

Libertarian Paternalism

By RICHARD H. THALER AND CASS R. SUNSTEIN*

Many economists are libertarians and consider the term “paternalistic” to be derogatory. Most would think that the phrase libertarian paternalism is an oxymoron. The modest goal of this essay is to encourage economists to rethink their views on paternalism. We believe that the anti-paternalistic fervor expressed by many economists is based on a combination of a false assumption and at least two misconceptions. The false assumption is that people always (usually?) make choices that are in their best interest. This claim is either tautological, and therefore uninteresting, or testable. We claim that it is testable and false—indeed, obviously false.

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 choices of some other people. The point applies to both private and public actors. Consider the problem facing the director of a company cafeteria who discovers that the order in which food is arranged influences the choices people make. To simplify, consider three alternative strategies: (1) she could make choices that she thinks would make the customers best off; (2) she could make choices at random; or (3) she could maliciously choose those items that she thinks would make the customers as obese as possible. Option 1

appears to be paternalistic, which it is, but would anyone advocate options 2 or 3?

The second misconception is that paternalism always involves coercion. As the cafeteria example illustrates, the choice of which order to present food items does not coerce anyone to do anything, yet one might prefer some orders to others on paternalistic grounds. Would many object to putting the fruit before the desserts at an elementary school cafeteria if the outcome were to increase the consumption ratio of apples to Twinkies? Is this question fundamentally different if the customers are adults? If no coercion is involved, we think that some types of paternalism should be acceptable to even the most ardent libertarian. We call such actions *libertarian paternalism*.

In our understanding, a policy counts as “paternalistic” if it is selected with the goal of influencing the choices of affected parties in a way that will make those parties better off. We intend “better off” to be measured as objectively as possible, and we clearly do not always equate revealed preference with welfare. That is, we emphasize the possibility that in some cases individuals make inferior choices, choices that they would change if they had complete information, unlimited cognitive abilities, and no lack of willpower. 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 sick, obese, or generally 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 paternalistic options.¹

[†] *Discussants:* Robert E. Hall, Stanford University; Gary Becker, University of Chicago; Robert Barro, Harvard University.

* Thaler: Department of Economics and Behavioral Science, University of Chicago Graduate School of Business, 1101 E. 58th Street, Chicago, IL 60637 (e-mail: thaler@gsb.uchicago.edu); Sunstein: Law School and Department of Political Science, University of Chicago, 1111 E. 60th Street, Chicago, IL 60637 (e-mail: csunstei@midway.uchicago.edu). We are grateful to Richard Epstein, Russ Fuller, Owen Lamont, Eric Posner, Richard Posner, and David Wilcox for valuable comments on a previous draft.

¹ Readers interested in this topic should also consult Colin Camerer et al. (2001) for an illuminating discussion of related issues. That paper shares with the papers in this

I. Are Choices Rational?

The presumption that individual choices should be free from interference is usually based on the assumption that people do a good job of making choices, or at least that they do a far better job than third parties could do. As far as we can tell, there is little empirical support for this claim. Research by psychologists and economists over the past three decades has raised questions about the rationality of the judgments and decisions that individuals make. People do not exhibit rational expectations, fail to make forecasts that are consistent with Bayes' rule, use heuristics that lead them to make systematic blunders, exhibit preference reversals (that is, they prefer A to B *and* B to A) and make different choices depending on the wording of the problem (for many examples, see the two recent collections of papers by Daniel Kahneman and Amos Tversky [2000] and by Thomas Gilovich et al. [2002]). Furthermore, in the context of intertemporal choice, people exhibit dynamic inconsistency, valuing present consumption much more than future consumption. In other words, people have self-control problems (see the other papers in this session [James Choi et al., 2003b; Ted O'Donoghue and Matthew Rabin, 2003] for details and references).

Many economists are skeptical of some of these findings, thinking that people may do a better job of choosing in the "real world" than they do in the laboratory. However, studies of actual choices for high stakes reveal many of the same problems. For example, the Surgeon General reports that 61 percent of Americans are either overweight or obese. Given the adverse effects obesity has on health, it is hard to claim that Americans are eating optimal diets.

Another illustration comes from the domain of savings behavior. Shlomo Benartzi and Thaler (2002) investigate 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 just 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 they 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 choosing investment portfolios for themselves.

II. Is Paternalism Inevitable?

As the cafeteria line example discussed above illustrates, planners are forced to make some design choices. A simple and important example is the selection of a "default option" to determine what happens if an agent fails to choose for himself. In a fully rational world such design choices would have little effect (at least in high-stakes situations) because agents would simply choose the best option for them regardless of the default. However, numerous experiments illustrate that there is a very strong "status quo" bias (see William Samuelson and Richard Zeckhauser, 1988; Kahneman et al., 1991). The existing arrangement, whether set out by private institutions or by government, tends to stick.

One illustration of this phenomenon comes from studies of automatic enrollment in 401(k) employee savings plans. 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 some default options for savings rates and asset allocation). In companies that offer a "match" (the employer matches the employee's contributions accord-

session the common goal of devising policies that help some agents who are making some mistake, while minimizing the costs imposed on others.

ing 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, Brigitte Madrian and Dennis Shea (2001) found that initial enrollments jumped from 49 percent to 86 percent, and Choi et al. (2002) find similar results for other companies.

Should the adoption of automatic enrollment be considered paternalistic? And, if so, should it therefore be seen as a kind of officious meddling with employee preferences? We answer these questions yes and no respectively. If the employer thinks (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. They are attempting to steer employees' choices in directions that will promote employees' welfare. But since no one is forced to do anything, we think this steering should be considered unobjectionable to 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, zero 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.

Quick-minded readers 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 a choice, either in or out. But some thought reveals that this is not at all a way out of the dilemma; rather, it is simply another option among many that the employer can elect. In fact, Choi et al. (2003a) find that this rule increases enrollments (relative to the opt-in rule) though by not as much as automatic enrollment. Furthermore, the very requirement that employees make a choice has a paternalistic element. Many employees do not want to have to make a choice (and would choose not to have to do so). Should employers really force them to choose?

Why, exactly, does the setting of defaults have such large effects? 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. But there is a separate explanation, involving inertia. For any employee, a change from any status quo entails time and effort, and many people seem to prefer to avoid both of these, especially if they are prone to procrastination. When default rules are "sticky" and affect choices as a result, inertia might be the major reason.

For present purposes, the choice among these various explanations does not much matter. The point is only that paternalism, in the form of effects on individual choices, is often unavoidable. When paternalism seems absent, it is usually because the starting point appears so natural and obvious that its preference-shaping effects are invisible to most observers. But those effects are nonetheless there. 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-paternalistic position is incoherent.

III. Beyond the Inevitable (but Still Libertarian)

The inevitability of paternalism is most clear when the planner has to choose default rules. It is reasonable to ask whether the planner should go beyond the inevitable. Take the cafeteria example discussed above. Putting the fruit before the desserts is a fairly mild intervention. A more intrusive step would be to place the desserts in another location altogether, so that diners have to get up and get a dessert after they have finished the rest of their meal. This step raises the transactions costs of eating dessert, and according to a standard economic analysis the proposal is unattractive: it seems to make dessert-eaters worse off and no one better off. But once self-control costs are incorporated, we can see that some diners would prefer this arrangement, namely, those who would eat a dessert if it were put in front of them but would resist temptation if given a little help. To the extent that the dessert location is not hard to find, and no choice is forbidden, this approach meets libertarian muster.

In the domain of employee savings, Thaler and Benartzi (2003) have proposed a method of increasing contributions to 401(k) plans that also meets the libertarian test. Under this plan,

called Save More Tomorrow, 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 plan. 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. This is successful libertarian paternalism in action.

IV. How to Choose: The Toolbox of the Libertarian Paternalist

How should sensible planners (a category we mean to include anyone who must design plans for others, from human-resource directors to bureaucrats to kings) choose among possible systems, given that some choice is necessary? We suggest two approaches to this problem.

If feasible, a comparison of possible rules should be done using a form of cost–benefit analysis. The goal of a cost–benefit study would be to measure the full ramifications of any design choice. To illustrate, take the example of automatic enrollment. Under automatic enrollment some employees will join the plan who otherwise would not. Presumably, some are made better off (especially if there is an employer match), but some may be made worse off (e.g., those who are highly liquidity-constrained). If the issue were just enrollment, we would guess that the gains would exceed the losses. We base this guess partly on revealed choices. Most employees do join the plan eventually, and very few who are automatically enrolled opt out when they figure out what has happened to them. We also judge that the costs of having too little saved up for retirement are typically greater than the costs of having saved too much.

In many cases, however, the planner will be unable to make a direct inquiry into welfare, either because too little information is available or because the costs of doing 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 approach is welfare-promoting without relying on unreliable 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*. Useful though it is, this market-mimicking approach raises its own problems. Perhaps the majority's choices would be insufficiently informed. Perhaps those choices, in fact, would not 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 is likely to go one way or the other regardless of the starting point; and to that extent, the market-mimicking strategy seems quite workable.

Second, the libertarian paternalist might select the approach *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 forced choices in the context of retirement plans; it would be easy to find other examples. 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. If it is likely that automatic enrollment is welfare-promoting, perhaps automatic enrollment should be preferred over forced choices. The only suggestion is that, where the social planner is unsure how to handle the welfare question, he might devise a strategy that requires people to choose.

Third, the libertarian paternalist might select the approach *that minimizes the number of opt-*

outs. For example, very few employees opt out of the 401(k) plan when they are automatically enrolled, though many opt in under the standard enrollment procedure. 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 automatic enrollment is better, if only because more people are sufficiently satisfied to leave it in place.

V. Conclusion

Our goal here has been to defend libertarian paternalism, an approach that preserves freedom of choice but that authorizes both private and public institutions to steer people in directions that will promote their welfare. Some kind of paternalism is likely whenever such institutions set out arrangements that will prevail unless people affirmatively choose otherwise. In these circumstances, the goal should be to avoid random, arbitrary, or harmful effects and to produce a situation that is likely to promote people's welfare, suitably defined.

REFERENCES

- Benartzi, Shlomo and Thaler, Richard H.** "How Much Is Investor Autonomy Worth?" *Journal of Finance*, August 2002, 57(4), pp. 1593–1616.
- Camerer, Colin; Issacharoff, Samuel; Loewenstein, George; O'Donoghue, Ted and Rabin, Matthew.** "Regulation for Conservatives: Behavioral Economics and the Case for Asymmetric Paternalism." Working paper, California Institute of Technology, 2001.
- Choi, James; Laibson, David; Madrian, Brigitte C. and Metrick, Andrew.** "Defined Contribution Pensions: Plan Rules, Participant Decisions, and the Path of Least Resistance," in James M. Poterba, ed., *Tax policy and the economy*, Vol. 16. Cambridge, MA: MIT Press, 2002, pp. 67–113.
- _____. "Benign Paternalism and Active Decisions: A Natural Experiment." Working paper, University of Chicago, Graduate School of Business, 2003a.
- _____. "Optimal Defaults." *American Economic Review*, May 2003b (*Papers and Proceedings*), 93(2), pp. 180–85.
- Gilovich, Thomas; Griffin, Dale and Kahneman, Daniel, eds.** *Heuristics and biases: The psychology of intuitive judgment*. Cambridge, U.K.: Cambridge University Press, 2002.
- Kahneman, Daniel; Knetsch, Jack L. and Thaler, Richard H.** "The Endowment Effect, Loss Aversion, and Status Quo Bias." *Journal of Economic Perspectives*, Winter 1991, 5(1), pp. 193–206.
- Kahneman, Daniel and Tversky, Amos, eds.** *Choices, values, and frames*. Cambridge, U.K.: Cambridge University Press, 2000.
- Madrian, Brigitte and Shea, Dennis.** "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior." *Quarterly Journal of Economics*, November 2001, 116(4), pp. 1149–87.
- O'Donoghue, Ted and Rabin, Matthew.** "Studying Optimal Paternalism, Illustrated by a Model of Sin Taxes." *American Economic Review*, May 2003 (*Papers and Proceedings*), 93(2), pp. 186–91.
- Samuelson, William and Zeckhauser, Richard J.** "Status Quo Bias in Decision Making." *Journal of Risk and Uncertainty*, March 1988, 1(1), pp. 7–59.
- Thaler, Richard H. and Benartzi, Shlomo.** "Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving." *Journal of Political Economy*, 2003 (forthcoming).