

Preferences, Paternalism, and Liberty

CASS R. SUNSTEIN

RICHARD H. THALER

Our goal in this chapter is to draw on empirical work about preference formation and welfare to propose a distinctive form of paternalism, libertarian in spirit, one that should be acceptable to those who are firmly committed to freedom of choice on grounds of either autonomy or welfare. Indeed, we urge that a kind of 'libertarian paternalism' provides a basis for both understanding and rethinking many social practices, including those that deal with worker welfare, consumer protection, and the family.

In the process of defending these claims, we intend to make some objections to widely held beliefs about both freedom of choice and paternalism. Our major emphasis is on the fact that in many domains, people lack clear, stable, or well-ordered preferences. What they 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. If social planners are asked to respect preferences, or if they are told that respect for preferences promotes well-being, they will often be unable to know what they should do.

Consider the question whether to undergo a risky medical procedure. When people are told, 'Of those who undergo this procedure, 90 percent are still alive after five years,' they are far more likely to agree to the procedure than when they are told, 'Of those who undergo this procedure, 10 percent are dead after five years' (Redelmeier, Rozin, & Kahneman, 1993, p. 73). What, then, are the patient's 'preferences' with respect to this procedure? Repeated experiences with such problems might be expected to eliminate this framing effect, but doctors too are vulnerable to it. Or consider the question of savings for retirement. 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

Richard H. Thaler

(Choi et al., 2002, p. 70; Madrian & Shea, 2001, pp. 1149–1150). Can workers then be said to have well-defined preferences about how much to save? This simple example can be extended to many situations involving the behavior of workers, consumers, voters, and family members.

As the savings problem illustrates, the design features of both legal and organizational rules have surprisingly powerful influences on people's choices. Preferences are formed in part by reference to those influences. We urge that the relevant 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, libertarian paternalists urge that people should be 'free to choose' (Friedman & Friedman, 1980). Hence we do not aim to defend any approach that blocks individual choices.

The paternalistic aspect consists in the claim that it is legitimate for private and public institutions to attempt to influence people's choices and preferences, even when third-party effects are absent. 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 choosers' 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 (see also VanDeVeer 1986, p. 22). Drawing on some well-established findings in behavioral economics and cognitive psychology, we emphasize the possibility that in some cases individuals make inferior decisions in terms of their own welfare—decisions that they would change if they had complete information, unlimited cognitive abilities, and no lack of self-control (Jolls, Sunstein, & Thaler, 1998, pp. 1477–1479). In addition, the notion of libertarian paternalism can be complemented by that of *libertarian benevolence*, by which plan design features such as default rules, framing effects, and starting points are enlisted in the interest of vulnerable third parties. We shall devote some discussion to this possibility.

Libertarian paternalism is a relatively weak and nonintrusive type of paternalism, because choices are not blocked or fenced off. In its most cautious forms, libertarian paternalism imposes trivial costs on those who seek to depart from the planner's preferred option. But the approach we recommend nonetheless counts as paternalistic, because private and public planners are not trying to

Preferences, Paternalism, and Liberty

track people's anticipated choices, but are self-consciously attempting to move people in welfare-promoting directions. It follows that one of our principal targets is the dogmatic anti-paternalism of numerous analysts of law and policy. We believe that this dogmatism 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 their 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. In fact, we do not think that anyone believes it on reflection. Suppose that a chess novice were to play against an experienced player. Predictably the novice would lose precisely because he made inferior choices—choices that could easily be improved by some helpful hints. More generally, how well people choose is an empirical question, one whose answer is likely to vary across domains.

As a first approximation, it seems reasonable to say that people make better choices in contexts in which they have experience and good information (say, choosing ice cream flavors) than in contexts in which they are inexperienced and poorly informed (say, choosing among medical treatments or investment options). So long as people are not choosing perfectly, it is at least possible that some policy could make them better off by improving their decisions.

The first misconception is that preferences predate social contexts and hence 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 and often even what they prefer. We are emphasizing, then, the possibility that people's preferences, in certain domains and across a certain range, are influenced by the choices made by planners (even those who do not understand themselves as such).

As a simple example, consider the cafeteria at some organization. The cafeteria must make a multitude of decisions, including which foods to serve, which ingredients to use, and in what order to arrange the choices. Suppose that the director of the cafeteria notices that customers have a tendency to choose more of the items that are presented earlier in the line. How should the director decide in what order to present the items? To simplify, consider

Richard H. Thaler

some alternative strategies that the director might adopt in deciding which items to place early in the line:

1. She could make choices that she thinks would make the customers best off, all things considered.
2. She could make choices at random.
3. She could choose those items that she thinks would make the customers as obese as possible.
4. She could give customers what she thinks they would choose on their own.

Option 1 appears to be paternalistic, but would anyone advocate options 2 or 3? Option 4 is what many anti-paternalists would favor, but it is much harder to implement than it might seem. 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.

The second misconception is that paternalism always involves coercion. As the cafeteria example illustrates, the choice of the order in which to present food items does not coerce anyone to do anything, yet one might prefer some orders to others on grounds that are paternalistic in the sense that we use the term. Would anyone object to putting the fruit and salad before the desserts at an elementary school cafeteria if the result were to increase the consumption ratio of apples to Twinkies? Is this question fundamentally different if the customers are adults? Since no coercion is involved, we think that some types of paternalism should be acceptable to even the most ardent libertarian. This point has large implications for planners who are seeking to promote social welfare.

Once it is understood that some organizational decisions are inevitable, that preferences are endogenous to social situations, 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. The Rationality of Choices

The presumption that individual choices should be respected is often based on the claim that people do an excellent job of making choices that promote their welfare, or at least that they do a far better job than third parties could possibly do.¹ As far as we can tell, there is little empirical support for this claim, at least if it is offered in this general form. Consider the issue of obesity. Rates of obesity in the United States are now approaching 20 percent, and over 60 percent of Americans are considered either obese or overweight. These numbers reflect a 61 percent increase in obesity from 1991 to 2001, with 38.8 million Americans now qualifying as obese (Centers for Disease Control and Prevention, 2003). There is a great deal of evidence that obesity causes serious health risks, frequently leading to premature death (Calle, Thun, Petrelli, Rodriguez, & Heath, 1999; National Institute of Diabetes & Digestive & Kidney Diseases, 2001). It is quite fantastic to suggest that everyone is choosing the optimal diet, or a diet that is preferable to what might be produced with third-party guidance.

Of course, rational people care about the taste of food, not simply about health, and we do not claim that everyone who is overweight is necessarily failing to act rationally. It is the strong claim that all or almost all Americans are choosing their diet *optimally* that we reject as untenable. What is true for diets is true as well for much other risk-related behavior, including smoking and drinking, which produce many thousands of premature deaths each year (Sunstein, 2002, pp. 8–9). In these circumstances, people's

¹ It is not always based on this claim. Some of the standard arguments against paternalism rest not on consequences but on autonomy—on a belief that people are entitled to make their own choices even if they err. Thus Mill (1972, p. 69) advances a mix of autonomy-based and consequentialist claims. Our principal concern here is with welfare and consequences, though as we suggest below, freedom of choice is sometimes an ingredient in welfare. We do not disagree with the view that autonomy has claims of its own, but we believe that it would be fanatical, in the settings that we discuss, to treat autonomy, in the form of freedom of choice, as a kind of trump not to be overridden on consequentialist grounds. In any case, the autonomy argument is undermined by the fact, discussed in Part II, that sometimes preferences and choices are a function of given arrangements. Most importantly, we think that respect for autonomy is adequately accommodated by the libertarian aspect of libertarian paternalism, as discussed below.

Richard H. Thaler

choices cannot reasonably be thought, in all domains, to be the best means of promoting their well-being.

On a more scientific level, research by psychologists and economists over the past three decades has raised questions about the rationality of many of our judgments and decisions. People fail to make forecasts that are consistent with Bayes's rule (Grether, 1980); use heuristics that can lead them to make systematic blunders (Kahneman & Frederick, 2002, p. 53; Tversky & Kahneman, 1973; Tversky & Kahneman, 1974); exhibit preference reversals (that is, they prefer A to B *and* B to A) (Thaler, 1992, pp. 79–91; Sunstein, Kahneman, Schkade, & Ritov, 2002); suffer from problems of self-control (Frederick, Loewenstein, & O'Donoghue, 2002, pp. 367–368); and make different choices depending on the framing of the problem (Camerer, 2000, pp. 294–295; Johnson, Hershey, Meszaros, & Kunreuther, 2000, pp. 224, 238). 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 (De Bondt & Thaler, 1990; Shiller, 2000, pp. 135–147; Camerer & Hogarth, 1999).

We do not intend to outline all of the relevant evidence here, but consider an illustration from the domain of savings behavior. Benartzi and Thaler (2002) have investigated how much investors like the portfolios they have selected in their defined contribution savings plans. Employees volunteered to share their portfolio choices with the investigators by bringing a copy of their most recent statement to the lab. They were then shown the probability distributions of expected retirement income for three investment portfolios simply labeled A, B, and C. Unbeknownst to the subjects, the three portfolios were their own and portfolios mimicking the average and median choices of their fellow employees. The distributions of expected returns were computed using the software of Financial Engines, the financial information company founded by William Sharpe. On average, the subjects rated the average portfolio equally with their own portfolio, and judged the median portfolio to be significantly more attractive than their own. Indeed, only 20 percent of the subjects preferred their own portfolio to the median portfolio. Apparently, people do not gain much, by their own lights, from choosing investment portfolios for themselves.

Or consider people's willingness to take precautions. In general, the decision to buy insurance for natural disasters is a product not

Preferences, Paternalism, and Liberty

of a systematic inquiry into the likely effects on individual welfare, but of recent events (Slovic, Kunreuther, & White, 1974, p. 14; Kunreuther, 1996, pp. 174–178). If floods have not occurred in the immediate past, people who live on flood plains are far less likely to purchase insurance (Kunreuther, 1996, pp. 176–177). In the aftermath of an earthquake, the level of insurance coverage for earthquakes rises sharply—but it declines steadily from that point, as vivid memories recede (Kunreuther, 1996, pp. 176–177; Slovic et al., 1974, p. 14). Findings of this kind do not establish that people's choices are usually bad or that third parties can usually do better. But they do show that some of the time, people do not choose optimally even when the stakes are high.

It is true that people sometimes respond to their own bounded rationality by, for example, hiring agents or delegating decisions to others (Sunstein & Ullman-Margalit, 1999). 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. 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.

In any event, our emphasis here is not on blocking choices, but on strategies that move people in welfare-promoting directions while also allowing freedom of choice. Evidence of bounded rationality and problems of self-control is sufficient to suggest that such strategies are worth exploring. Of course many people value freedom of choice as an end in itself, but they should not object to approaches that preserve that freedom while also promising to improve people's lives.

II. Is Paternalism Inevitable? On the Endogeneity of Preferences

A few years ago, the tax law was changed so that employees could pay for employer-provided parking on a pre-tax basis (Energy Policy Act of 1992, 2000). Previously, such parking had to be paid for with after-tax dollars. Our employer, and the employer of some of our prominent anti-paternalist colleagues, sent around an announcement of this change in the law, and adopted the following policy: Unless the employee notified the payroll department, deductions for parking would be taken from pre-tax rather than

Richard H. Thaler

post-tax income. In other words, the University of Chicago decided that the default option would be to pay for parking with pre-tax dollars, but employees could opt out of this arrangement and pay with after-tax dollars. Call this choice Plan A. An obvious alternative, Plan B, would be to announce the change in the law and tell employees that if they want to switch to the new pre-tax plan they should return some form electing this option. The only difference between the two plans is the default. Under Plan A the new option is the default, whereas under Plan B the status quo is the default. We will refer to the former as an 'opt-out' strategy and the latter as an 'opt-in' strategy.

How should the university choose between opt-in and opt-out? In the parking example, it seems to be the case that every employee would prefer to pay for parking with pre-tax dollars rather than after-tax dollars. Since the cost savings are substantial (parking costs as much as \$1200 per year) and the cost of returning a form is trivial, standard economic theory predicts that the university's choice will not really matter. Under either plan, all employees would choose (either actively under Plan B or by default under Plan A) the pre-tax option. In real life, however, had the university adopted Plan B, we suspect that many employees, especially faculty members (and probably including the present authors), would still have that form buried somewhere in their offices and would be paying substantially more for parking on an after-tax basis. In short, the default plan would have had large effects on behavior.

Throughout we shall be drawing attention to the effects of default plans on choices. Often those plans will be remarkably 'sticky.' Often people's choices, and even their valuations, are endogenous to the social context, including default rules. This point raises a serious problem for those who reject paternalism in the name of liberty, and who argue that people should be permitted to choose in accordance with their preferences.

A Savings and Employers

1. Data and default rules

Our conjecture that default plans affect outcomes is supported by the results of numerous experiments documenting a 'status quo' bias (Kahneman, Knetsch, & Thaler, 1991, pp. 197–199; Samuelson & Zeckhauser, 1988). The existing arrangement, whether set out by private institutions or by government, is often robust. One

Preferences, Paternalism, and Liberty

illustration of this phenomenon comes from studies of automatic enrollment in 401(k) employee savings plans (Choi et al., 2002, p. 70; Madrian & Shea, 2001, pp. 1149–1150), and we now elaborate the brief account with which we began. Most 401(k) plans use an opt-in design. When employees first become eligible to participate in the 401(k) plan, they receive some plan information and an enrollment form that must be completed in order to join. Under the alternative of automatic enrollment, employees receive the same information but are told that unless they opt out, they will be enrolled in the plan (with default options for savings rates and asset allocation). In companies that offer a ‘match’ (the employer matches the employee’s contributions according to some formula, often a 50 percent match up to some cap), most employees eventually do join the plan, but enrollments occur much sooner under automatic enrollment. For example, Madrian and Shea found that initial enrollments jumped from 49 percent to 86 percent (Madrian & Shea, 2001, pp. 1158–1159), and Choi et al. (2002, pp. 76–77) found similar results.²

Should the adoption of automatic enrollment be considered paternalistic? And if so, should it be seen as a kind of officious meddling with employee preferences? We answer these questions yes and no respectively. If employers think (correctly, we believe) that most employees would prefer to join the 401(k) plan if they took the time to think about it and did not lose the enrollment form, then by choosing automatic enrollment, they are acting paternalistically by our definition of the term. They are not attempting to protect against harms to third parties, but to steer employees’ choices in directions that will, in the view of employers, promote employees’ welfare. Since no one is forced to do anything, we think that this steering should be considered unobjectionable

² In a separate phenomenon, the default rule also had a significant effect on the chosen contribution rate (Madrian & Shea, pp. 116). The default contribution rate (3 percent) tended to stick; a majority of employees maintained that rate even though this particular rate was chosen by around 10 percent of employees hired before the automatic enrollment. The same result was found for the default allocation of the investment: While less than 7 percent of employees chose a 100 percent investment allocation to the money market fund, a substantial majority (75 percent) of employees stuck with that allocation when it was the default rule. The overall default rate (participation in the plan, at a 3 percent contribution rate, investing 100 percent in the money market fund) was 61 percent, but only 1 percent of employees chose this set of options prior to their adoption as defaults.

Richard H. Thaler

even to committed libertarians. The employer must choose some set of rules, and either plan affects employees' choices. No law of nature says that in the absence of an affirmative election by employees, 0 percent of earnings will go into a retirement plan. Because both plans alter choices, neither one can be said, more than the other, to count as a form of objectionable meddling.

2. Skeptics

Skeptical readers, insistent on freedom of choice, might be tempted to think that there is a way out of this dilemma. Employers could avoid choosing a default if they *required* employees to make an active choice, either in or out. Call this option *required active choosing*. Undoubtedly required active choosing is attractive in some settings, but a little thought reveals that this is not at all a way out of the dilemma. On the contrary, required 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. Some employees may not want to have to make a choice (and might make a second-order choice not to have to do so). Why should employers force them to choose?

Required active choosing honors freedom of choice in a certain respect; but it does not appeal to those who would choose not to choose, and indeed it will seem irritating and perhaps unacceptably coercive by their lights. In some circumstances, required choosing will not even be feasible. In any case, an empirical question remains: What is the effect of forced choosing? Choi et al. (2002, pp. 77, 86) find that required active choosing increases enrollments relative to the opt-in rule, though not by as much as automatic enrollment (opt-out). Our discussion in Part III below offers some suggestions about the circumstances in which it makes most sense to force people to choose.

Other skeptics might think that employers should avoid paternalism by doing what most employees would want employers to do. On this approach, a default rule can successfully avoid paternalism if it tracks employees' preferences. Sometimes this is a plausible solution. But what 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. The choices employees will make depend on the way the employer frames those choices. Employee 'preferences,' as such, do not exist in those circumstances.

Preferences, Paternalism, and Liberty

We think that savings is a good example of a domain in which preferences are likely to be ill-defined. Few households have either the knowledge or inclination to calculate their optimal life-cycle savings rate, and even if they were to make such a calculation, its results would be highly dependent on assumptions about rates of return and life expectancies. In light of this, actual behavior is highly sensitive to plan design features.

B. Government

Some enthusiasts for free choice might be willing to acknowledge these points and hence to accept private efforts to steer people's choices in what seem to be the right directions. Market pressures, and the frequently wide range of possible options, might be thought to impose sufficient protection against objectionable steering. But our emphasis has been on the inevitability of paternalism, and on this count, the same points apply to some choices made by governments in establishing legal rules.

1. Default rules

Default rules of some kind are inevitable, and much of the time those rules will affect preferences and choices (Sunstein, 2002b; Korobkin, 1998). In the neglected words of a classic article (Calabresi & Melamed, 1972, pp. 1090–1091):

[A] minimum of state intervention is always necessary ... When a loss is left where it falls in an auto accident, it is not because God so ordained it. Rather it is because the state has granted the injurer an entitlement to be free of liability and will intervene to prevent the victim's friends, if they are stronger, from taking compensation from the injurer.

If the entitlement-granting rules seem invisible, and seem to be a simple way of protecting freedom of choice, it is because they appear so sensible and natural that they are not taken to be a legal allocation at all. But this is a mistake. What we add here is that when a default rule affects preferences and behavior, it has the same effect as employer presumptions about savings plans. This effect is often both unavoidable and significant. So long as people can

Richard H. Thaler

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.

Consumers, workers, and married people,³ for example, are surrounded by a network of legal allocations that provide the background against which agreements are made. As a matter of employment law, and consistent with freedom of contract, workers might be presumed subject to discharge 'at will,' or they might be presumed protected by an implied right to be discharged only 'for cause.' They might be presumed to have a right to vacation time, or not. They might be presumed protected by safety requirements, or the employer might be free to invest in safety as he wishes, subject to market pressures. In all cases, the law must establish whether workers have to 'buy' certain rights from employers or vice versa (Sunstein, 2001, pp. 208–212). Legal intervention, in this important sense, cannot be avoided. The same is true for consumers, spouses, and all others who are involved in legal relationships. Much of the time, the legal background matters, even if transaction costs are zero, because it affects choices and preferences, as demonstrated by Korobkin (1998, pp. 633–64) and Kahneman et al. (1991, pp. 194–204). Here, as in the private context, a form of paternalism is unavoidable.

In the context of insurance, an unplanned, natural experiment showed that the default rule can be very 'sticky' (Camerer, 2000, pp. 294–95; Johnson et al., 2000, p. 238). New Jersey created a system in which the default insurance program for motorists included a relatively low premium and no right to sue; purchasers were allowed to deviate from the default program and to purchase the right to sue by choosing a program with that right and also a higher premium. By contrast, Pennsylvania offered a default program containing a full right to sue and a relatively high premium; purchasers could elect to switch to a new plan by 'selling' the more ample right to sue and paying a lower premium. In both cases, the default rule tended to stick. A strong majority accepted the default rule in both states, with only about 20 percent of New Jersey drivers acquiring the full right to sue, and 75 percent of Pennsylvanians retaining that right (Johnson et al., 2000, p. 238). There is no reason to think that the citizens of Pennsylvania have systematically different preferences from the citizens of New Jersey. The default plan is what produced the ultimate effects.

³ Okin (1989) is a good source of general information on marriage and legal rules.

Preferences, Paternalism, and Liberty

Indeed, controlled experiments find the same results, showing that the value of the right to sue is much higher when it is presented as part of the default package (Johnson et al., 2000, pp. 235–238).

In another example, a substantial effect from the legal default rule was found in a study of law student reactions to different state law provisions governing vacation time from firms (Sunstein, 2002b, pp. 113–114). The study was intended to be reasonably realistic, involving as it did a pool of subjects to whom the underlying issues were hardly foreign. Most law students have devoted a good deal of time to thinking about salaries, vacation time, and the tradeoffs between them. The study involved two conditions. In the first, state law guaranteed two weeks of vacation time, and students were asked to state their median willingness to pay (in reduced salary) for two extra weeks of vacation. In this condition, the median willingness to pay was \$6,000. In the second condition, state law provided a mandatory, non-waivable two-week vacation guarantee, but it also provided employees (including associates at law firms) with the right to two additional weeks of vacation, a right that could be ‘knowingly and voluntarily waived.’ Hence the second condition was precisely the same as the first, except that the default rule provided the two extra weeks of vacation. In the second condition, students were asked how much employers would have to pay them to give up their right to the two extra weeks. All by itself, the switch in the default rule more than doubled the students’ responses, producing a median willingness to accept of \$13,000.

We can imagine countless variations on these experiments. For example, the law might authorize a situation in which employees have to opt into retirement plans, or it might require employers to provide automatic enrollment and allow employees to opt out. Both systems would respect the freedom of employees to choose, and either system would be libertarian in that sense. In the same vein, the law might assume that there is no right to be free from age discrimination in employment, permitting employees (through individual negotiation or collective bargaining) to contract for that right. Alternatively, it might give employees a nondiscrimination guarantee, subject to waiver via contract. Our suggestion here is that one or another approach is likely to have effects on the choices of employees. This is the sense in which paternalism is inevitable, from government no less than from private institutions.

Richard H. Thaler

2. Anchors

In emphasizing the absence of well-formed preferences, we are not speaking only of default rules. Consider the crucial role of ‘anchors,’ or starting points, in contingent valuation studies, an influential method of valuing regulatory goods such as increased safety and environmental protection (Bateman & Willis, 1999). Such studies, used when market valuations are unavailable, attempt to ask people their ‘willingness to pay’ for various regulatory benefits. Contingent valuation has become prominent in regulatory theory and practice. Because the goal is to determine what people actually want, contingent valuation studies are an effort to elicit, rather than to affect, people’s values. Paternalism, in the sense of effects on preferences and choices, is not supposed to be part of the picture. But it is extremely difficult for contingent valuation studies to avoid constructing the very values that they are supposed to discover (Payne, Bettman, & Schkade, 1999). The reason is that in the contexts in which such studies are used, people do not have clear or well-formed preferences, and hence it is unclear that people have straightforward ‘values’ that can actually be found. Hence some form of paternalism verges on the inevitable: Stated values will often be affected, at least across a range, by how the questions are set up.

Perhaps the most striking evidence to this effect comes from a study of willingness to pay to reduce annual risks of death and injury in motor vehicles (Jones-Lee & Loomes, 2001, pp. 208–212). The authors of that study attempted to elicit both maximum and minimum willingness to pay for safety improvements. People were presented with a statistical risk and an initial monetary amount, and asked whether they were definitely willing or definitely unwilling to pay that amount to eliminate the risk, or if they were ‘not sure.’ If they were definitely willing, the amount displayed was increased until they said that they were definitely unwilling. If they were unsure, the number was moved up and down until people could identify the minimum and maximum.

The authors were not attempting to test the effects of anchors; on the contrary, they were alert to anchoring only because they ‘had been warned’ of a possible problem with their procedure, in which people ‘might be unduly influenced by the first amount of money that they saw displayed.’ To solve that problem, the study allocated people randomly to two subsamples, one with an initial display of 25 pounds, the other with an initial display of 75 pounds. The authors hoped that the anchoring effect would be small, with no

Preferences, Paternalism, and Liberty

significant consequences for minimum and maximum values. But their hope was dashed. *For every level of risk, the minimum willingness to pay was higher with the 75 pound starting point than the maximum willingness to pay with the 25 pound starting point!* For example, a reduction in the annual risk of death by 4 in 100,000 produced a *maximum* willingness to pay of 149 pounds with the 25 pound starting value, but a *minimum* willingness to pay of 232 pounds with the 75 pound starting value (and a maximum, in that case, of 350 pounds). The most sensible conclusion is that people are sometimes uncertain about appropriate values, and whenever they are, anchors have an effect—sometimes a startlingly large one.

It is not clear how those interested in eliciting (rather than affecting) values might respond to this problem. What is clear is that in the domains in which contingent valuation studies are used, people often lack well-formed preferences, and starting points have important consequences for behavior and choice.

3. Framing

We have suggested that in the important context of medical decisions, framing effects are substantial (Redelmeier et al., 1993, p. 73). Apparently, most people do not have clear preferences about how to evaluate a procedure that leaves 90 percent of people alive (and 10 percent of people dead) after a period of years. A similar effect has been demonstrated in the area of obligations to future generations (Frederick, 2003), a much-disputed policy question (Revesz, 1999, pp. 987–1016; Morrison, 1998). This question does not directly involve paternalism, because those interested in the valuation of future generations are not attempting to protect people from their own errors. But a regulatory system that attempts to track people's preferences would try to measure intergenerational time preferences, that is, to elicit people's judgments about how to trade off the protection of current lives and future lives (Revesz, 1999, pp. 996–1007).

Hence an important question, asked in many debates about the issue, is whether people actually make such judgments and whether they can be elicited. And indeed, an influential set of studies finds that people value the lives of those in the current generation far more than the lives of those in future generations (Cropper, Aydede, & Portney, 1994; Cropper, Aydede, & Portney, 1992, p. 472). From a series of surveys, Maureen Cropper and her coauthors (1994) suggest that people are indifferent between saving

Richard H. Thaler

1 life today and saving 44 lives in 100 years. They make this suggestion on the basis of questions asking people whether they would choose a program that saves '100 lives now' or a program that saves a substantially larger number '100 years from now.'

But it turns out that other descriptions of the same problem yield significantly different results (Frederick, 2003). Here, as in other contexts, it is unclear whether people actually have well-formed preferences with which the legal system can work. For example, most people consider 'equally bad' a single death from pollution next year and a single death from pollution in 100 years—implying no preference for members of the current generation. In another finding of no strong preference for the current generation, people are equally divided between two programs: one that will save 55 lives now and 105 more lives in 20 years; and one that will save 100 lives now and 50 lives 25 years from now. It is even possible to frame the question in such a way as to find that future lives are valued more, not less, highly than current lives. The most sensible conclusion is that people do not have robust, well-ordered intergenerational time preferences. If so, it is not possible for government to track those preferences, because they are an artifact of how the question is put. The point applies in many contexts. For example, people are unlikely to have context-free judgments about whether government should focus on statistical lives or statistical life-years in regulatory policy; their judgments will be much affected by the framing of the question (Sunstein, 2004).

C. Why Effects on Choice Can Be Hard to Avoid

1. Explanations

Why, exactly, do default rules, starting points, and framing effects have such large effects? To answer this question, it is important to make some distinctions.

a) Suggestion

In the face of uncertainty about what should be done, people might rely on one of two related heuristics: do what most people do, or do what informed people do. Choosers might think that the default plan or value captures one or the other. In many settings, any starting point will carry some informational content and will thus

Preferences, Paternalism, and Liberty

affect choices. When a default rule affects behavior, it might well be because it is taken to carry information about how sensible people usually organize their affairs. Notice that in the context of savings, people might have a mild preference for one or another course, but the preference might be overcome by evidence that most people do not take that course. Some workers might think, for example, that they should not enroll in a 401(k) plan and have a preference not to do so; but the thought and the preference might shift with evidence that the employer has made enrollment automatic.

With respect to savings, the designated default plan apparently carries a certain legitimacy for many employees, perhaps because it seems to have resulted from some conscious thought about what makes most sense for most people (Madrian & Shea, 2001). This interpretation is supported by the finding that the largest effects from the new default rule are shown by women and African-Americans. We might speculate 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.

b) Inertia

A separate explanation points to inertia. Any change from the default rule or starting value is likely to require some action. Even a trivial action, such as filling in some form and returning it, can leave room for failures due to memory lapses, sloth, and procrastination. Many people wait until the last minute to file their tax return, even when they are assured of getting a refund. Madrian & Shea (2001, p. 1171) note that, under automatic enrollment, individuals become 'passive savers' and 'do nothing to move away from the default contribution rate.' The power of inertia should be seen as a form of bounded rationality. Although the costs of switching from the default rule or the starting point can be counted as transaction costs, the fact that large behavioral changes are observed even when such costs are tiny suggests that a purely rational explanation is difficult to accept.

c) Endowment effect

A default rule might create a 'pure' endowment effect. It is well known that people tend to value goods more highly if those goods have been initially allocated to them than if those goods have been

Richard H. Thaler

initially allocated elsewhere (Korobkin, 1998; Thaler, 1991). And it is well known that, in many cases, the default rule will create an endowment effect (Kahneman et al., 1991, pp. 197–199; Samuelson & Zeckhauser, 1998). When an endowment effect is involved, the initial allocation, by private or public institutions, affects people's choices simply because it affects their valuations.

d) Ill-formed preferences

In the cases we have discussed, people's preferences are ill-formed and murky. Suppose, for example, that people are presented with various payouts and risk levels for various pension plans. They might be able to understand the presentation; there might be no confusion. But people might not have a well-defined preference for, or against, a slightly riskier plan with a slightly higher expected value. In these circumstances, their preferences might be endogenous to the default plan simply because they lack well-formed desires that can be accessed to overrule the default starting points. In unfamiliar situations, it is especially unlikely that well-formed preferences will exist. The range of values in the highway safety study is likely a consequence of the unfamiliarity of the context, which leaves people without clear preferences from which to generate numbers. The effects of framing on intergenerational time preferences attest to the fact that people do not have unambiguous judgments about how to trade off the interests of future generations with those of people now living.

2. The inevitability of paternalism

For present purposes, the choice among these various explanations does not greatly matter. The central point is that effects on individual choices are often unavoidable. Of course it is usually good not to block choices, and we do not mean to defend non-libertarian paternalism here. But in an important respect the anti-paternalist position is incoherent, simply because there is no way to avoid effects on behavior and choices. The task for the committed libertarian is, in the midst of such effects, to preserve freedom of choice.

Because framing effects are inevitable, it is hopelessly inadequate to say that when people lack relevant information the best response is to provide it. In order to be effective, any effort to inform people

Preferences, Paternalism, and Liberty

must be rooted in an understanding of how people actually think. Presentation makes a great deal of difference: The behavioral consequences of otherwise identical pieces of information depend on how they are framed.

Consider one example from the realm of retirement savings. Benartzi and Thaler (1999) asked participants in a defined contribution savings plan to imagine that they had only two investment options, Fund A and Fund B, and asked them how they would allocate their investments between these two funds. (The two funds were, in fact, a diversified stock fund and an intermediate term bond fund.) All subjects were given information about the historic returns on these funds. However, one group was shown the distribution of annual rates of return, whereas another group was shown *simulated thirty-year rates of return*. The long-term rates of return were derived from the annual rates of return (by drawing years at random from history), and so the two sets of information were, strictly speaking, identical. Nevertheless, participants elected to invest about 40 percent of their money in equities when shown the annual returns and 90 percent when shown the long-term rates of return. The lesson from this example is that plan sponsors cannot avoid influencing the choices their participants make simply by providing information. The way they display the information will, in some situations, strongly alter the choices people make.

The point that the presentation of information influences choice is a general one. In the face of health risks, for example, some presentations of accurate information might actually be counterproductive, because people might attempt to control their fear by refusing to think about the risk at all. In empirical studies, 'some messages conveying identical information seemed to work better than others, and . . . some even appeared to backfire' (Caplin, 2003, p. 443). When information campaigns fail altogether, it is often because those efforts 'result in counterproductive defensive measures.' Hence the most effective approaches go far beyond mere disclosure and combine 'a frightening message about the consequences of inaction with an upbeat message about the efficacy of a proposed program of prevention' (Caplin, 2003, p. 442).

There are complex and interesting questions here about how to promote welfare. If information greatly increases people's fear, it will to that extent reduce welfare—in part because fear is unpleasant, in part because fear has a range of ripple effects producing social costs. We do not speak to the welfare issue here. Our only suggestions are that if people lack information, a great deal of attention needs to be paid to information processing, and

Richard H. Thaler

that without such attention, information disclosure might well prove futile or counterproductive. 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.

D. Beyond the Inevitable (But Still Libertarian)

The inevitability of paternalism is most clear when the planner has to choose starting points or default rules. But if the focus is on welfare, it is reasonable to ask whether the planner should go beyond the inevitable, and whether such a planner can also claim to be libertarian.

In the domain of employee behavior, there are many imaginable illustrations. Employees might be automatically enrolled in a 401(k) plan, with a right to opt out, but employers might require a waiting period, and perhaps a consultation with an adviser, before the opt-out could be effective. Thaler and Benartzi (in press) have proposed a method of increasing contributions to 401(k) plans that also meets the libertarian test. Under the Save More Tomorrow plan, now in place in many institutions, employees are invited to sign up for a program in which their contributions to the savings plan are increased annually whenever they get a raise. Once employees join the plan, they stay in until they opt out or reach the maximum savings rate. In the first company to use this plan, the employees who joined increased their savings rates from 3.5 percent to 11.6 percent in a little over two years (three raises). Very few of the employees who join the plan drop out.

It should now be clear that the difference between libertarian and non-libertarian paternalism is not simple and rigid. The libertarian paternalist insists on preserving choice, whereas the non-libertarian paternalist is willing to foreclose choice. But in all cases, a real question is the cost of exercising choice, and here there is a continuum rather than a sharp dichotomy. A libertarian paternalist who is especially enthusiastic about free choice would be inclined to make it relatively costless for people to obtain their preferred outcomes. (Call this a *libertarian* paternalist.) By contrast, a libertarian paternalist who is especially confident of his welfare judgments would be willing to impose real costs on workers and consumers who seek to do what, in the paternalist's view, would not be in their best interests. (Call this a libertarian *paternalist*.)

Preferences, Paternalism, and Liberty

Rejecting both routes, a non-libertarian paternalist would attempt to block certain choices. But notice that almost any such attempt will amount, in practice, to an effort to impose high costs on those who try to make those choices. Consider a law requiring drivers to wear seat belts. If the law is enforced, and a large fine is imposed, the law is non-libertarian even though determined violators can exercise their freedom of choice—at the expense of the fine. But as the expected fine approaches zero, the law approaches libertarianism.

III. How to Choose: Preference Formation and Welfare

How should sensible planners choose among possible systems, given that some choice is necessary? The promotion of human well-being should be a principal goal, but it is far from clear how to do so. We suggest two approaches. If feasible, a comparison of possible rules should be done using a form of cost-benefit analysis, one that pays serious attention to welfare effects. In many cases, however, such analyses will be both difficult and expensive. As an alternative, we offer some rules of thumb that might be adopted to choose among various options.

A. Costs and Benefits

The goal of a cost-benefit study would be to measure the full ramifications of any design choice. In the context at hand, the cost-benefit study cannot be based on the economists' measure of willingness to pay (WTP), because WTP will be a function of the default rule (Kahneman et al., 1991, pp. 202–203; Korobkin, 1998, pp. 636–641). What is necessary is a more open-ended (and inevitably somewhat subjective) assessment of the welfare consequences. To illustrate, take the example of automatic enrollment. 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.

If the issue were only enrollment, we think it highly likely that the gains would exceed the losses. Because of the right to opt out,

Richard H. Thaler

those who need the money immediately are able to have it. In principle one could also compare the costs of foregone current consumption and the benefits of increased consumption during retirement, though this is, admittedly, difficult to do in practice. 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 (Choi et al., 2002, p. 78; Madrian & Shea, 2001, pp. 1158–1161). 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.

Some readers might think that our reliance on behavior as an indication of welfare is inconsistent with one of our central claims—that choices do not necessarily coincide with welfare. But in fact, there is no inconsistency. Compare rules calling for mandatory cooling-off periods. The premise of such rules is that people are more likely to make good choices when they have had time to think carefully and without a salesperson present. Similarly, it 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.

Once the other effects of automatic enrollment are included, the analysis becomes cloudier. Any plan for automatic enrollment must include a specified default savings rate. Some of those automatically enrolled at a 3 percent savings rate—a typical default in automatic enrollment—would have chosen a higher rate if left to their own devices (Choi et al., 2002, pp. 78–79). If automatic enrollment leads some or many people to save at a lower rate than they would choose, the plan might be objectionable for that reason. Hence we are less confident that this more complete cost-benefit analysis would support the particular opt-out system, though a higher savings rate might well do so. A more sophisticated plan, avoiding some of these pitfalls, is discussed below.

Similar tradeoffs are involved with another important issue: the appropriate default rule for organ donations. In many nations—Austria, Belgium, Denmark, Finland, France, Italy, Luxembourg, Norway, Singapore, Slovenia, and Spain—people are presumed to consent to allow their organs to be used, after death, for the benefit of others; but they are permitted to rebut the presumption, usually through an explicit notation to that effect on their drivers’ licenses (Presumed Consent Foundation, Inc., 2003b). In the United States,

Preferences, Paternalism, and Liberty

by contrast, those who want their organs to be available for others must affirmatively say so, also through an explicit notation on their drivers' licenses. The result is that in 'presumed consent' nations over 90 percent of people consent to make their organs available for donation, whereas in the United States, where people have to take some action to make their organs available, only 28 percent elect to do so (Presumed Consent Foundation, Inc., 2003b; Mardfin, 1998). We hypothesize that this dramatic difference is not a product of deep cultural differences, but of the massive effect of the default rule. Hence we would predict that a European-style opt-out rule in the United States would produce donation rates similar to those observed in the European countries that use this rule. Note in this regard that by one report, over 85 percent of Americans support organ donation—a statistic that suggests opt-outs would be relatively rare (Presumed Consent Foundation, Inc., 2003a).

A recent study strongly supports this prediction. Suggesting that preferences are constructed by social frames, Johnson and Goldstein (2004; Chapter 39) urge that with respect to organ donation, people lack stable preferences and that their decisions are very much influenced by the default rule. A controlled online experiment showed a substantial effect from the default rule: The opt-in system created a 42 percent consent rate, about half of the 82 percent rate for an opt-out system. The real-world evidence is even more dramatic. Presumed consent nations show consent rates ranging from a low of 85.9 percent (Sweden) to a high of 100 percent (Austria), with a median of 99 percent. The default also produces a significant, though less dramatic, increase in actual donations, meaning that many people are saved as a result of the presumed consent system.⁴ There is reason to believe that in the United States, a switch in the default rule could save thousands of lives.

The default rules for organ donation do not fit the usual definition of paternalism. The issue is the welfare of third parties, not of choosers. Here we are speaking not of libertarian paternalism, but of libertarian benevolence: an approach that

⁴ Many factors determine how many organs are actually made available and used for transplants. The transplant infrastructure is certainly important, and fewer organs will be available if family members and heirs can veto transplants, even under a presumed consent regime. Johnson and Goldstein estimate that switching to an opt-out system increases organs actually used by 16 percent, holding everything else constant.

Richard H. Thaler

attempts to promote benevolence, and to assist vulnerable people, without mandating behavior in any way. We suggest that changes in default rules, or a system of Give More Tomorrow, could produce large increases in public assistance—and that such approaches could do so in a way that avoids coercion. With respect to behavior, the analysis of libertarian benevolence is quite similar to that of libertarian paternalism. One of the advantages of that analysis is the demonstration that when third-party interests are at stake, the default rule will matter a great deal. It follows that planners can often deliver significant benefits to third parties simply by switching the default rule. In the case of organ donation, this is what we observe.

Does one or another default rule promote welfare? At first glance, the opt-out rule common in Europe seems better, simply because it should save a large number of lives without compromising any other important value. The most that can be said against the opt-out rule is that through inertia, perceived social pressure, or confusion, some people might end up donating their organs when they would not, all things considered, prefer to do so *ex ante*. (Their *ex post* preferences are difficult to infer!) If this objection (or some other) seems forceful, an alternative would be to require active choices—for example, to mandate, at the time of applying for a driver's license, that applicants indicate whether they want to allow their organs to be used for the benefit of others. We make only two claims about this example. First, the evaluative question turns in large part on empirical issues of the sort that it would be both possible and useful to investigate. Second, the opt-in approach is unlikely to be best.

B. Rules of Thumb

In many cases, the planner will be unable to make a direct inquiry into welfare, either because too little information is available or because the costs of conducting the analysis are not warranted. The committed anti-paternalist might say, in such cases, that people should simply be permitted to choose as they see fit. We hope that we have said enough to show why this response is unhelpful. What people choose often depends on the starting point, and hence the starting point cannot be selected by asking what people choose. In these circumstances, the libertarian paternalist would seek indirect proxies for welfare—methods that test whether one or another

Preferences, Paternalism, and Liberty

approach promotes welfare without relying on guesswork about that question. We suggest three possible methods.

First, the libertarian paternalist might select the approach *that the majority would choose if explicit choices were required and revealed*. In the context of contract law, this is the most familiar inquiry in the selection of default rules (Ayres & Gertner, 1989, pp. 90–91)—provisions that govern contractual arrangements in the absence of express decisions by the parties. Useful though it is, this market-mimicking approach raises its own problems. Perhaps the majority's choices would be insufficiently informed, or a reflection of bounded rationality or bounded self-control. Perhaps those choices would not, in fact, promote the majority's welfare. At least as a presumption, however, it makes sense to follow those choices if the planner knows what they would be. A deeper problem is that the majority's choices might themselves be a function of the starting point or the default rule. If so, the problem of circularity dooms the market-mimicking approach. But in some cases, at least, the majority might go one way or the other regardless of the starting point; and to that extent, the market-mimicking strategy is workable. Note that in the cafeteria example, some options would not fit with the majority's *ex ante* choices (healthy but terrible-tasting food, for example), and that for savings, some allocations would certainly violate the choices of ordinary workers (say, an allocation of 30 percent or more to savings). In fact a clear understanding of majority choices might well support a default rule that respects those choices even if the planner thinks that an inquiry into welfare would support another rule. At the very least, planners should be required to have real confidence in their judgment if they seek to do something other than what a suitably informed majority would find to be in its interest.

Second, the libertarian paternalist might select the approach that we have called required active choices, one *that would force people to make their choices explicit*. This approach might be chosen if the market-mimicking strategy fails, either because of the circularity problem or because the planner does not know which approach would in fact be chosen by the majority. We have seen the possibility of requiring active choices in the context of retirement plans and organ donations; it would be easy to multiply examples. In the law of contract, courts sometimes choose 'penalty defaults'—default rules that penalize the party in the best position to obtain a clear statement on the question at hand, and hence create an incentive for clarity for the person who is in the best position to produce clarity (Ayres & Gertner, 1989, pp. 101–106).

Richard H. Thaler

Libertarian paternalists might go along the same track; in fact penalty defaults can be seen as a form of libertarian paternalism.

Here too, however, there is a risk that the choices that are actually elicited will be inadequately informed or will not promote welfare. In the case of retirement plans, for example, forced choices have been found to produce higher participation rates than requiring opt-ins, but lower rates than requiring opt-outs (Choi et al., 2002, pp. 77, 86). If it is likely that automatic enrollment promotes people's welfare, perhaps automatic enrollment should be preferred over requiring active choices. The only suggestion is that where social planners are unsure how to handle the welfare question, they might devise a strategy that requires people to choose.

Third, the libertarian paternalist might select the approach *that minimizes the number of opt-outs*. Suppose, for example, that when drivers are presumed to want to donate their organs to others, only 10 percent opt out, but that when drivers are required to signal their willingness to donate their organs to others, 30 percent opt in. This is an ex post inquiry into people's preferences, in contrast to the ex ante approach favored by the market-mimicking strategy. With those numbers, there is reason to think that the presumption in favor of organ donation is better, if only because more people are sufficiently satisfied to leave it in place.

IV How Much Choice Should Be Offered?

It is far beyond our ambition here to venture a full analysis of the question of how much choice to offer individuals in various domains (Loewenstein, 2000, pp. 89–94; Dworkin, 1988, pp. 62–81). Instead, we identify some questions that a libertarian paternalist might ask to help decide how much (reasonable) choice to offer. Any such libertarian would obviously want to reduce the frequency and severity of errors, and the costs of making decisions. If an approach increases the costs of decisions for choosers, there is less reason to adopt it, and it should be selected only if it is likely to improve the match of choices to actual welfare. If an approach increases errors and their costs by leading people to make choices that do not promote their welfare, that is a strong point against it. We now trace some considerations that help answer the question whether more choices would increase the costs of errors and the costs of decisions.

Preferences, Paternalism, and Liberty

A. Do Choosers Have Informed Preferences?

In some domains, consumers and workers are highly informed—so much so that they will not even be influenced by default rules. Most adults have experimented enough over the course of their lives to have a good sense of what flavors of ice cream they like. They can do a decent job of picking even in a shop offering dozens of flavors. If the default option is asparagus-flavored ice cream, they will be unlikely to choose it, and might well be annoyed. But when faced with a menu listing many unfamiliar foods in a foreign country, customers would be unlikely to benefit from being required to choose among them, and they might prefer a small list or ask the waiter for a default suggestion (for example, what do other tourists like?). In such settings, clever restaurants catering to tourists often offer a default ‘tourist menu.’ Many actual choices fall between the poles of ice cream flavors and foreign menus. When information is limited, a menu of countless options increases the costs of decisions without increasing the likelihood of accuracy. But when choosers are highly informed, the availability of numerous options decreases the likelihood of error and does not greatly increase decision costs, simply because informed choosers can more easily navigate the menu of options.

B. Is the Mapping from Options to Preferences Transparent?

If we order a coffee ice cream cone, we have a pretty good idea what we will consume. If we invest \$10,000 in a mix of mutual funds, we have little idea (without the aid of sophisticated software) what a change in the portfolio will do to our distribution of expected returns in retirement. When we choose between health plans, we may not fully understand all the ramifications of our choice. If I get a rare disease, will I be able to see a good specialist? How long will I have to wait in line? When people have a hard time predicting how their choices will end up affecting their lives, they have less to gain from having numerous options from which to choose. If it is hard to map from options to preferences, a large set of choices is likely to be cognitively overwhelming, and thus to increase the costs of decisions without also increasing welfare by reducing errors.

Richard H. Thaler

C. How Much Do Preferences Vary across Individuals?

Some people smoke; others hate the smell of smoke. Some people like hard mattresses; others like soft ones. How do hotels deal with this problem? Most choose to cater to differences in tastes with respect to smoking but not with respect to mattresses. The mattress that appeals to the median hotel guest seems to be good enough to satisfy most customers, but the threat of a smoky room (or a night without cigarettes) is enough to scare customers away. Here is a case in which many people have well-formed preferences that trump default rules. Many planners, both private and public, must make similar tradeoffs. Since offering choice is costly, sensible planners make multiple choices available when people's preferences vary most. The argument for a large option set is thus strongest in cases of preferences that are both clear and heterogeneous. In such cases, people's welfare is likely to be promoted if each can choose as he sees fit, and homogeneity will lead to inaccuracy and thus widespread error costs.

D. Do Consumers Value Choosing for Themselves As An Intrinsic Good?

Freedom of choice is itself an ingredient in welfare. In some situations people derive welfare from the very act of choosing. But sometimes it is a chore to have to choose, and the relevant taste can differ across individuals. (One of us derives pleasure from reading and choosing from a wine list; the other finds that enterprise basically intolerable.) A more serious example comes from evidence that many patients do not want to make complex medical decisions and would prefer their doctors to choose for them (Schneider, 1998, pp. 35–46). The point very much bears on the decision whether to force explicit choices or instead to adopt a default rule that reflects what the majority wants. If making choices is itself a subjective good, the argument for forced choices is strengthened. 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. By contrast, a thoughtfully chosen default rule, steering them in sensible directions, is a blessing.

Preferences, Paternalism, and Liberty

Conclusion

Our central empirical claim here has been that in many domains, people's preferences are labile and ill-formed, and do not predate social and legal contexts. For this reason, starting points and default rules are likely to be quite sticky. Building on empirical work involving rationality and preference formation, we have sketched and defended libertarian paternalism – an approach that preserves freedom of choice but that encourages both private and public institutions to steer people in directions that will promote their own welfare.

Some kind of paternalism, we believe, is likely whenever such institutions set out default plans or options. Unfortunately, many current social outcomes are both random and inadvertent, in the sense that they are a product of default rules whose behavior-shaping effects have never been a product of serious reflection. In these circumstances, the goal should be to avoid arbitrary or harmful consequences and to produce contexts that are likely to promote people's welfare, suitably defined.

References

- Ayres, I., & Gertner, R. (1989). Filling gaps in incomplete contracts: An economic theory of default rules. *Yale Law Journal*, *99*, 87–130.
- Bateman, I. A., & Willis, K. G. (1999). *Valuing environmental preferences*. New York: Oxford University Press.
- Benartzi, S., & Thaler, R. H. (1999). Risk aversion or myopia? Choices in repeated gambles and retirement investments. *Management Science*, *45*, 364–381.
- Benartzi, S., & Thaler, R. H. (2002). How much is investor autonomy worth? *Journal of Finance*, *57*, 1593–1616.
- Calabresi, G., & Melamed, A. D. (1972). Property rules, liability rules, and inalienability: One view of the cathedral. *Harvard Law Review*, *85*, 1089–1128.
- Calle, E. E., Thun, M. J., Petrelli, J. M., Rodriguez, C., & Heath, C. W., Jr. (1999). Body-mass index and mortality in a prospective cohort of U.S. adults. *New England Journal of Medicine*, *341*, 1097–1105.
- Camerer, C. F. (2000). Prospect theory in the wild: Evidence from the field. In D. Kahneman & A. Tversky (Eds.), *Choices, values, and frames*. Cambridge, UK: Cambridge University Press.
- Camerer, C. F., & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty*, *19*, 1–3, 7–42.

Richard H. Thaler

- Centers for Disease Control and Prevention. (2003). *1991–2001 Prevalence of obesity among U.S. adults, by characteristics*: Retrieved May 6, 2004, from http://www.cdc.gov/nccdphp/dnpa/obesity/trend/prev_char.htm.
- Choi, J. J., Laibson, D., Madrian, B. C., & Metrick, A. (2002). Defined contribution pensions: Plan rules, participant decisions, and the path of least resistance. In J. M. Poterba (Ed.), *Tax policy and the economy* (Vol. 16, pp. 67–113). Cambridge, MA: MIT Press.
- Cropper, M. L., Aydede, S. K., & Portney, P. R. (1994). Preferences for life-saving programs: How the public discounts time and age. *Journal of Risk and Uncertainty*, 8(3), 243–265.
- De Bondt, W. F. M., & Thaler, R. H. (1990). Do security analysts overreact? *American Economic Review*, 80(2), 52–57.
- Dworkin, G. (1988). *The theory and practice of autonomy*. Cambridge, UK: Cambridge University Press.
- Frederick, S. (2003). Measuring intergenerational time preference: Are future lives valued less? *Journal of Risk and Uncertainty*, 26, 39–53.
- Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, 40, 351–401.
- Friedman, M., & Friedman, R. (1980). *Free to choose: A personal statement*. New York: Harcourt Brace Jovanovich.
- Grether, D. M. (1980). Bayes rules as a descriptive model: The representativeness heuristic. *Quarterly Journal of Economics*, 95, 537–557.
- Johnson, E. J., Hershey, J., Meszaros, J., & Kunreuther, H. (1993). Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty*, 7, 35–51.
- Jolls, C., Sunstein, C. R., & Thaler, R. H. (1998). A behavioral approach to law and economics. *Stanford Law Review*, 50, 1471–1550.
- Jones-Lee, M., & Loomes, G. (2001). Private values and public policy. In E. U. Weber, J. Baron, & G. Loomes (Eds.), *Conflict and tradeoffs in decision making* (pp. 205–230). Cambridge, UK: Cambridge University Press.
- Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics of intuitive judgment: Extensions and applications* (pp. 49–81). New York: Cambridge University Press.
- Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991, Winter). Anomalies: The endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives*, 5, 193–206.
- Korobkin, R. (1998). The status quo bias and contract default rules. *Cornell Law Review*, 83, 608–687.
- Kunreuther, H. (1996). Mitigating disaster losses through insurance. *Journal of Risk and Uncertainty*, 12(2–3), 171–187.
- Loewenstein, G. F. (2000). How people make decisions: Costs and benefits of health- and retirement-related choice. In S. Burke, E. Kingson, & U.

Preferences, Paternalism, and Liberty

- Reinhardt (Eds.), *Social Security and Medicare: Individual versus collective risk and responsibility* (pp. 87–113). Washington, DC: National Academy of Social Insurance.
- Madrian, B. C., & Shea, D. F. (2001). The power of suggestion: Inertia in 401(k) participation and savings behavior. *Quarterly Journal of Economics*, *116*, 1149–1187.
- Mill, J. S. (1972). On liberty. In H. B. Acton (Ed.), *Utilitarianism, on liberty, considerations on representative government*. London: Dent.
- Morrison, E. R. (1998). Comment, judicial review of discount rates used in regulatory cost-benefit analysis. *University of Chicago Law Review*, *65*, 1333–1359.
- National Institute of Diabetes and Digestive and Kidney Diseases. (2001). *Understanding adult obesity*: Retrieved May 4, 2005, from <http://win.niddk.nih.gov/publications/understanding.htm>.
- Okin, S. M. (1989). *Justice, gender, and the family*. New York: Basic.
- Payne, J. W., Bettman, J. R., & Schkade, D. A. (1999). Measuring constructed preferences: Towards a building code. *Journal of Risk and Uncertainty*, *19*, 243–270.
- Qualified Transportation Fringe Benefits, 26 CFR Parts 1 and 602 (2001).
- Redelmeier, D., Rozin, P., & Kahneman, D. (1993). Understanding patients' decisions: Cognitive and emotional perspectives. *Journal of the American Medical Association*, *270*, 72–76.
- Revesz, R. L. (1999). Environmental regulation, cost-benefit analysis, and the discounting of human lives. *Columbia Law Review*, *99*, 941–1017.
- Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision-making. *Journal of Risk and Uncertainty*, *1*, 7–59.
- Schneider, C. E. (1998). *The practice of autonomy: Patients, doctors, and medical decisions*. New York: Oxford University Press.
- Shiller, R. J. (2000). *Irrational exuberance*. Princeton, NJ: Princeton University Press.
- Slovic, P., Kunreuther, H., & White, G. F. (1974). Decision processes, rationality and adjustment to natural hazards. In G. F. White (Ed.), *Natural hazards: Local, national, global* (pp. 187–205). New York: Oxford University Press.
- Sunstein, C. R. (2001). Human behavior and the law of work. *Virginia Law Review*, *87*, 205–276.
- Sunstein, C. R. (2002a). *Risk and reason: Safety, law, and the environment*. Cambridge, UK: Cambridge University Press.
- Sunstein, C. R. (2002b). Switching the default rule. *New York University Law Review*, *77*(106–134).
- Sunstein, C. R. (2004). Lives, life-years, and willingness to pay. *Columbia Law Review*, *104*, 205–252.
- Sunstein, C. R., Kahneman, D., Schkade, D., & Ritov, I. (2002). Predictably incoherent judgments. *Stanford Law Review*, *54*, 1153–1215.
- Thaler, R. H. (1991). *Quasi-rational economics*. New York: Russell Sage.

Richard H. Thaler

- Thaler, R. H., & Benartzi, S. (2004). Save More Tomorrow™: Using behavioral economics to increase employee saving. *Journal of Political Economy*, 112, S164-S187.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207-232.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124-1131.
- Tversky, A., & Thaler, R. H. (1990). Anomalies: Preference reversals. *Journal of Economic Perspectives*, 4(2), 201-211.
- VanDeVeer, D. (1986). *Paternalistic intervention: The moral bounds on benevolence*. Princeton, NJ: Princeton