies intended to affect behavior to start from a recognition that actions

Rationality within behavioral economic and psychological tests of rational choice theory requires only a consistency among expressed preferences, with consistency defined by the axioms of rational choice.⁸³ Rationality does not require the promotion of welfare in any objective sense (nor do behavioral tests of rational choice theory inspect welfare promotion in any objective sense⁸⁴). Indeed, one individual may rationally make choices that others would deem objectively bad choices (i.e., this person acts in ways that maximize her utility, but others deem her chosen ends to be objectionable, as with the rational sadist or masochist), while another individual irrationally makes choices that others would deem objectively good choices (i.e., the irrational sadist or masochist who fails to choose in ways that will maximize her sadistic or masochistic preferences).⁸⁵

are taken because they serve the interests of the actor as the actor sees them.”); William H. Riker, *The Political Psychology of Rational Choice Theory*, 16 POL. PSYCHOL. 23, 24-25 (1995) (“In the rational choice model . . . there is no particular goal. Rather, there is a set of possible goals that must be ordered. . . . It is . . . important to point out that whether the goals in the set are good or evil or neither is irrelevant, as long as the participants order them.”).

⁸³ See MARK BLAUG, *THE METHODOLOGY OF ECONOMICS, OR HOW ECONOMISTS EXPLAIN* 229 (2d ed. 1992) (“For the economist . . . , rationality means choosing in accordance with a preference ordering that is complete and transitive, subject to perfect and costlessly acquired information; where there is uncertainty about future outcomes, rationality means maximizing expected utility, that is, the utility of an outcome multiplied by the probability of its occurrence.”); Daniel Kahneman, *New Challenges to the Rationality Assumption*, in CHOICES, VALUES, AND FRAMES, note 25, at 759 (“[T]echnical discussions of rationality generally adopt a *logical* conception, in which an individual’s beliefs and preferences are said to be rational if they obey a set of formal rules such as complementarity of probabilities, the sure thing principle or independence of irrelevant alternatives. In the laissez-faire spirit of modern economics and decision theory, the content of beliefs and of preferences is not a criterion of rationality—only internal coherence matters.” (citation omitted)); see also Amartya Sen, *Internal Consistency of Choice*, 61 ECONOMETRICA 495, 495 n.3 (1993) (“The fulfillment of internal consistency is sometimes seen as the central feature of ‘rational choice.’”).

⁸⁴ There is, however, a growing interest in the psychology of utility, happiness, and well-being. For an introduction to empirical studies in “hedonic psychology,” see WELL-BEING: THE FOUNDATIONS OF HEDONIC PSYCHOLOGY (Daniel Kahneman et al. eds., 1999).

⁸⁵ See Daphna Lewinsohn-Zamir, *The Objectivity of Well-Being and the Objectives of Property Law*, 78 N.Y.U.L. REV. 1669, 1678 (2003) (“A major objection to the actual preference . . . theory is that people may often desire what is bad for them.”). Cf. JONATHAN BARON, *THINKING AND DECIDING* 55 (3d ed. 2000) (“Rationality is . . . not the same as accuracy, and irrationality is not the same as error. We can use good methods and reach erroneous conclusions, or we can use poor methods and be lucky, getting a correct answer.”).

Therefore, when using an ordinal definition of utility or individual welfare, it is *impossible by definition* for a third party to make judgments about another individual's utility, because the ranking of preferences is purely subjective, with no objective goal implied or possible. The most that can be said about individual welfare from evidence of an individual's irrational choice behavior is that, based on revealed preferences, the individual failed to maximize his or her own subjective expected utility. This evidence cannot mean that a third party could do better than, or even as well as, the individual with respect to the maximization of ordinal utility. Only the individual can order his or her own preferences for purposes of maximizing subjective utility.

In addition, any system that makes welfare synonymous with ordinal utility, as rational choice theory and behavioral tests of rational choice theory do, will not permit interpersonal comparisons of welfare because welfare is not measured on a cardinal, or interval, scale (i.e., a scale that measures intensity of preferences and not their order) that would permit such

interpersonal comparisons.⁸⁶ To make interpersonal welfare comparisons, the welfare paternalist needs some cardinal measure of utility or some objective measure of welfare.⁸⁷

And here is the rub for Sunstein and Thaler’s welfare-based libertarian paternalism, for it is unlikely that the central planner can choose any objective measure of welfare that will not be objectionable to some set of people in any given context, and the adoption of a form of utilitarianism is likely to trample on libertarian principles.⁸⁸ In practice it will often be difficult

⁸⁶ Except in a way that is not helpful to the libertarian paternalist, that is. A small number of principles, including Rawls’ maximin principle, which gives priority to the worst off in a society, arguably permit interpersonal comparisons of ordinal utility to guide the formation of social policy. See Amartya Sen, *Rationality and Social Choice*, 85 AMER. ECON. REV. 1, 8 (1995) (“Even without any cardinality, ordinal interpersonal comparisons permit the use of such rules of social judgment as maximin, or lexicographic maximin.” (footnote omitted)). It is not necessary to dwell on this possibility, however, for the libertarian central planner, who is tasked with promoting the welfare of all affected parties and not simply the “worst off” (and simply defining “worst off” begins to raise questions about making interpersonal ordinal utility comparisons), cannot choose such a principle of welfare. A maximin rule, in other words would tread greatly on the welfare of “better off” individuals and thus greatly trouble the committed libertarian. (Indeed, Nozick’s theory of the libertarian minimal state is often set in opposition to Rawls’ theory of justice. For Nozick’s argument against Rawls, see ROBERT NOZICK, ANARCHY, STATE, AND UTOPIA 183-231 (1974).)

To be clear, the point here is not that interpersonal comparisons of utility should never be attempted—only that in most instances they simply cannot be conducted using ordinal utility. Cf. Adler & Posner, *supra* note 64, at 1111 n.8 (“Skeptics of interpersonal welfare comparisons sometimes point out that it is not practicable for government to make such comparisons. We agree. But this is not the same as saying that interpersonal comparisons are impossible. To think otherwise is to conflate decision procedures and normative criteria.”).

⁸⁷ See Martin Shubik, *A Curmudgeon’s Guide to Microeconomics*, 8 J. ECON. LIT. 405, 411 (1970) (“It is worth noting that if one is committed to the proposition that consumer preferences cannot be determined beyond an ordinal measure, then nothing whatsoever can ever be said about welfare schemes involving any form of fair division or equitable settlement.”). Or to put it another way, to make interpersonal comparisons a “thicker” conception of utility must be employed. See Russell B. Korobkin & Thomas S. Ulen, *Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics*, 88 CAL. L. REV. 1051, 1060-66 (2000) (discussing “thin” and “thick” conceptions of utility, with wealth maximization being the most common thick form within law and economics).

⁸⁸ For an excellent brief discussion of points of conflict between utilitarianism and libertarianism, see Larry Alexander & Maimon Schwarzschild, *The Uncertain Relationship Between Libertarianism and Utilitarianism*, 19 QUINNIPIAC L.R. 657 (2000). As Sen famously pointed out with his “Prude” and “Lewd” example involving *Lady Chatterley’s Lover*, “free choice in personal matters can easily conflict with utilitarian principles.” Samuel Brittan, *Two Cheers for Utilitarianism*, 35 OXFORD ECON. PAPERS 331, 342 (1983) (citing AMARTYA SEN, CHOICE, WELFARE AND MEASUREMENT (1982) (introductory chapter)). The Sen problem arises from “interdependence effects” on utility: “[i]f negative interdependence effects are taken into account in public policy, people will be penalized for carrying out private personal acts, which affect others only through thinking making it so.” *Id.* Notwithstanding potential conflicts, an affinity still exists between utilitarianism and a pragmatic variety of libertarianism, which sees the market and private contracts as the best means to economic efficiency. See, e.g., EDWARD FESER, ON NOZICK 13-14 (2004) (“Some libertarians would appeal to *utilitarianism* . . . and argue that it is

to find an uncontroversial measure of objective individual welfare even among non-libertarians. Many may agree in the abstract that better health is preferable to worse health, but when the choice is framed as enjoying life-shortening but intensely pleasurable vices during one's college days versus abstaining during college to gain a couple of extra boring years at an advanced age, then better health may not look quite as good. Thus, picking health as the welfare measure entails an interpersonal comparison of welfare that forces the preferences of future potential selves onto present selves. And greater wealth, the other obvious objective measure of welfare, has its own obvious and less obvious problems.⁸⁹

The welfare paternalist must navigate these difficult waters that will ultimately lead to the imposition of some conception of welfare on irrational people that some subset is sure to find objectionable. Thus, when Sunstein and Thaler avoid the issue of what exactly welfare enhancement should look like under their version of libertarian paternalism, they avoid the hardest but most important question raised by their welfare-focused paternalism.⁹⁰ The liberty-focused paternalist, on the other hand, avoids these difficult issues, because he has chosen individual liberty over welfare, happy to trust free individuals to make their own welfare decisions and let them live with the consequences.⁹¹

precisely a libertarian society which produces the best consequences. (Utilitarian moral arguments for libertarianism thus, for obvious reasons, very quickly reduce to pragmatic arguments).”)

⁸⁹ Among the obvious problems are concerns about the marginal utility and inherently instrumental nature of money. Among the less obvious problems is the fact that increasing wealth and increasing happiness are not as closely related as might be assumed. See Ng, *supra* note 81, at 1849 (“for economically advanced countries (the number of which is increasing) there is evidence suggesting that, for the whole society and in the long run (in real purchasing power terms) money does not buy happiness, or at least not much” (citations omitted)).

⁹⁰ Cf. Richard H. Fallon, Jr., *Should We All Be Welfare Economists?*, 101 MICH. L. REV. 979, 983 (2003) (noting, with respect to the credence of Kaplow and Shavell’s argument that only welfare calculations should drive normative legal decisions with no place given to fairness considerations, “that nearly everything depends on how ‘well-being’ and being ‘better-off’ are defined”).

⁹¹ Cf. Amartya Sen, *Utilitarianism and Welfarism*, 76 J. PHIL. 463, 484 (1979) (“I have to report that it was a little difficult to get an outcome libertarian for the dialogue, since libertarians tend to be anti-consequentialist and ferociously ‘deontic’ . . .”).

III. The Redistributive Consequences of Libertarian Paternalism

We may return to the example of the potential conflict between present and future selves with respect to current spending versus retirement savings to see another problem with Sunstein and Thaler's attempted reconciliation of libertarian principles with paternalistic goals. In order to enhance the welfare of irrational individuals in their golden years, Sunstein and Thaler propose using a libertarian paternalist retirement plan effectively to redistribute wealth from the present self to the future self. Such redistribution violates the libertarian prohibition on state-based redistribution of resources for any purpose except rectification of involuntary transfers,⁹² but Sunstein and Thaler might respond that the holdings of the present irrational self are not legitimate and entitled to protection because they arise from the choices of an incompetent individual.⁹³ That is, no "true" preferences are disturbed by the paternalistic policy and hence no libertarian objection should be raised.

However, there is a way that libertarian paternalism may lead to redistribution that interferes with "true" preferences which presents itself once we inquire into the prevalence of the irrational choice tendencies that Sunstein and Thaler hope to redirect through their libertarian

⁹² See, e.g., Child, *supra* note 15, at 722 ("As usually set out [within libertarian writings], adults having normal capacities must be free to engage in any market transaction absent force or fraud. The appropriate role of the state is merely to see that these two disruptions of free market transactions do not occur." (footnote omitted)); Geoffrey Sampson, *Liberalism and Nozick's 'Minimal State,'* 87 *Mind* 93, 93 (1978) ("One of the key principles of liberalism, as Nozick realizes, is that the State should not enforce 'redistribution': as far as possible it should avoid interfering with the pattern of rewards which emerges from the free play of market forces, either by progressive taxation or by providing a so-called 'social wage,' i.e., benefits paid from tax revenue rather than by charges to the consumers."); James P. Sterba, *From Liberty to Welfare,* 105 *ETHICS* 64, 67 (1994) ("libertarians are opposed to any coercively supported welfare program").

⁹³ Sunstein and Thaler might also respond by arguing that it makes no sense to prefer present over future interests and that we should focus on well-being of the person over the whole lifetime. For a discussion of the different views of the single self versus multiples selves and the relevance of these views to the rational choice literature, see Shane Frederick, *Time Preference and Personal Identity,* in *TIME AND DECISION* 89 (George Lowenstein et al. eds., 2003). If the individual is treated as a continuous self, then the self-paternalism problem becomes much less acute. See, e.g., Adam J. Hirsch, *Spendthrift Trusts and Public Policy: Economic and Cognitive Perspectives,* 73 *WASH. U.L.Q.* 1, 89 (1995) ("Both from a moral and utilitarian standpoint, self-paternalism appears even easier to justify than parental paternalism. Self-paternalism entails no interference with individual autonomy." (footnote omitted)); *id.* at 89 n.333 (discussing objections to self-paternalism).

paternalist policies. Sunstein and Thaler do not squarely address the question of prevalence with respect to the irrational choice tendencies that they discuss, but they do at various points speak in terms that suggest general tendencies toward irrationality within the population at large.⁹⁴ At other points, however, Sunstein and Thaler's language admits of some variation in irrational tendencies,⁹⁵ as it must, for individuals differ reliably in their propensities to exhibit various irrational behaviors.⁹⁶ When these individual differences in rationality are brought to the fore,

⁹⁴ See, e.g., Sunstein & Thaler, *supra* note 1, at 1168 ("People fail to make forecasts that are consistent with Bayes's rule, use heuristics that can lead them to make systematic blunders, exhibit preference reversals . . . , suffer from problems of self-control, and make different choices depending on the framing of the problem." (footnotes omitted)). Sunstein and Thaler do acknowledge that efforts at self-control and learning may help overcome irrational tendencies, but they ultimately discount the importance of these corrective factors:

It is true that people sometimes respond to their own bounded rationality by, for example, hiring agents or delegating decisions to others. It is also true that learning frequently enables people to overcome their own limitations. But many of the most important decisions (for example, buy a home or choosing a spouse) are made infrequently and typically without the aid of impartial experts. The possibilities of delegation and learning are insufficient to ensure that people's choices always promote their welfare or that they always choose better than third parties would.

Id. at 1170 (footnote omitted).

⁹⁵ Sunstein and Thaler at times qualify their unspecific "people exhibit this or that irrational tendency" by saying "many" or "most" people exhibit this or that tendency. See, e.g., *id.* at 1184 ("many people face problems of self-control"); see also *id.* at 1181 (noting that women and African-American have been found to be most affected by a change in a default rule for a retirement savings plan and speculating that "members of such groups tend to be less confident in their judgments in this domain and may have less experience in assessing different savings plans"). These general statements, whether qualified or not, in addition to obscuring the reliable individual differences in rationality, imply a level of prevalence of irrationality within the findings from judgment and decision-making research that simply does not often exist. See Mitchell, *supra* note 22, at 1805-07 (discussing how a minority of experimental subjects exhibiting an irrational behavior results in a statistically significant finding of some "irrational tendency"); see also 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. (forthcoming 2004) (manuscript at 18; footnote omitted) (manuscript on file with author). ("Indeed, some biases reach significance although the modal response is identical with the demands of the normative model."). Although Sunstein and Thaler's argument suggests that the libertarian paternalist should treat a statistically significant research result (e.g., a statistically significant percentage of the subjects exhibited a preference reversal) as practically significant for policy purposes, that conclusion does not follow. See Mitchell, *supra* note 8, at 1954-60 (distinguishing between statistical and practical significance, the latter of which is more important to policymakers but which cannot be inferred from the former).

⁹⁶ See generally KEITH E. STANOVICH, WHO IS RATIONAL? STUDIES OF INDIVIDUAL DIFFERENCES IN REASONING (1999). Below I discuss individual differences in self-control, but self-control is only one of several dimensions along which individuals exhibit differing degrees of rationality. For a discussion of several of these other dimensions, see Mitchell, *supra* note 10, at 83-105, 139-60. In addition to self-control differences, perhaps most relevant to the policies of libertarian paternalism are (1) the substantial differences in risk attitudes and risk-taking propensities between men and women, which lead women as a group to be exhibit a greater tendency toward

we see that many libertarian paternalist policies are likely to have redistributive effects on true preferences, with rational persons bearing the costs of paternalistic changes to the preferences of irrational persons.⁹⁷

A fundamental tenet of libertarianism is that “no social goal can justify forcing an innocent individual to be a resource for others.”⁹⁸ Thus, to the extent that libertarian paternalism

gain/loss framing effects than men as a group, and (2) substantial sex differences in levels of overconfidence with respect to financial decision-making, with men exhibiting greater overconfidence. See, e.g., Brad M. Barber & Terrance Odean, *Boys Will Be: Gender, Overconfidence, and Common Stock Investment*, 116 Q.J. ECON. 261, 289 (2001) (“Psychological research has established that men are more prone to overconfidence than women, particularly so in male-dominated realms such as finance.”); N.S. Fagley & Paul M. Miller, *The Effect of Framing on Choice: Interactions with Risk-Taking Propensity, Cognitive Style, and Sex*, 16 PERS. & SOC. PSYCHOL. BULL. 496, 504-505, 507-09 (1990) (finding sex differences in framing effects, with only women being affected by the frame); N.S. Fagley & Paul M. Miller, *Framing Effects and Arenas of Choice: Your Money or Your Life?*, 71 ORG’L BEHAV. & HUMAN DECISION PROCESSES 355, 367-68 (1997) (“when one examines the significant interaction between frame and sex, it is clear that framing affected choice among the women, but not among the men”); Melanie Powell & David Ansic, *Gender Differences in Risk Behaviour in Financial Decision-Making: An Experimental Analysis*, 18 J. ECON. PSYCHOL. 605, 623-24 (1997) (“The evidence supports the view that gender differences in financial risk preference exist in management populations and are not explained by the context instance of familiarity, ambiguity or gains and loss framing. Gender differences in risk propensity are also associated with a difference in decision strategy, which may arise from an underlying difference in motivation.”); see also Itamar Simonson, Thomas Kramer & Maia J. Young, *Effect Propensity*, ORG’L BEHAV. & HUMAN DECISION PROCESSES (forthcoming 2004) (“[S]ome people rather consistently exhibit risk-seeking preferences, others tend to be risk averse, and a third group does not exhibit any consistent risk propensity and their choices are most sensitive to the particular conditions.” (manuscript at 3, on file with author)).

Note that I am speaking here of reliable, or systematic, differences in behavior. If individuals randomly exhibited rational and irrational behaviors, then one could argue that over time individuals would be equally likely to benefit from or bear the costs of a libertarian paternalist policy.

⁹⁷ In these situations, we may assume that the rational person’s holdings are legitimate within a libertarian framework because they did not result from coercive or fraudulent transactions with irrational persons. See Child, *supra* note 15, at 722 (“The notion of a market transaction is a central feature of libertarianism. . . . As usually set out, adults having normal capacities must be free to engage in any market transaction absent force or fraud.” (footnote omitted)). Indeed, in most of the examples Sunstein and Thaler discuss, there is no direct transaction between the rational and irrational that is being affected by the central planner’s paternalism but rather some choice behavior of irrational persons within an institution that has monopoly power over the means of choice manipulation.

⁹⁸ Jeffrey H. Reiman, *The Fallacy of Libertarian Capitalism* 92 ETHICS 85, 85 (1981) (footnote omitted); see also Frank J. Garcia, *Building a Just Trade Order for a New Millennium*, 33 GEO. WASH. INT’L L. REV. 1015, 1053 (2001) (“The basis for libertarian opposition to egalitarian redistribution is that it undercuts formal self-ownership. To a libertarian, egalitarianism undermines moral equality in seeking to respond to natural inequality.”); Thomas Nagel, *Libertarianism Without Foundations*, 85 YALE L.J. 136, 137 (1975) (“[L]ibertarianism exalts the claim of individual freedom of action, and asks why state power should be permitted even the interference represented by progressive taxation and public provision of health care, education, and a minimum standard of living.”).

Certainly many see this tenet as indefensible, including Reiman, as his title suggests. See Reiman, *supra*, at 86 (“I think my argument amounts to a general refutation of any defense of [the doctrine of libertarian capitalism]”). My

seeks to improve the position of irrational individuals at the expense of rational individuals, then the committed libertarian is likely to oppose a libertarian paternalist policy. Such opposition flows not from opposition to the social goal of helping irrational persons *per se*, but from using or disadvantaging rational persons to achieve this goal without obtaining their consent.

If four conditions hold, a libertarian paternalist policy will redistribute resources from rational to irrational individuals: (1) individuals differ in their propensity to act rationally within some domain; (2) within this domain, a rational choice results in greater wealth for the chooser than the irrational choice (or, alternatively, the rational choice is more costly to the public or private institution than the irrational choice); (3) within this domain, the introduction of a libertarian paternalist policy causes some persons to make the rational choice who would not otherwise have done so; (4) within this domain, resources are limited or zero-sum, such that as more individuals make the rational choice, the mean outcome of the rational choice (in terms of wealth) decreases.⁹⁹

We may consider a simplified example from the pension plan domain to see how libertarian paternalism may redistribute wealth from rational to irrational persons.¹⁰⁰ If libertarian paternalism is applied to retirement savings plans that include an employer contribution tied to employee contribution levels, as through, say, implementation of the Save More Tomorrow program for accelerated savings over time, then the four conditions from above

purpose here, however, is not to debate the philosophical merits of libertarian opposition to redistribution but to point out an inconsistency within libertarian paternalism that a committed libertarian would likely find problematic.

⁹⁹ For present purposes, we need not assume an equilibrium point would be reached where the benefits of acting rationally disappear. We need only assume that some redistribution occurs.

¹⁰⁰ For a discussion of the variety of retirement savings plans, see Richard L. Kaplan, *Enron, Pension Policy, and Social Security Privatization*, 46 ARIZ. L. REV. 53, 54-81 (2004).

will likely hold.¹⁰¹ First, individuals vary widely in their degree of self-control and temporal discounting rates, with greater self-control and lower discount rates being associated with greater retirement savings.¹⁰² Second, Sunstein and Thaler assume that the rational choice to make in retirement planning is the choice that increases contributions to the plan,¹⁰³ which concomitantly will increase employer contributions in those plans with an employer match up to some cap. Third, Sunstein and Thaler present evidence that introduction of a libertarian paternalist plan such as Save More Tomorrow, causes a greater number of individuals to contribute more

¹⁰¹ Retirement savings plans serve as one of the main candidates for libertarian paternalist interventions. See Sunstein & Thaler, *supra* note 1, at 1159-60, 1172-73, 1185. Vanguard has implemented a variation on the Save More Tomorrow program, see Vanguard Introduces ‘One Step’ Program to Simplify Retirement Decisions For 401k, available at <http://www.advisorpage.com/modules.php?name=News&file=print&sid=1088> (last visited Sept. 17, 2004), and TIAA-CREF has begun piloting the Save More Tomorrow program at some institutions, see Pilot Program Promotes Saving More for Tomorrow, available at http://www.tiaa-cref.org/siteline/siteline_20_785_12201.html (last visited Sept. 17, 2004).

¹⁰² These individual differences occur both within individuals over time, see, e.g., Leonard Green et al., *Discounting of Delayed Rewards Across the Life Span: Age Differences in Individual Discounting Functions*, 46 BEHAV. PROCESSES 89, 94 (1999) (“As individuals get older, they appear to discount the value of delayed rewards less steeply and, in addition, the shape of the discounting function changes systematically.”), and between individuals, see, e.g. Leonard Green et al., *Temporal Discounting in Choice Between Delayed Rewards: The Role of Age and Income*, 11 PSYCHOL. & AGING 79, 82 (1996) (“The results of this study revealed that delayed rewards were discounted at similar rates by adults of different ages but similar income levels, but at different rates by adults of similar age but different income levels. Specifically, the lower income older age group showed a greater degree of temporal discounting than did either of the upper income groups, but there were no age differences in discounting between the upper income groups.”); John M. Hinson et al., *Impulsive Decision Making and Working Memory*, 29 J. EXPERIMENTAL PSYCHOL.: LEARNING, MEMORY & COGNITION 298, 299 (2003) (“[S]ome people are highly tolerant of delay, whereas others show a strong preference for immediately obtained outcomes. This individual difference in discounting may help explain important functional differences in individual choice and decision making.”).

On the topic of self-control and savings, see, for example, Thaler & Benartzi, *supra* note 36, at S167 (“[S]aving for retirement requires self-control. When surveyed about their low savings rates, many households report that they would like to save more but lack the willpower.”); John Ameriks et al., *Measuring Self-Control* 14 (May 2004) available at <http://papers.nber.org/papers/W10514.pdf> (“In quantitative terms, the equation suggests that the average over-consumer accumulates some 18% less than one with no self-control problem, while the average under-consumer accumulates some 27% more. The finding of a powerful impact of self-control on wealth accumulation is very robust.” (footnote omitted)).

¹⁰³ See Sunstein & Thaler, *supra* note 1, at 1172-73 (arguing that steering employees to joining a 401(k) plan promotes employee welfare); *id.* at 1185 (describing increased enrollment and savings rates from Save More Tomorrow plan as an example of “successful libertarian paternalism in action”). Cf. Amy B. Monahan, *Addressing the Problem of Impatients, Impulsives and Other Imperfect Actors in 401(k) Plans*, 23 VA. TAX REV. 471, 486-87 (2004) (“There are, of course, rational actors who make informed decisions regarding retirement savings and who elect to contribute below the level economists would consider adequate. However, there is also evidence that many participants irrationally elect a contribution rate that is too low.” (footnote omitted)).

towards retirement than they would have otherwise.¹⁰⁴ Fourth, employers confronted with increased participation must either redistribute funds among plan participants by reducing individual match amounts or infuse the plan with additional funds that may result in degradations of other employee benefits (alternatively, public employers could seek to pass the cost on to taxpayers, which may still have very marginal redistributive effects from the rational to the irrational, or private employers may seek to pass the increased plan costs on to customers).¹⁰⁵

Or consider Sunstein and Thaler's libertarian paternalist cafeteria director who seeks to fight obesity. If the cafeteria director helps diners with low self-control gain some greater control over their impulses and eat less, then she may need to raise prices, meaning that diners with high self-control may now bear some of the cost of the libertarian paternalist policy.¹⁰⁶ In general, for any domain where resources are limited and the increased costs of a libertarian paternalist policy cannot be externalized, the cost of increasing benefits to irrational persons is likely to be borne by rational persons. Such redistribution from rational to irrational individuals

¹⁰⁴ See *id.* at 1161 ("It is now clear that if an employer requires employees to make an affirmative election in favor of savings, with the default rule devoting 100 percent of wages to current income, the level of savings will be far lower than if the employer adopts an automatic enrollment program from which employees are freely permitted to opt out." (footnote omitted)); *id.* at 1172 ("In companies that offer a 'match' . . . most employees eventually do join the plan, but enrollments occur much sooner under automatic enrollment.").

¹⁰⁵ For instance, mandated vesting in retirement plans has arguably led to overall reductions in benefits. See Peter M. van Zante, *Mandated Vesting: Suppression of Voluntary Retirement Benefits*, 75 NOTRE DAME L. REV. 125, 217 (1999) ("Mandated vesting probably has reduced the fraction of workers covered by private retirement plans and the benefits that long-tenure employees receive."). Also, Sunstein has discussed how a switch in an employment default rule for the purpose of redistributing resources from employers to employees may not be particularly effective as employers recoup the costs caused by the switch. See Sunstein, *supra* note 68, at 127 ("But there is a significant risk that any redistributive gain will be nullified, at least in part, through readjustment of the wage package. On plausible assumptions, workers will lose, in wages, some or much of what they gain as a result of the switched entitlement." (footnote omitted)).

¹⁰⁶ One could argue, alternatively, that the low self-control diners who were overeating were keeping the price of the foods eaten by the high self-control diners artificially low. Also, the cafeteria owner may try only to redirect diners to healthier foods that are priced the same as unhealthy foods. In the unlikely case of no change in consumption and revenue, then the libertarian paternalist approach would have no redistributive effects.

should be objectionable to the committed libertarian, at least to the extent the libertarian paternalism results from coercive government intervention.¹⁰⁷

This redistributive concern is likely to apply with greater force to the welfare-focused paternalism found within Sunstein and Thaler's libertarian paternalism than the liberty-focused paternalism I discuss in Part II, because Sunstein and Thaler's central planner seeks to change the welfare of the irrational and often these welfare changes will be tied to changes in wealth. Nevertheless, the liberty-focused paternalism discussed above may also lead to changes in the choice behavior of the irrational that impose costs on the rational.¹⁰⁸ In any case where the liberty-enhancing default rule results in reduced consumption by the irrational or coincides with material welfare increases by the irrational, the rational person's non-fraudulent arbitrage

¹⁰⁷ Self-control is only one of several dimensions along which individuals exhibit different degrees of rationality. For a discussion of several of these other dimensions, see Mitchell, *supra* note 10, at 83-105, 139-60. In addition to self-control differences, perhaps most relevant to the policies of libertarian paternalism are (1) the substantial differences in risk attitudes and risk-taking propensities between men and women, which lead women as a group to be exhibit a greater tendency toward gain/loss framing effects than men as a group, and (2) substantial sex differences in levels of overconfidence with respect to financial decision-making, with men exhibiting greater overconfidence. See, e.g., Brad M. Barber & Terrance Odean, *Boys Will Be: Gender, Overconfidence, and Common Stock Investment*, 116 Q. J. ECON. 261, 289 (2001) ("Psychological research has established that men are more prone to overconfidence than women, particularly so in male-dominated realms such as finance."); N.S. Fagley & Paul M. Miller, *The Effect of Framing on Choice: Interactions with Risk-Taking Propensity, Cognitive Style, and Sex*, 16 PERS. & SOC. PSYCHOL. BULL. 496, 504-505, 507-09 (1990) (finding sex differences in framing effects, with only women being affected by the frame); N.S. Fagley & Paul M. Miller, *Framing Effects and Arenas of Choice: Your Money or Your Life?*, 71 ORG'L BEHAV. & HUMAN DECISION PROCESSES 355, 367-68 (1997) ("when one examines the significant interaction between frame and sex, it is clear that framing affected choice among the women, but not among the men"); Melanie Powell & David Ansic, *Gender Differences in Risk Behaviour in Financial Decision-Making: An Experimental Analysis*, 18 J. ECON. PSYCHOL. 605, 623-24 (1997) ("The evidence supports the view that gender differences in financial risk preference exist in management populations and are not explained by the context instance of familiarity, ambiguity or gains and loss framing. Gender differences in risk propensity are also associated with a difference in decision strategy, which may arise from an underlying difference in motivation."); see also Itamar Simonson, Thomas Kramer & Maia J. Young, *Effect Propensity*, ORG'L BEHAV. & HUMAN DECISION PROCESSES (forthcoming 2004) ("[S]ome people rather consistently exhibit risk-seeking preferences, other tend to be risk averse, and a third group does not exhibit any consistent risk propensity and their choices are most sensitive to the particular conditions." (manuscript at 3, on file with author)).

¹⁰⁸ See the discussion above of how welfare-focused and liberty-focused paternalism may lead to similar outcomes in Sunstein and Thaler's cafeteria example. See *supra* notes 71-72 & accompanying text.

opportunities and material resources obtained through voluntarily transactions with other rational persons may be adversely affected.¹⁰⁹

Although the discussion here has emphasized redistribution from the rational to the irrational to demonstrate how libertarian paternalism may interfere with rational preferences, the libertarian need not limit his opposition to cases of intra-institution redistribution (i.e., cases of redistribution from the rational to the irrational among the class of people targeted by the libertarian paternalist planner). The libertarian may find equally objectionable efforts to externalize the costs of government-sponsored libertarian paternalism to those outside the institution. That is, the costs of government-sponsored libertarian paternalism may be conceived of as taxes on both the rational members of an institution and otherwise disinterested third parties that are imposed for the benefit of the irrational.¹¹⁰ Certainly there may be an argument that could quell the libertarian objection to such redistribution, but Sunstein and Thaler do not offer one.¹¹¹ Absent such an argument, the committed libertarian may well conclude that it is better to

¹⁰⁹ Thanks to Adam Hirsch and Jon Klick for pushing me on this point.

¹¹⁰ Nozick, of course, equates taxation with forced labor. See NOZICK, *supra* note 86, at 169 (“Taxation of earning from labor is on a par with forced labor.”).

¹¹¹ Sunstein and Thaler seek to provide a libertarian-friendly justification for libertarian paternalism without recourse to third-party effects as a justification. See Sunstein & Thaler, *supra* note 1, at 1162 (“The paternalistic aspect consists in the claim that it is legitimate for private and public institutions to attempt to influence people’s behavior even when third-party effects are absent.”). To be a part of the project reconciling libertarianism and paternalism, such an argument would have to show how libertarian paternalism, despite its redistributive potential, does not do harm to libertarian values. Loevinshohn provides an argument for why libertarians should favor some redistribution of property over none that echoes Sunstein and Thaler’s claim about the inevitability of paternalism and might suggest a possible reconciling argument for Sunstein and Thaler. Loevinshohn argues that governments inevitably curtail liberty even in the minimal state favored by libertarians, for the legal penalties protecting against coercive takings will cause some persons to refrain from liberty-enhancing actions. Loevinshohn contends that some redistribution of property can be used to bring in line wants and opportunities for voluntary transfers and thereby reduce the number of involuntary transfers, a result that should be favorable to libertarians if correct. See Ernest Loevinshohn, *Liberty and the Redistribution of Property*, 6 PHIL. & PUB. AFFAIRS 226, 239 (1977) (“by redistributing property, a government can decrease the overall extent to which people’s desires for material goods go unsatisfied, so that overall it will become less important to people to use or consume property from which they are legally barred”). Similarly, Sunstein and Thaler might contend that libertarian paternalism, by bringing actions and welfare more in line for the irrational, will reduce the overall level of redistributive aid to the irrational. Thus, some minimal intrusion on the liberty of rational persons and disinterested third parties should be allowed to avoid greater intrusion overall. This argument takes advantage of the political reality that some welfare redistribution is certain to

leave the irrational to fend for themselves than to try to better them at the expense of rational or innocent others.¹¹²

IV. Conclusion

On close scrutiny, Sunstein and Thaler's attempt to reconcile the traditionally opposed concepts of libertarianism and paternalism fails to hold together. Most significantly, Sunstein and Thaler neglect alternative approaches to dealing with irrational choice behavior that are more consistent with libertarian principles and that make choice-framing paternalism evitable, they would subjugate the liberty of irrational individuals to the central planner's paternalistic welfare judgments, and they fail to deal with the redistributive consequences of libertarian paternalism. Libertarian paternalism, as currently formulated, is not designed to liberate individuals from their irrational tendencies but to capitalize on irrational tendencies to move citizens in directions that the paternalistic planner deems best. It is true that libertarian paternalism leaves rational persons a way out of the central planner's paternalism, but often the exit will not be costless, as the paternalistic costs of trying to improve the welfare of irrational persons are shifted to the rational persons.

occur, and so the realist libertarian should favor programs that reduce overall levels of redistribution rather than seek the unachievable truly minimal, "night-watchman" state.

¹¹² It would be possible to formulate scenarios in which libertarian concerns about self-ownership and opposition to redistribution appear "fanatical," and Sunstein and Thaler feel no obligation to deal with such instances of libertarianism, *see* Sunstein & Thaler, *supra* note 1, at 1167 n.22 ("We do not disagree with the view that autonomy has claims of its own, but we believe it would be fanatical, in the settings we discuss, to treat autonomy, in the form of freedom of choice, as a kind of trump not to be overridden on consequentialist grounds."), but it is also possible to formulate scenarios where libertarian opposition to redistribution to favor the irrational appears quite reasonable in light of the relative costs and benefits of a libertarian paternalist policy or genuine ambiguities about what options make people "better off" in the long run (indeed, the redistributive examples in the text would surely raise questions about whether the gains in supposed welfare are worth the sacrifices to liberty among even non-fanatical but committed libertarians). Thus, it would seem insufficient in a project aiming to reconcile libertarianism with paternalism to simply dismiss out of hand the core libertarian concern about governmental redistribution of resources, especially if the goal is to influence committed libertarians and even dogmatic anti-paternalists.

Notwithstanding libertarian paternalism's shortcomings from a libertarian perspective, Sunstein and Thaler craft a clever argument that draws on an impressive range of evidence to reconsider the relation of libertarianism to paternalism, and they offer examples of libertarian paternalistic policies that I suspect even libertarians will find intuitively compelling. Who could oppose programs to improve the self-control of Americans so that they achieve better health, save more wisely, and make their acts of benevolence more meaningful, especially when told that the affected individuals can freely opt out of these programs?¹¹³ But the argument for libertarian paternalism functions better as an "intuition pump" that drives intuitions toward a desired end, than as a carefully constructed argument for limited paternalism from the perspective of a committed libertarian.¹¹⁴

As committed behavioralists, Sunstein and Thaler undoubtedly know that the intuitive mode of thought will often trump the analytical mode of thought with respect to the creation of beliefs and preferences.¹¹⁵ Therefore, it is often wise to make intuitively appealing arguments rather than drab, analytical arguments when trying to persuade others. And so what we may

¹¹³ With respect to the last item in the list, on benevolence, see Sunstein and Thaler's discussion of "libertarian benevolence." See Sunstein & Thaler, *supra* note 1, at 1191-93 (discussing how the libertarian paternalist approach may be utilized to improve assistance to vulnerable people, such as through a change in the default rule on organ donation).

¹¹⁴ See DANIEL C. DENNETT, *ELBOW ROOM: THE VARIETIES OF FREE WILL WORTH WANTING* 12 (1984) ("A popular strategy in philosophy is to construct a certain sort of thought experiment I call an *intuition pump*. Such thought experiments (unlike Galileo's or Einstein's, for instance) are *not* supposed to clothe strict arguments that prove conclusions from premises. Rather, their point is to entrain a family of imaginative reflections in the reader that ultimately yields not a formal conclusion but a dictate of "intuition." Intuition pumps are cunningly designed to focus the reader's attention on "the important" features, and to deflect the reader from bogging down in hard-to-follow details.").

¹¹⁵ See, e.g., Daniel Kahneman, *A Psychological Point of View: Violations of Rational Rules as a Diagnostic of Mental Processes*, 23 *BEHAV. & BRAIN SCI.* 681, 682 (2000) ("Tversky and I always thought of the heuristics and biases approach as a two-process theory. We proposed that intuitive judgments of probability or frequency are based on 'natural assessments'—that is, of similarity, causality, affective valence, or past frequency—which are effortless and largely automatic. In the terminology of the target article, judgment by heuristics is a manifestation of System 1. However, we also believed that System 2 can override intuitive judgments that conflict with a general rule—but only if the relevant rule is evoked." (citation omitted)); see also Cass R. Sunstein, *Moral Heuristics*,

have here is the ultimate example of libertarian paternalism at work: Sunstein and Thaler framing their argument to take advantage of the cognitive limitations of the libertarian to lead him to believe that libertarian paternalism really is good for him. It might just work, but perhaps this brief response will serve some debiasing function and help libertarians to form rational beliefs about the desirability of libertarian paternalism.