

CHAPTER II

Libertarian paternalism, utilitarianism, and justice

Jamie Kelly

INTRODUCTION

In a number of recent publications, Cass Sunstein and Richard Thaler have argued for a novel approach to the design of public policy.¹ Their proposal has received a great deal of attention, both within academic circles and the public at large. Drawing upon evidence from behavioral economics and empirical psychology, the authors attempt to demonstrate that the conventional antagonism between libertarians and paternalists within political theory dissolves in conditions that obtain widely in public decision-making. Where free choice and the promotion of individual welfare can coexist, the authors believe that designers of public policy ought to be libertarian paternalists.

In this paper I criticize their proposal on grounds that the authors are unable to sufficiently motivate the paternalistic element of their approach. I argue that the empirical evidence cited by the authors is capable of supporting a number of competing approaches, including what I call libertarian utilitarianism and libertarian justice. Since the evidence that the authors draw upon does not provide us any grounds for selecting between these rival approaches, I conclude that Sunstein and Thaler are unable to provide us with a convincing guide for the design of public policy. In order to show that this is the case, I consider three arguments in favor of libertarian paternalism, and find each lacking. I end with some comments about what we can properly conclude on the basis of Sunstein and Thaler's arguments.

In constructing their argument for libertarian paternalism, Sunstein and Thaler draw upon a rich and interesting body of empirical results from the Heuristics and Biases literature.² This literature got its start in empirical

¹ I will focus on Thaler and Sunstein, *Nudge*, "Libertarian Paternalism," and Sunstein and Thaler, "Libertarian Paternalism Is Not an Oxymoron."

² Major contributions to this research have been anthologized in three volumes: Kahneman, Slovic, and Tversky, eds., *Judgment Under Uncertainty*; Kahneman and Tversky, eds., *Choices, Values, and Frames*; and Gilovich, Griffin, and Kahneman, eds., *Heuristics and Biases*.

psychology, but has recently had enormous impact upon a number of other disciplines including economics, law, and finance. Most important for our purpose is the portion of this literature that concerns the relationship between individual choice and the framing of decisions.

Since the 1970s, a vast amount of empirical evidence has been accumulated showing that individual choices in a wide variety of domains are not invariant over equivalent presentations of a decision problem. That is, individuals respond in different ways to a given choice, depending upon how it is framed. Sunstein and Thaler interpret this evidence as follows:

Our 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.³

Drawing on one of the most memorable experiments in the Heuristics and Biases literature,⁴ Sunstein and Thaler explain:

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.” 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.⁵

The authors assert that our susceptibility to framing effects means that, in a whole host of common situations, we lack stable preferences. In proposing libertarian paternalism, they aim to show that such situations often involve contexts of both public and private decision-making. As a result, they argue that planners ought to design choices in such a way that individuals are “nudged” towards making good decisions, even if they might not do so on their own. One of their central examples concerns Carolyn, the director of food services for a large city school system. In their example, Carolyn must make a decision about how to arrange items in the cafeterias under her control:

³ Sunstein and Thaler, “Libertarian Paternalism Is Not an Oxymoron,” 1161.

⁴ McNeil et al., “On the Elicitation of Preferences for Alternative Therapies.”

⁵ Sunstein and Thaler, “Libertarian Paternalism Is Not an Oxymoron,” 1161.

Here are some suggestions she has received from her usually sincere but occasionally mischievous friends and coworkers:

1. Arrange the food to make students best off, all things considered.
2. Choose the food order at random.
3. Try to arrange the food to get the kids to pick the same foods they would choose on their own.
4. Maximize the sales of the items from suppliers that are willing to offer the largest bribes.
5. Maximize profits, period.⁶

Sunstein and Thaler claim that the cafeteria director described above cannot help but make a choice that affects the decisions of her customers. Food must be arranged in some way, and the organizational strategy that is adopted will help to determine what people eat. The reason that this is the case stems from the (in this case, unproven) assertion that food ordering influences food choice. That is, we are invited to imagine that in this situation individuals lack clear, stable, or well-defined preferences about what to eat. Given this lack, an opportunity presents itself: The cafeteria director has the chance to promote better nutrition without having to eliminate any of her customers' options. For example, should she choose strategy 1, more people might be induced to choose healthy options, but those with a well-defined preference for a cheeseburger and fries will not have had their freedom of choice restricted in any way. This is the sense in which their proposal is "libertarian": Their aim is to influence decisions without constraining liberty. Thus, no options are to be eliminated, and no significant costs are to be imposed on the selection of these options.

Central to Sunstein and Thaler's proposal is their claim that evidence from the social sciences justifies a distinctively paternalistic approach to public policy. I will argue that this claim is false, that this research cannot justify paternalism in particular. In what follows, I will consider three different arguments purporting to justify the use of a paternalist approach to public policy. The first argument comes from a series of publications by Sunstein and Thaler written prior to *Nudge*. Here the authors assert directly that paternalism in public policy is inevitable. The second reconstructs the argument of *Nudge*, where the authors merely claim that nudges are inevitable. Finally, the third provides a friendly argument of my own, intended to establish that nudges ought to be paternalistic. I show that none of these arguments succeed in establishing libertarian paternalism as the proper approach to the design of public policy.

⁶ Thaler and Sunstein, *Nudge*, 2.

PATERNALISM IS INEVITABLE

Starting in 2003, Sunstein and Thaler published a series of articles advocating for a new way of thinking about paternalism. Libertarian paternalism was proposed as a response to standard anti-paternalist arguments concerning government ineptitude and intrusion. Relying upon insights from the social sciences, the authors sought to demonstrate that their form of paternalism was not subject to standard anti-paternalist objections. In particular, in these early publications,⁷ the authors claimed that the instability of our preferences made it the case that paternalism in public policy is inevitable:

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.⁸

Unfortunately, the authors appear to be running together two claims here: The first is the claim that some sort of influence is inevitable; the second is that the appropriate response must be paternalistic. In the present context, “influence” merely entails that the decisions of public planners have an impact upon the decisions of citizens or consumers. This kind of influence is demonstrated by a large number of studies where the framing of a decision directly influences the rate at which different options are chosen.⁹ It is clear to me that existing empirical evidence is sufficient to warrant the conclusion that this kind of influence in public policy is indeed inevitable: It is simply not possible to avoid having an impact upon the expressed preferences of individuals when we go about designing public policy instruments. As a result, I think we should accept this premise.

However, paternalism is not coextensive with influence: There are a great many sorts of influence that do not count as paternalistic. If, in the cafeteria example above, the director were to select strategy 4 (i.e., maximize bribes), she would hardly be attempting to make choosers better off. As such, it should be clear that the inevitability of influence does not entail that policy

⁷ See especially Thaler and Sunstein, “Libertarian Paternalism,” 176–177, and Sunstein and Thaler, “Libertarian Paternalism Is Not an Oxymoron,” 1171–1190.

⁸ Thaler and Sunstein, “Libertarian Paternalism,” 175.

⁹ For a statistically striking example of this impact, see the famous Asian Disease Problem from “Choices, Values, and Frames” in Kahneman and Tversky, eds., *Choices, Values, and Frames*.

designers must necessarily act paternalistically. For their part, the authors define the paternalistic element of their proposal in the following way:

The paternalistic aspect consists in the claim that it is legitimate for private and public institutions to attempt to influence people's behavior even when third-party effects are absent. In other words, we argue for self-conscious efforts, by 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.¹⁰

What, then, is going on? Even on the authors' own understanding of paternalism, it cannot be strictly true that paternalism is inevitable. In order to understand Sunstein and Thaler's thesis, I think we need to weaken our understanding of "inevitability." What the authors seem to be claiming is that once we eliminate all the strategies that are morally intolerable, some form of paternalism is the only viable option. That is, assuming that policy designers are even moderately well-intentioned, they cannot but engage in paternalism. Even on this weaker reading, however, Sunstein and Thaler's claim is false.

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.¹¹

Even if we assume that policy planners are well-intentioned, however, the instability of preferences should not lead us to conclude that paternalism in public policy is – in any sense – inevitable. According to Sunstein and Thaler, a policy is paternalistic if it attempts to make choosers better off. It is simply not the case, however, that a planner's good intentions will lead them to choose policies that promote the welfare of choosers. To see why this is the case, let us examine the predicament of a policy planner from a very abstract level.

Imagine a choice architect who has to present a choice to the public. There are a number of different potential frames for the choice, and she knows that her decision about which frame to use will have an influence

¹⁰ Sunstein and Thaler, "Libertarian Paternalism Is Not an Oxymoron," 1162.

¹¹ Thaler and Sunstein, "Libertarian Paternalism," 177.

upon some people's choices. As a result, she must rank the potential frames in accordance with some standard, using the predicted influence of each frame as a guide. Once she has determined which frame fares best against her standard, she should employ that frame for the public choice.

Notice, however, that the above description tells us nothing about which standard our choice architect ought to use to rank the frames for the decision in question. Sunstein and Thaler assume that the appropriate standard to use when evaluating frames is paternalistic: Frames that promote the welfare of choosers are good; frames that do not are bad (similarly, frames that promote the welfare of choosers more are better than frames that promote the welfare of choosers less). The data on framing effects is, however, unable to guide our selection of such a standard. This is because the scientific evidence is entirely descriptive: It tells us that we are often responsive to frames, and it tells us which sorts of frames tend to have greater or lesser purchase on our decisions. On its own, however, this data generates no prescriptions for public policy. To arrive at a guide for the design of public decisions, we must import a moral standard from elsewhere. It is only by relying upon extraneous moral considerations that we can decide how we should rank frames. Since empirical evidence is silent on this issue, the research cited by Sunstein and Thaler can be used to support a number of rival guides for the design of public policy.

NUDGES ARE INEVITABLE

In *Nudge: Improving Decisions About Health, Wealth, and Happiness* we find the authors once again arguing for an approach to public policy that is both libertarian and paternalist. Whereas their prior arguments seemed to have been intended for a largely academic audience, in *Nudge* they propose a version of the same basic argument designed for popular consumption. Once again, they endorse the design of public policies that “maintain or increase freedom of choice” (this is the libertarian aspect of their proposal) while promoting paternalist goals:

The paternalist aspect lies in the claim that it is legitimate for choice architects to try to influence people's behavior in order to make their lives longer, healthier, and better. In other words, we argue for self-conscious efforts, by institutions in the private sector and also by government, to steer people's choices in directions that will improve their lives. In our understanding, a policy is “paternalistic” if it tries to influence choices in a way that will make choosers better off, *as judged by themselves*.¹²

¹² Thaler and Sunstein, *Nudge*, 5, emphasis in the original.

In *Nudge*, however, Sunstein and Thaler abandon the explicit claim that paternalism in public policy is inevitable. Given my concerns in the last section, this ought to be a good thing, but their new argument seems to rely upon a similar ambiguity in their understanding of “nudges.” The authors claim that:

A nudge, as we will use the term, is any aspect of the choice architecture that alters people’s behavior in a predictable way without forbidding any options or significantly changing their economic incentives. To count as a mere nudge, the intervention must be easy and cheap to avoid. Nudges are not mandates. Putting the fruit at eye level counts as a nudge. Banning junk food does not.¹³

Once again, I have no quarrel with this description of nudges, and I find persuasive the authors’ claim that nudges in public policy are inevitable. Our susceptibility to defaults, starting points, and the framing of decisions shows that choice architects cannot avoid nudging us in certain predictable directions. Sunstein and Thaler state: “In many cases, some kind of nudge is inevitable, and so it is pointless to ask government simply to stand aside. Choice architects, whether private or public, must do *something*.”¹⁴

Unfortunately, the claim that nudges are inevitable can provide little support for libertarian paternalism. We can agree that choice architects necessarily influence the decisions of the public without agreeing about what kind of standard ought to govern the design of public policy. Unless we interpret Sunstein and Thaler to be making the purely stipulative claim that nudges – in order to be nudges – need to aim at the benefit of choosers, then there is no reason to think that all nudges must be paternalist. Sunstein and Thaler don’t provide any argument for the view that all or even most nudges will aim at paternalist goals. Instead, they rely upon an unstated assumption that nudges must be paternalistic. This assumption, however, is false. In what follows, I present two examples of how we could go about designing non-paternalist nudges. These nudges seek to preserve liberty while attempting to steer behavior towards morally appealing but non-paternalistic choices.

A utilitarian nudge

Consider the following modification to Sunstein and Thaler’s cafeteria example: Our fictional cafeteria director has to choose which item to spotlight at the front of the line. She knows that placing an item in this location is likely

¹³ *Ibid.*, 6. ¹⁴ *Ibid.*, 237.

to result in a large number of customers (those with unstable preferences) choosing this item. Her two options for the spotlight today are as follows:

- (a) Californian Mixed Greens with Poached Chilean Sea Bass;
- (b) Farm-fresh egg salad on homemade bread.

Further, let us assume that the cafeteria director has the following beliefs:

- (i) A diet low in saturated fat is good for her customers' health;
- (ii) Transporting food from far away is environmentally unsustainable.

On the above assumptions,¹⁵ I think that it is clear that if the cafeteria director has her customers' welfare in mind, she will choose to place the greens and sea bass at the beginning of the line (since this option helps to improve the health of her customers). However, if instead she has the best interest of people more generally in mind, then she should choose the egg salad (since this yields greater overall utility). If she chooses (a), she is acting paternalistically, whereas if she chooses (b), she is being a utilitarian. In neither case, however, is she making her decision on morally unreasonable grounds. If she chooses (b) she is still well-intentioned, but she has simply chosen a different, utilitarian, standard against which to evaluate potential frames for her customers' decisions.

My concern here isn't to establish how the cafeteria director ought to structure her menu, but rather to point out that there is nothing in the empirical literature that ought to incline her towards one option rather than the other. Just as Sunstein and Thaler have shown that libertarian concerns are – when preferences are unstable – reconcilable with paternalism, we can see that the same is equally true of utilitarianism. Sunstein and Thaler's proposal ought to have no traction for the cafeteria director, unless she has independently decided that her nudges ought to be paternalist in nature. If she instead believes that she should nudge people towards decisions that promote overall utility, then there is nothing in Sunstein and Thaler's proposal that should change her mind.

A Rawlsian nudge

We can also use empirical evidence of framing effects to construct nudges aimed at promoting Rawls' theory of justice (or any other theory of justice, for that matter). Advocates of this sort of approach would claim that in

¹⁵ Let us also assume: (1) That this particular cafeteria is in Ohio, and (2) that, at least in this instance, the utility of promoting environmentally sustainable agriculture outweighs that of promoting the health of her customers.

instances where individuals do not have stable preferences and where the choice of frame does not curtail anyone's options, those with the opportunity ought to select the frame that serves to promote policies that best comply with Rawls' two principles of justice.¹⁶

As an example, let us consider a situation akin to the sorts of policy decisions to which Sunstein and Thaler wish to see libertarian paternalism applied: the situation of a legislator drafting a referendum on property taxes. Imagine that our legislator knows that people are affected by framing, and that she also knows that individuals are hostile to frames that represent a move away from the status quo.¹⁷ She must decide between two equivalent formulations of a proposal to raise taxes:

- (a) Raise taxes by 2 percent over last year;
- (b) Maintain the same rate of increase (2 percent) as in previous years.

Since (a) frames the proposal as a deviation from the status quo while (b) frames it as a continuation of the status quo, it is likely that the second proposal will garner more support than the first. Given people's strong bias in favor of the status quo, her choice could significantly influence the referendum results.

Now, let's further assume that our legislator believes that it is in the interest of voters to have lower property taxes. If she follows Sunstein and Thaler's proposed standard for the evaluation of frames, she will rank the above formulations in terms of their impact upon the welfare of choosers. In this case, she should select (a), since it will incline voters with unstable preferences to vote against the increase. But let us also assume that she believes it to be the case that higher property taxes serve to improve the condition of the worst-off citizens (who, for whatever reason, tend not to vote). If she chooses to evaluate these decision frames in terms of how they affect justice (as described by Rawls), she will rank them in terms of their likely impact upon social fairness. In this case, she should select (b), since it nudges those with unstable preferences towards endorsing the proposed increase.

¹⁶ See Rawls, *A Theory of Justice*.

¹⁷ This is the highly pervasive status quo bias (see Kahneman, Knetsch, and Thaler, "Anomalies: The Endowment Effect, Loss Aversion and Status Quo Bias," in Kahneman and Tversky, eds., *Choices, Values, and Frames*, 159–170.) When an option is framed in terms of a deviation from the status quo, it tends to be greatly dispreferred when compared with frames that cast that same choice as a continuation of the current state of affairs. Thus, people display a bias in favor of the status quo, even when there is no cost to changes from the current state. This is important since it is often the case that planners have discretion with regard to which state of affairs comes to be regarded as the status quo.

To be clear, I have not provided (and will not attempt to provide) an argument demonstrating that nudges ought to be constructed in order to promote justice or utility. In both of the above cases, I hope it is clear, however, that it is possible to design nudges that are not paternalist. As a result, we should conclude that empirical evidence of the instability of preferences and of our susceptibility to the framing of decisions cannot support Sunstein and Thaler's endorsement of libertarian paternalism. Since paternalism in public policy is not inevitable, and since nudges need not be paternalistic, we require a different kind of argument in order to establish libertarian paternalism as the proper guide to public policy. In the next section, I try to construct such an argument on Sunstein and Thaler's behalf.

NUDGES OUGHT TO BE PATERNALISTIC

In order to be charitable to Sunstein and Thaler, I will here construct an argument attempting to show why, given the diversity of normative standards available to us in the design of public choices, we ought to settle specifically upon a paternalistic one. I have not been able to find such an argument in their writings, but I intend this argument to be friendly to Sunstein and Thaler's approach. Despite its friendliness, however, I believe it fails in its attempt to rescue their proposal.

It might be claimed that the moral and political conditions under which choice architects operate necessitate, or at least ought to strongly incline them towards, being paternalistic (as opposed to utilitarian or Rawlsian). In particular, one might argue that choice architects, in virtue of their social position, are morally required to act in ways that seek to benefit the citizens and consumers whom they influence. This might stem, for example, from a general obligation that those who wield power ought to use that power for the benefit of those over whom the power is wielded, or it might result from a kind of fiduciary duty to care for those who are under one's control.

A fuller version of the moral argument in favor of paternalism might be something like this: Choice architects have the ability to influence the choices of all those who act within the scope of the choices they design. As a result, whenever they have the ability to structure choices in a way that will lead individuals to make certain predictable choices, they have a moral duty to ensure that the welfare of choosers is maximized.¹⁸

¹⁸ I thank the audience at the Second Copenhagen Conference on the Epistemology of Liberal Democracy for suggesting this argument.

I think it is plausible to ascribe something like this argument to Sunstein and Thaler, given how they respond to concerns about the possibility of “evil nudgers and bad nudges.” In response to concerns that nudges might be deployed not to the benefit of choosers but rather in order to serve the interests of choice architects themselves, Sunstein and Thaler reply that:

We should create rules of engagement that reduce fraud and other abuses, that promote healthy cooperation, that restrict interest-group power, and that create incentives to make it more likely that the architects will promote the public interest. In both public and private sectors, a primary goal ought to be to increase transparency.¹⁹

Here they articulate the beginnings of a moral argument that would attempt to ensure that choice architects operate with the public interest in mind. The rules of engagement they sketch seem aimed at increasing trust in institutions: They want to cultivate a reasonable belief that choice architects will design public policy instruments in ways that serve the interests of choosers.

There are, of course, a number of different ways that we might go about trying to develop such an argument. We might try to tie the duty to act paternalistically to something like the categorical imperative, or we might try to ground it in widely held moral intuitions about trust and public service. We should note, however, that any such argument must argue on moral and not empirical grounds that choice architects are obliged to act paternalistically. Thus, what this argument requires is a moral justification of paternalism in these cases, and such a justification would have to establish that the exercise of power over an individual morally obligates the powerful to act in the interests of those whom they influence.

I suspect, however, that a utilitarian would be unmoved by any such argument. Because choice architects have the power to influence decisions, the utilitarian will claim that they ought to act in such a way as to maximize overall utility (rather than maximizing the welfare of the subset of individuals who happen to be the choosers). According to the utilitarian, the moral thing for a choice architect to do will be to exercise their power to influence others in such a way as to maximize utility. If that requires that they act in ways that preserve or enhance the public trust, then a utilitarian ought to be sensitive to those concerns. Much of the time, however, it seems likely that the way to maximize utility will not involve any paternalist goals whatsoever. Instead, the utilitarian will advocate for

¹⁹ Thaler and Sunstein, *Nudge*, 240.

designing policies that sometimes sacrifice the welfare of choosers in favor of greater increases in the welfare of non-choosers (i.e., third parties).

Similar considerations could be advanced for the Rawlsian, arguing that the exercise of power ought to oblige the powerful to abide by principles of justice. Just as the utilitarian will see the design of public choices as an opportunity to increase utility, the Rawlsian will approach the design of choice architecture as an opportunity to promote justice. Although it might be the case that a Rawlsian will sometimes recognize the need to promote social trust through the design of paternalist institutions, it isn't clear why such concerns would dominate. For citizens in a well-ordered society, it seems likely that the baseline expectation would be that the design of public institutions would be oriented towards the maintenance of justice, and not the promotion of individual welfare. As a result, a Rawlsian should not be particularly receptive to moral arguments in favor of paternalism in the design of public policy. Given that the paternalist, the utilitarian, and the Rawlsian disagree about what is just, we should expect this disagreement to extend to questions about how we ought to attempt to steer behavior. Further, these disagreements cannot be settled by recourse to empirical evidence from the Heuristics and Biases literature; such controversies are thoroughly moral.

If we interpret Sunstein and Thaler as merely trying to establish that paternalism is sometimes appropriate, then their argument is relatively plausible. But if we interpret the authors as making the stronger claim that paternalism is the only, or the most appropriate, standard for design of choice architecture, then their arguments are insufficient. Indeed, the idea that choice architects should not be sensitive to considerations of general utility or social fairness seems obviously false. At times, even Sunstein and Thaler seem to be inclined towards a number of non-paternalist policy goals. Although they attempt to cast these goals in paternalist terms, I think that a more careful analysis reveals that even they recognize that non-paternalist goals are often perfectly appropriate for choice engineers.

To see this, consider first the humorous real-world example Sunstein and Thaler use to introduce libertarian paternalism: the men's rooms at Schiphol Airport in Amsterdam: "There the authorities have etched the image of a black housefly into each urinal. It seems that men usually do not pay much attention to where they aim, which can create a bit of a mess, but if they see a target, attention and therefore accuracy are much increased."²⁰

²⁰ *Ibid.*, 3–4.

With this example, the authors seem to be claiming that the designers of airport bathrooms are justified in nudging users to improve their aim on libertarian paternalist grounds. Upon inspection, however, this is implausible. Is the choice of aim-improving urinals really to be justified in virtue of the welfare of those who use the urinals? It seems, rather, that it is consideration of the welfare of others (e.g., the janitorial staff) or of costs to taxpayers that provides the real justification for this kind of nudging. After all, aside from extreme cases of bad aim, it seems that the welfare of those who use the urinals is relatively unaffected by the choices of airport planners. This, then, seems better described as a case of libertarian utilitarianism than libertarian paternalism.

Something similar seems to be the case with Sunstein and Thaler's example of organ donation. Here the authors canvass a number of potential nudges aimed at increasing the size of the pool of organ donors through the design of the choice architecture surrounding the default position for potential donors: "A policy that can pass libertarian muster by our standards is called presumed consent . . . Under this policy, all citizens would be presumed to be consenting donors, but they would have the opportunity to register their unwillingness to donate, and they could do so easily."²¹

Although the libertarian aspect of their proposal here is clear, the paternalistic aspect is not. I take it to be rather implausible to think that this kind of choice architecture can be justified on purely paternalistic grounds. The likelihood of any one individual ever benefiting from an organ transplant is very small, but there is a great social good that is promoted by increasing the pool of donors. As a result, the justification of such a plan should center not on the welfare of would-be donors (i.e., the choosers in this situation) but on would-be recipients (this would include choosers, but would also encompass a much larger class of citizens). As a result, the most straightforward moral justification of this kind of proposal appears once again to be utilitarian rather than paternalistic.²²

Because both utilitarianism and paternalism share a focus on individual welfare, it can be difficult to distinguish utilitarian and paternalistic positions in political theory. Further, because the good of choosers is a part of the good of all, paternalistic interventions in public policy will

²¹ *Ibid.*, 177–178.

²² Similarly, nudges that are intended to save the world from environmental degradation (Thaler and Sunstein, *Nudge*, chap. 12) seem to be better motivated by concerns about fairness to future generations or overall utility than they are by paternalistic concerns.

sometimes coincide with utilitarian ones. Unfortunately, this can lead us to conflate utilitarian and paternalist justifications in public decision-making. I hope it is clear, however, that the moral arguments supporting these positions can lead to strikingly different practical consequences. Because utilitarianism seeks to promote the welfare of all, it can require us to sacrifice the well-being of the narrower set of individuals that occupy the attention of paternalists. And because paternalism is focused on the good of a particular class of people, it will sometimes conflict with the greater good. Thus, utilitarians will demand that choice architects make decisions that undermine the welfare of choosers, and paternalists will expect them to make decisions that are, from the standpoint of overall utility, clearly suboptimal. Although these kinds of conflicts may sometimes be obscured by the superficial similarities of paternalism and utilitarianism, they are pervasive in public policy.

The above arguments are not intended to show that Sunstein and Thaler do not really endorse a paternalist guide for public policy. Many of their policy proposals are clearly motivated in this way.²³ What I hope to have shown here is that their account cannot establish that the only acceptable goals for public policy are paternalistic ones. Given that they are themselves sometimes swayed by non-paternalistic motivations, we ought to acknowledge that paternalism is not the only appropriate goal for the design of public policy.

CONCLUSION

At this point, I should say something about what we ought to conclude on the basis of Sunstein and Thaler's arguments. I agree with the authors that the empirical research upon which they draw is of crucial importance for thinking seriously about the design of public policy, and the authors should be commended for bringing this research into broader public awareness.²⁴ I also appreciate their attempt to present a version of paternalism that is compatible with freedom of choice.²⁵ If the authors were

²³ This is particularly evident in Part II of Thaler and Sunstein, *Nudge* ("Money"), where the authors are concerned specifically with the personal finances of individuals, rather than with the global or even national effects of financial decision-making. One can easily imagine circumstances where the economic well-being of the individual investors under scrutiny in that section fails to coincide with the greater good.

²⁴ I rely upon much of the same research in developing my own account of the relationship between framing effects and theories of democracy in Kelly, *Framing Democracy*.

²⁵ For criticisms of this attempt, see Klein, "Statist Quo Bias"; Mitchell, "Libertarian Paternalism Is an Oxymoron"; and Glaeser, "Paternalism and Psychology."

merely interested in pointing out a previously unheralded form of weak paternalism, I would entirely applaud their efforts. Finally, I think that the authors have succeeded in augmenting the rhetorical arsenal of committed paternalists who must defend their view against concerns about government ineptitude and intrusion.²⁶ Unfortunately, the authors overreach, claiming to have justified a paternalistic approach to the design of public policy, when in fact they have not provided a justification of paternalism at all.

What I hope my arguments above demonstrate is that the terrain in which choice architects work is the familiar moral one in which we all find ourselves. Just as we confront difficult choices between the welfare of those close to us and those far away in our moral lives, or encounter seemingly irresolvable conflicts between fairness and utility in the domain of politics, we should expect the dilemmas facing choice architects to be deep, pervasive, and difficult. The moral dilemmas facing choice architects will, therefore, mirror those that confront individual agents, government regulators, politicians, and legislators. In other words, there is nothing special about the situation of the choice architect: They must contend with the same set of moral and political controversies that we all encounter in public life.

As a result, choice architects and those who direct them should not be blind to the non-paternalist goods that they can help to secure, nor should they think that their own moral landscape is simpler than it really is. Sunstein and Thaler have shown that it is possible to steer our decision-making without coercing us, without altering our incentives, and even without changing the information we have available to us. For this I think they should be commended. Unfortunately, they seem to think that this makes choices about public policy easier than it in fact does. If anything, the fact that our choices are susceptible to the framing of decisions does just the opposite: It moves difficult choices about political philosophy directly into the design of public life.

²⁶ This, for example, seems to be their aim in an earlier publication, when the authors interpreted much of the same empirical literature as motivating a kind of anti-anti-paternalist argument: “[B]ounded rationality pushes toward a sort of anti-antipaternalism – a skepticism about antipaternalism, but not an affirmative defense of paternalism” (Jolls, Sunstein, and Thaler, “A Behavioral Approach to Law and Economics,” 1541).