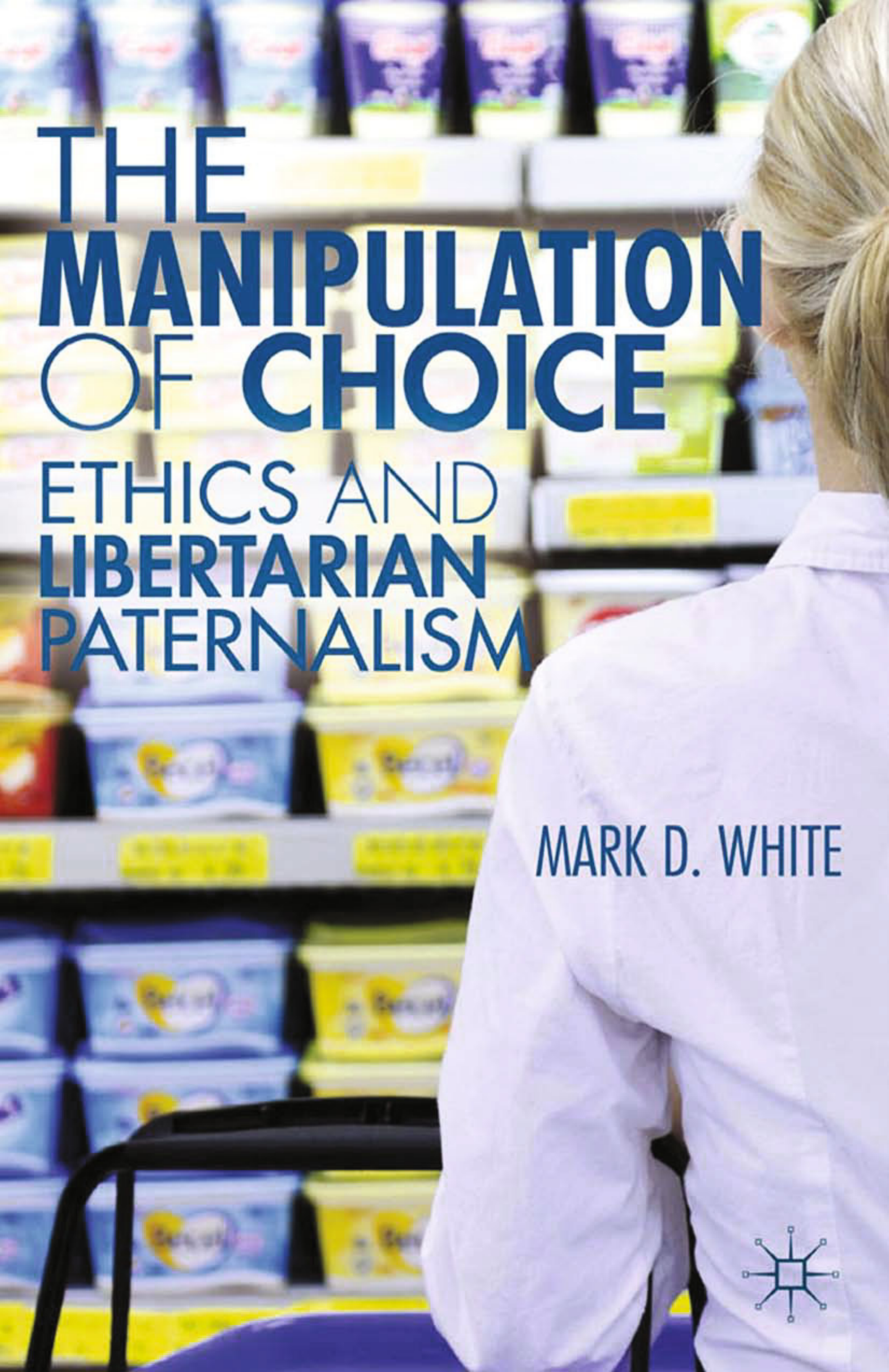

# THE MANIPULATION OF CHOICE

## ETHICS AND LIBERTARIAN PATERNALISM

MARK D. WHITE

THE MANIPULATION OF CHOICE

# The Manipulation of Choice

## Ethics and Libertarian Paternalism

Mark D. White

palgrave
macmillan

THE MANIPULATION OF CHOICE

Softcover reprint of the hardcover 1st edition 2013 978-1-137-28775-5



First published in 2013 by
PALGRAVE MACMILLAN®
in the United States—a division of St. Martin's Press LLC,
175 Fifth Avenue, New York, NY 10010.

Where this book is distributed in the UK, Europe and the rest of the world, this is by Palgrave Macmillan, a division of Macmillan Publishers Limited, registered in England, company number 785998, of Houndmills, Basingstoke, Hampshire RG21 6XS.

Palgrave Macmillan is the global academic imprint of the above companies and has companies and representatives throughout the world.



ISBN 978-1-137-28776-2 ISBN 978-1-137-31357-7 (eBook)
DOI 10.1057/9781137313577

Library of Congress Cataloging-in-Publication Data

White, Mark D., 1971–
The manipulation of choice : ethics and libertarian paternalism / Mark D. White.
p. cm.
Includes index.

1. Economics—Psychological aspects. 2. Decision making. 3. Social choice. 4. Welfare economics. I. Title.

HB74.P8W49 2013
330.01′9—dc23 2012029970

A catalogue record of the book is available from the British Library.

Design by Newgen Imaging Systems (P) Ltd., Chennai, India.

First edition: January 2013

10 9 8 7 6 5 4 3 2 1

# Contents

# Acknowledgments

This book enjoyed a long journey from conception to publication, but I won't bore you with that. (I will have plenty of chances to bore you soon enough.) I would, however, like to express my gratitude to the people involved in this book along the way, whether in a professional, collegial, or personal capacity—or, in many case, all three. My sincere appreciation and gratitude goes out to Giles Anderson, Lynn Beighley, Thom Brooks, Maryanne Fisher, Margo Beth Fleming, Jerry Gaus, Lauren Hale, Bill Irwin, Rob Lovering, Barbara Montero, Steve Morris, Muireann Quigley, Peter Simpson, Martha Starr, and Jonathan Wight. I owe special thanks to Samantha Hasey, Leila Campoli, and Charlotte Maiorana at Palgrave Macmillan (plus one anonymous reviewer) for doing such a great job getting this book off the ground. Finally, I would like to thank Cass Sunstein and Richard Thaler—as you'll soon see, I disagree with their positions vehemently, but nonetheless they started a discussion well worth having.

# Introduction: We Start with a Donut . . .

I ate a donut this morning. *Mmm*—it was good (thanks for asking). More precisely, I made the choice to eat a donut rather than something else, when I had other options, including healthier fare such as an apple or a whole-grain bagel. So why did I decide to eat a donut?

There are many possible reasons (not all of them true, of course), ranging from the banal and obvious to the very specific but just as reasonable. I like the taste, the smell, or the texture. It goes well with coffee. I was walking by the shop and it seemed like a nice change—you know, from my normal apple and whole-grain bagel. I was rewarding myself for finishing writing a book chapter the night before. I ran into an old friend who offered to buy me a donut and, despite being utterly satisfied by the apple and whole-grain bagel I had eaten earlier, I accepted his offer out of my natural impulse to be friendly and sociable. Or the stunning woman with the wonderfully green eyes and dazzlingly charming smile behind the counter made the donut sound like heaven.

But to some—especially behavioral economists and advocates of libertarian paternalism, the focus of this book—I made an irrational choice because of some inborn, evolved flaw in my decision-making process that led me to choose to eat an unhealthy donut instead of a healthy apple or a whole-grain bagel. These people, who do not know me, may decide for me that my one and only "true interest" must be my health, and will possibly take it upon themselves to pronounce judgment on my choice. They may even decide to take measures to ensure that I make "better" choices in the future. In other words, I was "predictably irrational" when I chose to eat the donut and therefore must be "nudged" into better choices—for my own good.

Wrong and wrong, as I argue in this book—"wrong" in the sense of both "incorrect" and "unethical." First, these people have no way of knowing my true interests, or why I chose to eat the donut rather than an apple or a whole-grain bagel. As a result, they have no basis on which to judge that my choice was irrational or to know what I would have chosen if I had chosen "rationally." Second, they have no right interfering in my choices to make them "better," even if they did somehow know what "better" meant to me. The real questions here are why they think they do know better and why they think they have the right to do anything about it—questions I try to answer in this book.

* * *

More generally, this book is about *choice*—how most economists think we make choices, how behavioral economists think we make choices, and why they're all wrong. Economists are social scientists, and as social scientists their job is to try to understand, explain, and predict human behavior and larger social phenomena. Economics is definitely the most successful of the social sciences in terms of developing a standard model of human behavior and applying it to a wide range of choices, from consumer purchases to health-care decisions, and from voting to mate selection (and abandonment!). But as all scientists know, models are necessarily imperfect and incomplete—they have to abstract from reality to be useful. A useful road map leaves out details that drivers don't need (such as the location of sidewalks and trees) to make the ones that they do need easier to see (such as, well, the roads). Likewise, economic models of choice must leave out some details to focus on the important ones, ideally those that are most relevant to the economic situation being modeled.

But as Albert Einstein is said to have remarked, "Make everything as simple as possible, but no simpler." Many people have criticized the traditional economic model of choice for being too simplistic to describe the way people actually make decisions, foremost among them *behavioral economists*. Relying on research by experimental psychologists, behavioral economists have identified quirks and anomalies in our reasoning abilities that cause living and breathing people to make decisions differently than what the textbook economic models predict. Popular books such as Dan Ariely's *Predictably Irrational* have introduced people around the world to many of these cognitive biases and heuristics, which lead us to place values on things that don't reflect their true usefulness to us, assess risk so poorly that we put ourselves in more dangerous situations for no good reason, and use "rules of thumb" that systematically result in bad decisions. With an enormous amount of experimental data to support them, behavioral economists have thrown down the gauntlet to mainstream economists: real people do not make decisions as well as traditional economic models of choice assume they do.

Once we accept that we naturally and inevitably make poor decisions, we might wonder if there's any way to help us make better ones, perhaps using rules that already exist to guide our behavior—the law. This is where behavioral economics meets *law and economics*, which is simply economic theory applied to legal issues. Law and economics takes the traditional model of economic choice and uses it to study how people react to legal incentives, such as liability rules or criminal penalties. It can then recommend legal reforms, based on the results of the models of choice, to create the most benefit for society as a whole. For example, law and economics would recommend higher penalties for grand larceny if that would lower the frequency of theft without imposing excessive costs on the prison system. Also, it recommends certain liability rules in accident cases to give people incentive to take the efficient level of precaution (after all, there's no use spending an extra $200 on precaution to save $100 in accident costs).

Although law and economics is mostly concerned with behavior that affects other people (and is therefore subject to the law), behavioral economics focuses on how well a person's choices promote his or her well-being, especially in the face of the psychological factors that significantly affect our choices. One way to influence those choices, in hopes of improving them, is to change the background against which those choices are made, which includes regulations and laws. Thus, *behavioral law and economics* was born, combining behavioral economics' research on people's decision-making flaws with law and economics' emphasis on regulating behavior. The result is a field of study that not only looks into how people make choices, but also attempts to modify those choices in their own interests (not just the well-being of society as a whole).

In 2008, behavioral economist Richard Thaler and law professor Cass Sunstein published a book titled *Nudge: Improving Decisions About Health, Wealth, and Happiness*, which suggested a new framework for policies designed to help people make better choices. *Nudge* quickly became a bestseller, and Thaler and Sunstein soon became influential figures in politics in both the United States and the United Kingdom. The book is full of examples of subtle and seemingly unobtrusive regulations that change the context of people's choices and thereby can influence their decisions relying on the cognitive biases and heuristics identified by behavioral economists. For example, arguing that people often fail to sign up for 401(k) plans when they start new jobs because of natural tendencies toward laziness or procrastination, Thaler and Sunstein recommend changing the default choice to automatic enrollment. New employees can still choose *not* to enroll, but such a choice would now take a little extra effort. Under this plan, if they succumb to their irrational impulses toward inaction, at least the better choice will be made—in Thaler and Sunstein's opinion, and presumably in new employees' opinion as well.

Thaler and Sunstein call this approach to public policy *libertarian paternalism*: paternalistic because its goal is to help individuals, make better decisions in their own interests, and libertarian because it's relatively "hands-off," tinkering with the options people choose from rather than interfering with the choices themselves. They contrast their brand of paternalism with more direct versions, such as motorcycle-helmet laws, which force people to do things to protect themselves; bans on trans fats, which limit people's food options; and cigarette taxes, which alter economic incentives directly. Instead, rearranging options or default choices just changes the way options or choices *look* to people, so they have the same options as before and are not forced into making any particular one. Based on people's imperfect decision-making skills, libertarian paternalism claims to "nudge" them into making the choices they *would* have made if their decision-making skills were better.

That all sounds very reasonable: since we don't make good choices, other people will help us make better ones. But... *can* they? And *should* they? As I argue in this book, there are serious problems with libertarian paternalism and nudges, based on information, ethics, and practical considerations, which combine to make such policies ineffective, unethical, and dangerous to individual choice, interests, and autonomy. Briefly, policymakers have no way to know whether a particular choice made by a person is good or bad—only that person can make that judgment because only that person knows his or her true interests and motivations for that choice. The only way a policymaker can judge another person's choice is by the policymaker's own standards and interests, not the standards and interests of the person whose choices are being judged. And regardless of whether a particular choice is judged as good or bad by the person who made it, policymakers are not justified in "nudging" that person to make a different choice which suits the interests imposed by the policymakers—especially by relying on the same cognitive biases and heuristics that motivated the nudge in the first place.

* * *

Remember the donut I ate this morning? I chose to eat that donut because it served my interests to do so—my true interests, which are complex and multifaceted, and include desires, goals, principles, and much more. I may have eaten that donut for any of the reasons I listed, for countless other reasons that I didn't list, or any combination of these. I may not have been aware, even, of the reason I ate that donut, but it was the choice I made. Was it the best choice I could have made, or even a good one? I could certainly ask myself that question, but to answer it I would have to know my own interests. It may have been a terrible decision, but that's for me—and only me—to assess.

Of course, behavioral economists may judge my choice to eat the donut as irrational. But how can they? They don't know anything about me, what my interests are, or which one (or ones) motivated my decision. They would have to assume I have certain interests and use those to judge my choice as irrational. For instance, if my only interest were my health, then my choice to eat a donut would be clearly irrational. Maybe the behavioral economists assume my only interest is my health—or even that my only interest *should* be my health. But whichever it is, their assumption or their opinion, it's *their* idea of my interests, not mine. I can decide that eating the donut was a bad idea, because I know why I made that choice—but they *don't* and they *can't*.

Nonetheless, they may take it upon themselves to act on their judgment and design policies that nudge me—and others who may presume to enjoy a donut from time to time—into making choices in their idea of our interests, not our own. In general, no matter how well-intentioned they may be, policymakers simply have no information regarding people's true interests. Instead, they judge people's choices to be bad ones based on the interests assumed and imposed by the policymakers, and then nudge people into making choices that the policymakers want them to make. Finally, they measure the success of their policies by how well they elicited the desired behavior—not by how well they improved people's choices as determined by them, because they have no way of doing this. In the end, libertarian paternalism is not about helping people make better choices—it's about getting people to make the choices policymakers want them to make.

We have a lot to learn from behavioral economists about how we make decisions in various choice situations, and we can use those insights to help ourselves make better choices. But policymakers have no information and no justification to do that for us. Let people make their own decisions—within the laws and regulations that rightfully protect them from the wrongful actions of others (and vice versa), but free of any manipulation that is claimed to be in their own interests. Some of those choices will be bad ones, but that is for people themselves to judge, not anyone else. If people realize they've made bad choices, they'll learn from it, and make better decisions in the future. Although behavioral economists can continue to treat people as "predictably irrational" for the purposes of research, policymakers should treat people as "presumably rational" and leave the choosing to us.

# Chapter 1

# The Problems with Traditional Economic Models of Choice

*Most economists understand choice to be a matter of picking the best option available to a person, such as filling your shopping cart with great bargains or selecting an automobile based on getting the best options for a good price. Sounds reasonable, right? Sure, but once we tease out the meanings of "best" and "available"—as well as other terms that economists use when they discuss choice, such as "preferences" and "well-being"—we see that the standard economic model of choice is anything but reasonable. When you get down to it, it doesn't involve any actual choosing or deciding: you see, and you want, so you get. This might be fine to describe your dog's "choices," but not yours or mine—we deserve a better model, and economists need one.*

*This chapter highlights three of the most serious problems with standard economic models of choice. First, they have far too narrow an idea of what we are choosing from. We don't merely choose goods and services based on what they can do for us or how happy they make us. For instance, many people buy fair-trade coffee or environmentally safe laundry detergent, despite the higher price, because they believe it's* right *to do so. In general, we make choices based on principles, ideals, and values as well as desires and preferences, all of which make up our interests. Second, our decision-making processes are much more elaborate than economists think. We don't just choose the* *shiniest car we can afford; we make judgments based on the wide range of considerations identified above, weighing some factors more heavily than others, or considering some before we even think about others. Finally, we have to follow through with our choices, using resolve or willpower, with or without external help (friends, support groups, and so forth). We can choose to make New Year's resolutions, but we all know how difficult it is to maintain them (assuming we even start them)!*

*Including these three factors in economic models of choice would complicate them, of course, but these additional details are necessary to understand choice well enough to explain it and make predictions based on it—not to mention use it to make judgments and policy decisions. Once we've described the problems with traditional economic model of choice, we will have laid the foundation for our critique of behavioral economics, libertarian paternalism, and nudges.*

* * *

Like many of us at one time or another, Chris and Pat have found themselves in a need of a new car. They agree on the basics—what size car they want, what features they consider essential, and what price they're willing and able to pay—but they don't have much time to look for the best deal. They've checked out the new and used car ads in the newspaper and online, they've called a few local dealers to check on availability, and they're planning to spend the next day driving around to various dealers to see what they have to offer, test-drive a few cars, and then haggle with the salespeople. At the end of the day, they hope to have settled on a car and a price with a dealer. (Wish them luck!)

Human behavior is complex, and this includes "economic" behavior such as shopping for a car. To study economic behavior, economists (like all social scientists) need to simplify it to get to the core of what leads us to make certain choices, in the hope of explaining our past choices and predicting our future ones. To do this, they have developed a simple model called *constrained preference satisfaction*, which essentially has three parts. The model starts with *preferences*, which represent what we want or need, and take the form of rankings of which things we like better than which other things. Chris and Pat have preferences about the size and features of their ideal car and alternatives, and they place values on each of these options. For instance, they may greatly prefer a midsize sedan to a compact car, but have a very weak preference for a sedan over a small SUV. They'd rather have blue exterior than red, but this preference over color is much less important to them than size is, and so forth.

But although our wants and needs may be unlimited, and we'd like to satisfy as many preferences as we can, the *resources* we have to do this are limited. Chris and Pat would love their car to be loaded with all the bells and whistles, but they can't (or won't) spend that kind of money. They'd also love to spend weeks researching all the various models and visiting all the dealers in their area, but they can't devote that much time to this one task. Money, time, energy—they're all limited, or as economists say, *scarce*. It is this scarcity that requires us to make tough choices about how to use our limited resources to achieve our most important wants and needs, which also serves as a common definition to economics. Chris and Pat may very well have the money and time to find the perfect car for themselves, but it is not worth using it to find the perfect car because they have other things they need to spend their money and time on also. (More on this point soon.)

Finally, how well we use our scarce resources to satisfy our preferences depends on the *information* we have, which is never complete and never perfectly accurate, but doesn't have to be. Chris and Pat are not automotive engineers, but they know what they want in a car. What they don't know—but would like to—is what models and options the various dealers in their area are offering and at what prices. But they can't just "know" this; they have to spend resources getting this information. In general, we need some degree of information to know how to use our resources to fulfill our wants and needs. The more information we have, and the better information we have, the more effective we'll be at solving this basic economic problem of choice, but how much information we want to get is a choice problem in itself![1]

The most basic economic activity that most of us engage in—shopping, whether for cars, groceries, clothes, or plastic surgery—fits within this framework rather well. Consumers have a certain amount of money to spend on various items that they need or want, and they have to decide which are worth their cost, given the limited amount of the money (or credit) they have to spend. Although money seems like the most important constraint when it comes to shopping, time can be a limiting factor as well—especially when it comes to last-minute Christmas shopping! Time constraints have been loosened tremendously by online shopping, but ironically, the easy, 24/7 availability of a wider array of goods and services than ever only makes our limited money seem even more scarce!

But economists don't see cost merely in terms of dollars and cents, or even hours and minutes, but rather what you can *do* with that money or time. For example, we could buy a more expensive variety of cereal (to satisfy a higher preference), but this will come at the cost of buying something else, such as bananas. We could watch the new Adam Sandler movie in the theater rather than wait to get the DVD, but this will mean we can't spend that two hours (and ten bucks) doing something else. Chris and Pat can spend an extra day visiting car dealers, but that would be a day they can't spend doing other things together. Economists call this sacrifice the *opportunity cost* of a choice, and it exists whenever a decision has to be made: if you choose to buy the premium cereal, you have to give up the bananas, and if you watch the Adam Sandler movie, you can't do something that's actually worth the time and money (such as buying a baseball bat and whacking yourself over the head with it). In other words, you made a *trade-off*, and economists are trained to analyze trade-offs like these to see when a certain trade-off will increase the person's level of preference-satisfaction and when it won't.

Preferences can cover more than just consumer goods, though. They can be defined over anything a person may devote his or her resources to, all of which will generate trade-offs when the person's resources have to be devoted to one thing or another. For example, this model is often applied to labor supply decisions, in which a person chooses how many of his or her waking hours to devote to labor or leisure. Each of these "activities" satisfies different preferences: time spent at work generates money that can be used to purchase goods the person wants or needs, and time spent relaxing provides obvious benefits—most important, enjoying the fruits of the person's labor! Each person has to find his or her own personal "best" trade-off between labor and leisure, which comes down to a number of factors, including the wage or salary the person earns at work, his or her need for money (which depends on family, mortgage, debts, and so on), and the strength of his or her preference for time off from work (within the limits provided by his or her job). This way of analyzing individual decisions has a great deal of influence on business and policymakers concerned with workplace issues such as taxes, vacation, family and sick leave, and flextime scheduling.

The same model can also be used to study the effort put into getting information. We already saw this in terms of Chris and Pat's search for information about the makes, options, and prices of cars, but it can apply to any search for information—including looking for the "best" person to form a romantic relationship with. When dating, in hopes of finding that special someone, a person needs to balance his or her picture of an ideal mate with the time and money spent looking for him or her. Some people place a high value on their time compared to finding the ideal mate—in other words, they have a high opportunity cost of search—so they settle down early but with a person who may not be as good a match as they could have found if they had been willing to spend more time looking. Others place a higher value on finding the "right one," and those people spend more time and money looking for that person (because they perceive a lower opportunity cost, and higher benefit, to more time spent searching), and on the average will likely find better matches. This model is used to study dating patterns and relationship patterns in areas that differ in terms of income, as well as other implications of the model.

Also, preferences don't have to be just about what a person needs or wants for himself or herself. Self-interest is a convenient assumption for economists because it is to a large extent reliable—at the most basic level, everyone needs to look after himself or herself—and it is the one motivation that, to some degree, we all share. But not all of our preferences have to be in our self-interest, because economists understand preferences in a very formal way: a ranking of alternative uses of resources. This can include a preference for premium cereal over bananas or working overtime over having a free Saturday, but can also include preferences for buying a birthday gift for your best friend, spending time helping a cousin move, or donating time or money to a local food shelter, over spending that time and money on yourself. These are all examples of *benevolent* or *altruistic preferences*, but a person can also have *malevolent* or *anti-social preferences*, such as donating to a hate group or spending time ruining a rival's marriage. (Don't judge me—he never deserved her!) Generally, we call these *other-regarding preferences*, because they're based on other people's well-being rather than our own. Just like self-regarding preferences, however, they generate trade-offs: if you donate an extra hour to the food shelter, it comes at the expense of an hour spent doing something else, whatever that may be. (Even watching half an Adam Sandler movie.)

The modeling framework of constrained preference-satisfaction has proven very useful for economics, helping it earn the title of "queen of the social sciences." This model and its depiction of rational decision-making—symbolized by *homo economicus*—has allowed the economic approach to spread to other fields, such as sociology, political science, and law. (Detractors call this "economic imperialism," but economists call it "more jobs.") This model is not perfect, and no responsible economist would claim that it is. After all, it's just a model, an abstraction, a simplification of reality, which is not meant to be perfect or complete. But any model that successfully explains and predicts behavior should capture the *most* important aspects of decision-making for *most* people in *most* situations, and this is where constrained preference-satisfaction falls short. There are important elements of choice that this model leaves out—and these same elements are ignored by behavioral economists, despite the various improvements they've made to mainstream economists' model of choice. (We'll discuss them in the next chapter.) Without these additional, crucial elements, economists' model of choice is overly simplistic, more appropriate to choices made by animals and computers than human beings, and we'll spend the rest of this chapter seeing why.

## Choice Is More Than Preferences

As useful as the concepts of preferences, constraints, and trade-offs are, they don't cover everything that enters into our decision-making processes. There are some factors that contribute to choices that are difficult to describe in terms of preferences, are not subject to trade-offs in the same way, and often have more in common with constraints. For instance, Chris and Pat may refuse to buy cars made by a certain company because of business or labor practices they find despicable, or will not buy from a certain dealer who made racist comments in the press. No matter how low a price or how option-packed a car they can get from that car manufacturer or dealer, no matter how many preferences could be satisfied or how many resources could be freed up for other purposes, they simply *will not* buy from these companies. We'll call these factors *principles*, referring to something that can't be measured, such as the moral principles of honesty and courage (although not all such principles need be moral in nature). Principles are elements in decision-making that are not easily traded off against preferences and are not as responsive to opportunity cost as are preferences. Obviously, this makes the traditional economist's focus on trade-offs difficult, which in turn represents a problem for the economic model of choice to the extent that principles are an important and pervasive element of choice.

Here's another example. Jodi normally gives $100 each month to her local animal shelter out of her love and sympathy for abandoned pets. If this choice were based on a preference, we would say that she chose to donate the $100 because the value of that donation to her is greater than its opportunity cost (the value of whatever else she would do with the money). If Jodi's circumstances changed and she wants (or needs) to use the money for something else—or simply has less to give—that trade-off may no longer be attractive to her. She might give only $75 to the shelter next month, because she now has more pressing uses for the other $25. Of course, the opposite may also happen instead, where she has less need for her last bit of money (or more money to use to satisfy her preferences), and so she might increase her donation to $150. As the opportunity cost of her donation changes, so does the best trade-off she can make, and her donation would change with it.

But now let's say Jodi *promised* to give $100 to the animal shelter each month. Her promise is a *commitment*, which is supported by a principle—specifically, a moral principle—to honor commitments, and possibly also by a principle that motivated the promise in the first place.[2] In this case, when

Jodi's circumstances change, she will be less likely to lower her donation. To maintain her commitment, she will try to find other ways to meet her changing wants or needs without compromising her promise to the shelter. Her promised donation, which she knows the shelter counts on each month, is not something she is willing to adjust slightly in response to an increase in the opportunity cost of giving. Her promise, and the commitment behind it, takes this aspect of choice to a different level. This is not to say that there is no circumstance that would lead Jodi to break or compromise her promise: she might lose her job, for instance, or find another cause that she feels serves abandoned animals better. Commitments are not absolute, but neither are they subject to the small changes and adjustments that preferences are, wherein we devote a few more resources to this and a few less to that when our circumstances change. It will take more than a slight change in the opportunity cost of giving to make Jodi break her promise to the shelter, just as it would take a lot to make most of us break promises—that's what makes them *promises*.

In fact, many principles have this property of limiting our discretion to make different trade-offs among preferences when circumstances changes. (Some principles play a more positive role, guiding us to devote our resources to certain purpose: Jodi's preference for supporting animal shelters may be based on the principle that every animal deserves care.) Jodi's promise of a $100 donation every month limits her ability to give less money to the shelter and use it to satisfy another preference. In the same fashion, maintaining a principle of not lying prevents you from fudging the truth on a job application to increase your chances of landing a great position, and a principle of fairness keeps you from cutting in line at the taxi stand so you can get home from the airport a few minutes earlier. And as we saw before, principles can also enter into our consumer behavior, such as Chris and Pat's refusal to buy cars made by a certain manufacturer or sold by a certain dealer.

Some economists would like to maintain that principles such as these are still preferences; they would say that we prefer the option of "not lying and losing the job" to the option of "lying and getting the job." But this implies too much flexibility on the part of the decision-maker to make slight changes in principled behavior based on circumstances. For instance, we could imagine that a small increase in the pay, benefits, or prestige of the position you're applying for could challenge the strength of your preference for honesty if it's "just" a preference. But you would be much less likely to compromise a principle in the face of small changes in the reward from doing so. In this way, principles act more like constraints than preferences, because they keep us from making certain trade-offs that we find unacceptable for reasons above and beyond preference-satisfaction.[3]

## Constraints Are Not Always So Constraining

Even though many principles act like constraints in that they limit the choices and trade-offs people can make to satisfy their preferences, they don't seem like the kind of constraints we mentioned when we introduced the economic model of choice earlier in this chapter. Those were constraints on available resources, primarily money and time, over which we would seem to have no control. But we can decide whether or not we maintain principles, and we saw earlier that principles are not absolute—we won't compromise them as easily as we would adjust trade-offs among preferences, but they can be compromised, or violated completely, if there is enough reason to do so. The job applicant may not lie on his application to earn an extra $1000 a year, but if he can earn enough in this job to pay for his son's lifesaving operation, he might consider lying to be a necessary evil.

There are two ways to look at the issue of treating principles like other constraints on choice: the first is to show that principles *can* be truly binding; and the second is to show that the usual resource constraints are *not* always binding. If we refuse to accept that principles can be binding, that would imply that we act solely out of self-interest. We would behave ethically only insofar as it benefits us, including any happiness we get for doing good and any punishment we can avoid from not doing bad. This is clearly consistent with the economic model of choice, but it is not consistent with everyday experience. We observe and maintain principles every day; society could barely survive, much less thrive, if we didn't. We do the right thing even when it's inconvenient, whether it's stopping to help a fallen senior citizen even when we have an urgent appointment to keep, or passing on an opportunity to lie, cheat, or steal, even when there's little chance of being caught. Even the market economy, all too often characterized as a morality-free zone, depends on honesty and trust to function. Most people are not looking for a chance to get away with cheating, lying, and stealing when buying or selling in the marketplace, which allows for the market to function relatively smoothly. (The recent financial crisis shows what can happen when these basic moral principles are not observed.)

To put it more positively, we do choose freely to follow principles such as moral ideals, but in doing so we bind ourselves to them. Principles are an important part of most ethical systems, whether they resemble the equal concern for all embodied in utilitarianism, the duties in Kantian deontology, or the virtuous character traits in the virtue ethics of the ancient Greeks and Romans. Most importantly, the principles we choose to follow and be constrained by become part of who we are, our moral character, and identity. They may not be binding in the same sense that we think of resource constraints as binding, but they are binding in the sense that when we compromise them, we compromise ourselves. The fact that they are chosen makes the fact that people feel bound by them all the more ethically meaningful. Of course we can't spend more money than we have (counting credit), but holding oneself to saving a certain percent of your income every month—or donating it to an animal shelter—even when there are many other things you could do with the money is admirable. In general, people give up their careers, their freedom, even their lives for the sake of principles. We shouldn't think of principles as constraints that are "only" chosen, because chosen principles are actually much more important to us as moral individuals than imposed ones are.

What about ordinary constraints on money and time? Surely we have to take them as given, don't we? Actually, not always; ironically, these constraints often rely on principles to make them binding. The father who can't afford food for his family within his budget doesn't have to accept his budget constraint as given—he can steal the money he needs. The executive who can't finish her earnings report before the Friday deadline doesn't have to turn in an unfinished or incomplete report—she can lie about being sick and finish it over the weekend. To the extent that we treat our resource constraints as binding, it is only because we hold ourselves to principles that will not allow us to bend those constraints. There is almost no resource constraint that can't be stretched if a person is willing to do what it takes—except the constraint of 24 hours in a day! This shows that the true limits on our resources are based on our moral principles, and the constraints of our money and time are only as binding as the moral principles that underlie them.

### Another Thing Preferences Don't Do Well

However, economists don't use preferences simply to model choice; they also use them to measure people's well-being or welfare. They assume that the more a person's preferences are satisfied, the better-off that person is, so in their view preference-satisfaction ends up being equivalent to well-being. But this is problematic, because many of the things we prefer are not necessarily good for us in the way we usually understand "well-being." The two best examples are self-destructive preferences and other-regarding preferences. On the one hand, self-destructive preferences, by definition, do not serve our well-being. Well-being is a vague concept in itself, but most versions of it would include some measure of health, and most of us have preferences for unhealthy foods, if not drugs or self-violence.[4] Other-regarding preferences, on the other hand, don't contribute to decision-makers' own well-being precisely because such preferences are focused on other people, not the decision-makers themselves. Even if a person has altruistic preferences (or even malevolent ones) toward other people, these preferences will lead him or her to make choices intended to benefit (or harm) other people, not to make himself or herself better-off (other than whatever amount of "warm glow" or self-satisfaction the person gets from helping or hurting someone else).[5]

We've said that a person's preferences do not represent his or her well-being, and also that they're not the whole picture when it comes to explaining his or her choices (due to the role of principles).[6] Well, guess what: there is no necessary link between choice and well-being either. People do not always make choices consistent with their well-being, whether because of self-destructive or other-regarding preferences, or principles that they consider binding and more important than their personal well-being. People have a much wider range of possible motivations for their decisions, including preferences, their personal well-being and the well-being of others, moral principles, or more general social ideals such as justice, fairness, or equality—and that's just to name the positive ones! More generally, any of these factors can be put in terms of *reasons* for choice, and we can refer to the reasons people have to make decisions, whatever they care enough about to act in pursuit of, as their *interests*.

My argument at this point—which is essential to our discussion of behavioral economics and libertarian paternalism later in the book—is that *a person's interests are the basis of his or her choices*. These interests cannot be reduced to preferences alone, but rather are based on preferences, the person's individual well-being, the well-being of others, personal principles, societal ideals, and whatever else matters to the person—and cannot be reduced to any one of them. Any particular choice can be explained in a number of different ways, from simple preference-satisfaction to a principled stance, but economists cannot even consider the range of possible explanations if they do not admit the possibility of choice outside of preferences.

## The Crucial Role of Judgment

As we said earlier, one virtue of the traditional economic method of choice is its simplicity: the economic actor makes choices to reach the highest level of preference-satisfaction possible within his or her resource constraints. (Economists refer to this as *utility-maximization*, although "utility" here means a number referring to how many preferences are satisfied, not well-being or happiness as utilitarian philosophers such as Jeremy Bentham and John Stuart Mill understood it.[7]) Even the simplest consumption decisions are more complex than this, but in many such cases consumers may behave "as if" they maximized their level of preference-satisfaction, in the same way that the master pool player and NBA free-throw shooter acts "as if" they are also masters of geometry and physics. In many case, however, people face conflicts between their preferences and principles, between what's good for them and what's good for others, or between what seems best and what seems right. When this happens, their choices cannot be reduced to seeing how far up the preference ladder his or her income will allow him or her to go, since an adherence to principles may prevent climbing the ladder at all (or shift the person to a new one with certain rungs that are incompatible with his or her principles knocked out).[8] This is why choices cannot be reduced to preferences, well-being, or any simple explanation based on a single reason or motivation.

Let's say Jim is shopping for coffee and feels strongly that he should buy fair-trade coffee based on a principle of solidarity with poor coffee growers. He pays the higher price for fair-trade coffee, even though it may not taste better, since it serves a principle that is important to him. This assumes, of course, that he can afford the higher price of fair-trade coffee; there is still an opportunity cost to following principle, specifically the higher price, but this opportunity cost is acceptable to him. Under different circumstances, however, it may not be: suppose his household income were to drop or his other expenses were to rise. Money becomes tighter, and the opportunity cost of buying fair-trade coffee rises because the sacrifices Jim must make to buy it become more severe. Even though Jim may not be as responsive to this rise in opportunity cost as he would be if it arose from buying better-tasting premium coffee, there will likely be a cost high enough to force him to compromise his principle and buy conventional coffee (or go without coffee altogether—no, really, it could happen!).

This is very similar to Jodi and her donations to the animal shelter, and both are cases of *conflicting principles*: Jim feels he should pay the extra money for fair-trade coffee, but he also has other household needs that must be met with his shrinking resources, such as feeding and clothing his family. Where he once had enough money to do both, now he must make a choice. Note that this is not as straightforward as simply accepting a lower level of preference-satisfaction or utility; this dilemma forces Jim to choose between his principle of supporting fair trade and his principle of supporting his family. When it just meant taking the kids out for ice cream one fewer time per month, Jim was satisfied acting out of the principle to support fair trade, but when it becomes an issue of putting enough food on the table for dinner, Jim may have to compromise his support for fair trade. His family comes first, and his obligation to support them takes precedence over his devotion to supporting coffee growers.

If these were all preferences, Jim could rank them and use the standard economic model of decision-making like usual. However, his fair-trade principle is not a preference and is not easily traded off against other preferences, although at some point, as his resources shrink, it must bend to them. But how can he value a principle, especially a moral principle, which is an incalculable, qualitative ideal? Such things are supposed to be above value, but in a world of scarcity they cannot be. This does not mean that the value is easily determined; it can only be inferred from the point at which Jim lets go of his fair-trade principle and buys conventional coffee. Jim must use his *judgment* to balance two principles that are both important to him, and he must arrive at the answer that maintains the integrity of his moral character as he envisions it.

The idea of resolving conflicts of principles, or *moral dilemmas*, is well-known to moral philosophers, especially those that deal in duties or virtues, but would seem strange to economists and their utilitarian cousins, who simply rank options according to preferences and then see how many they can satisfy with the resources at hand.[9] In their framework, everything is reducible to a numerical value, which in turn makes everything commensurate and comparable. But principles do not work like that; they resist being quantified or adjusted to fit some mathematical procedure or algorithm. They are ideals, absolutes, that only bend to the tragedy of scarcity, and even then only with a scream and a howl.

Once Jim acknowledges that putting healthy food on the table to feed his kids is more important to him than supporting fair trade, he buys conventional coffee. But he doesn't make this choice in the spirit of having successfully made an optimal trade-off that maximized his level of preference-satisfaction. He does so with a heavy heart because he realizes he had to compromise an important principle, one he still believes in but cannot support given his present circumstances. Another person, John, with a preference for fair-trade coffee may very well make the same decision but along very different lines, and the choices would mean very different things to Jim and John. When principles are involved, decision-making takes on a distinctly more elaborate and less calculative flavor, one that involves judgment and balancing rather than changing the numbers in a formula, as the standard economic model of choice would have us do—because it can do nothing else.

### The Difference between Deciding and Acting

The final way in which the traditional economic model of choice falls short of an accurate representation of the real world is its neglect of human beings' weakness of will. In the economic model, once a person makes a choice that maximizes his or her level of preference-satisfaction or utility (within constraints), he or she carries out that choice—obviously! If Janet determines that the large can of baked beans is a better deal than the small can, or if spending three hours studying for an exam is the best use of her time, or if running ten miles every morning will be the optimal way to lose weight, then of course she will buy the large can of baked beans, spend three hours studying . . . and go running every morning. Without fail. *Every* morning. Because that's what she decided to do. Just like all of us who decide to give up smoking, eat less fast food, procrastinate less, and stop watching reality TV. We *all* act according to what we decide is best. Right?

If you said "right," then feel free to move to the next section—you just won't understand. But if you're like the rest of us, you get the point. Deciding to do something and going through with it are two very different things. The economic model, however, does not recognize this distinction, because it's all about choosing, not acting on that choice. And that's fine in some cases: most people would have no problem buying the large can of baked beans once they realize it's the better value. But studying for three hours straight may be difficult, especially if there are ample distractions, and getting up every morning to go running is nearly impossible (or, umm, so I've heard).

To put it another way, the traditional economic model of choice has no concept of *willpower*. Willpower is the necessary bridge between making a decision and acting on it, and the economic model doesn't make this distinction. To be sure, economists have tried to model the effects of willpower, but they do it within their framework of preferences and opportunity costs, arguing that a person's preferences change once the decision is made, or some preferences are stronger at some times and weaker at others. These changes in preferences and costs then cause people to *change* their decision, not fail to carry out the decision they may still maintain was the best one. But very few economists recognize the existence of a will that either carries out the decisions that our judgment tells us is best, or leads to another action altogether. And because of this blind spot, economists can explain when people are more likely to exhibit weakness of will (when their preferences dictate), but they cannot explain how people *resist* their weakness of will and soldier on despite it![10]

Let's say Fred walks by a bakery. One thing you should know about Fred is that he loves his cherry danish. But Fred also knows he shouldn't have cherry danish because Fred's doctor told him he has to lose weight, *blah blah blah*.[11] His judgment tells him he should just walk on by the bakery, but the temptation threatens to draw him in. (Come on, people—it's cherry danish.) Fred has trouble acting on his better judgment because of the immediate draw of the cherry danish, and he has to use his willpower to keep him on the straight and narrow. The stronger his will, the more often he will able to abstain and keep walking, and the weaker his will, the more often he will succumb to the fruity, flaky treat.

Forget Fred and the cherry danish—well, at least forget Fred—and substitute your own personal struggle with self-control and temptation. Maybe you're in recovery but find yourself at a party with alcohol. Or you gave up smoking a week ago but pass through a group of smokers lighting up outside their office building. Or you're supposed to be finishing an important project for work, but instead you're reading *this book*. (Thank you!) This isn't a matter of preference and principle; your judgment settled all that when you arrived at a choice. This is about sticking to your choice in the face of temptation, which is an all-too-common situation, and success depends on your willpower—nothing more.

Furthermore, there is no formula or algorithm that can determine what you will actually do once you've made a choice, again because it's simply a matter of your willpower. Philosopher John Searle writes about a "gap" in our rational processes between decision and action. This gap is where some sense of free will resides—not necessarily free will in the big picture metaphysical sense, but in the sense of being free to make choices against what our judgment tells us to do (based on our preferences and principles).[12] This is where we find true choice, which the traditional economic model of choice lacks entirely. In that model, we do what our preferences dictate, and we have no role to play in that whatsoever. Even if we admitted other reasons such as principles into the economic model, the results of the model would become our choice, without us actually doing any "choosing." But with Searle's gap, there is space in our thought process for us to make a truly free choice, one that isn't determined by our preferences and principles but one that we can make for whatever reason we want—or with no reason at all! Searle anticipates the question, "What fills the gap?" and offers this answer: "What fills the gap? Nothing. Nothing fills the gap: you make up your mind to do something, or you just haul off and do what you are doing to do."[13] This seems irrational, but to Searle, it is the very sign of true rationality: if choice were totally determined by preferences and principles, we would be mere machines or beasts, not human beings. True rationality is more complex and less precise than any formula or model can represent, no matter how advanced our modeling techniques and computing power become, because there is a point at which we simply "flip a switch," regardless of what our best judgment tells us to do. And it is this, to Searle, that explains weakness of will, in which action is divorced from judgment.

Because of economists' neglect of the crucial role of the will in bridging the gap between decision and action, they can try to explain the former but have no tools to deal with the latter. Although they may construct insightful and brilliant models based on odd configurations of preferences and costs that can explain why people are prone to weakness of will, they cannot explain why—in the face of these clear incentives to lapse—people nonetheless show resolve, sticking to diets and avoiding procrastination. We can consider such resolve to be another example of adherence to a principle, specifically a principle of self-respect that impels us to stick to the goals we set for ourselves regardless of the pull of preferences and desires that drag us away from them.[14] But just as devotion to our principles must sometimes bend to scarcity of resources, it also succumbs on occasion to simple human weakness. Economists do not have the tools to deal with this due to their deterministic model of decision-making in which preferences and constraints run the show. Because of this shortcoming the economic model fails to explain a wide range of common behaviors correctly, and as a result cannot predict or anticipate them with any accuracy either.

### Next Step: Getting Behavioral

In this chapter, we discussed how the traditional economic model of choice falls short in explaining ordinary human decision-making due to its strict adherence to preferences and material constraints, and the resulting ignorance of principles, judgment, and willpower. Models have to abstract from details, of course, but the economic model of choice leaves out critical factors of decision-making that explain much of real-world behavior. The takeaway from this chapter is that (1) choices are never as simple as economic models claim, (2) there is much more to choice than preferences and constraints, and (3) people make choices in their own interests, which may not correspond with either their preferences or their well-being (however we decide to define it).

In the next chapter we'll turn our attention to *behavioral economics*, which claims to remedy the shortcomings of the traditional model by incorporating the insights of experimental psychology. Although behavioral economics certainly corrects some serious mistakes in the economic model of choice, adding a healthy dose of realism, we'll see that it does little to deal with the problems identified in this chapter. Despite its more elaborate description of the process of decision-making, it still preserves the exclusive focus on preferences and constrains that serves as a straitjacket for economics. Due to this shortcoming, behavioral economists leave the door open for the paternalistic policies that are often "read into" their results, policies that we will spend much of this book examining.

# Chapter 2

# How Behavioral Economics Makes the Same Mistakes

*Behavioral economists recognize how simplistic the standard economic model of choice is and have proposed more elaborate models as alternatives. Using the work of experimental psychologists who studied decision-making in various contexts and situations, behavioral economists pointed out specific quirks or anomalies in our decision-making processes, systematic and predictable errors that prevent us from reaching our true goals. These various cognitive biases and heuristics include framing effects, referring to the effect that the presentation of options has on the choices people make, and endowment effects, which tend to make you demand more to give up something that you own than what you would offer to get it if you didn't. In many cases, these cognitive effects represent significant deviations from the textbook models of rationality used by mainstream economists, and have helped to generate more accurate explanations and predictions of choice in studies of consumer behavior, financial markets—and legal studies, as we'll see in the next chapter.*

*Through best-selling books that explain how these cognitive dysfunctions relate to decisions each of us makes every day, behavioral economics has become an important part of scholarly and popular discussions about human beings' decision-making abilities. Although their findings have indisputable value in terms of better understanding choice, behavioral economics suffers from the same essential problems as the traditional economic model of* *choice: a focus on preferences and constraints with limited consideration of principles, judgment, or willpower. As this chapter explains, behavioral economists focus on minor flaws within the existing economic model of decision-making, making incremental improvements rather than questioning the entire model and asking if anything needs to be changed or added (as suggested in the last chapter). More important, these shortcomings only contribute to the more serious problems to come when the scientific conclusions of behavioral economics are developed into government policy.*

* * *

Despite its stature in economics and its application to other fields in social science—or perhaps because of them—the traditional economic model of choice has been roundly criticized on many grounds other than the ones we discussed in the first chapter.[1] One critique in particular has caught on in a big way, not only with academic economists, but also with policymakers and the general public. It comes from *behavioral economics*, which criticizes the traditional model in much the same spirit that we did: arguing that it may be elegant and simple, but it does not adequately describe or predict real-world behavior. However, whereas in chapter 1 we made an argument for extending the model to include more decision-making factors (such as principles and interests) and layers of choice (judgment and willpower), behavioral economists make an intensive argument, specifically that the model needs to modify the existing structure of preferences, constraints, and information to make it more realistic.

According to behavioral economists, the model of constrained preference-satisfaction demands far too much from economic actors in real-world choice situations, and as a result it doesn't account for how real people handle the trade-offs and opportunity costs. Thorstein Veblen, a founding figure in institutional economics who coined the term "conspicuous consumption" in his 1899 book *The Theory of the Leisure Class*, ridiculed the traditional model—decades before it even took its current form—for assuming that the economic actor was a "lightning calculator of pleasures and pains who oscillates like a homogeneous globule of desire of happiness under the impulse of stimuli that shift him about the area, but leave him intact."[2] Although the part about "oscillating" under the "impulse of stimuli" recalls the point we made in the last chapter about the absence of true choice in the standard economic model of choice, it is the "lightning calculator" part of the quote that appeals to behavioral economists. It successfully parodies the traditional model's implicit assumption that no matter how complex a choice situation, people can instantly perform all the mental calculations of benefit and cost, or "pleasures and pains," to arrive at the utility-maximizing choice.

Half a century later, Herbert A. Simon, a Nobel laureate in economics and noted scholar in many other fields including political science, computer science, and management, proposed an alternative conception of decision-making called *bounded rationality*, designed to be realistic concerning human beings' abilities to process information. Here is how Simon described the traditional model:

> [Economic man] is assumed to have knowledge of the relevant aspects of his environment which, if not absolutely complete, is at least impressively clear and voluminous. He is assumed also to have a well-organized and stable system of preferences, and a skill in computation that enables him to calculate, for the alternative courses of action that are available to him, which of these will permit him to reach the highest attainable point on his preference scale.[3]

Simon makes the same point that Veblen did, but in terms of the standard model of choice which by then had been formalized, with its emphasis on preferences, constraints, and information—and "lightning calculation" of it all.

Let's imagine a simple choice situation, and see how quickly it grows complex when understood within the traditional model.[4] Susan needs to fill her car with gasoline on the way home from work today. There are a dozen gas stations she could go to, scattered throughout the area, with an array of prices, quality of gasoline, congestion (lines), and additional services available (convenience stores, credit cards with benefits, toy trucks, and so on). Naturally, she has preferences for saving time and money, but also for getting quality gasoline and the amenities she enjoys. So she has to assess each of the 12 options in terms of price, quality, other goods, and the time it adds to her trip home (plus added hassle due to traffic and lines, which differs from one gas station to the next)—assuming she has all of this information to begin with—then rank them all according to her preferences for each, make the appropriate trade-offs, and make her choice. All that for gas (and maybe a toy truck)!

Simon's model of bounded rationality, in contrast, was intended to be a more reasonable and realistic picture of human decision-making. He proposed a number of simplifications to the traditional model to bring it in line with actual cognitive capacities of real people, such as narrowing down the range of options to choose from and eliminating all but the essential details of these options, thereby reducing the information and calculation needed. His model accounts much better for how Susan will probably make the decision: only consider the two or three closest or most convenient gas stations, use the information she has and not worry about what she doesn't know, and only consider the one or two most important features of each station (perhaps price and travel time). She might not make the same perfectly optimized decision she would have if she considered every detail, but even if we assume she could balance the myriad aspects of the decision problem, the time she saves in overthinking the matter is well worth it. In the same way that searching for information is subject to opportunity cost, so too is lengthy, complex decision-making. More significant decisions, such as buying a house, are worth long deliberations (and extensive search for information), but most choices we make on a day-to-day basis are not.

Another term Simon uses to describe this process is *satisficing*, by which a person settles on a merely satisfactory choice rather than making the perfectly optimal choice. Given the cognitive demands of arriving at the best answer possible, Susan narrows down her choice problem and arrives at a choice of gas station that she regards as "pretty good" if not ideal, which is fine given the relative insignificance of the consequence of a less-than-optimal choice (a few extra cents paid at the pump or a couple extra minutes spent in traffic). Even with the purchase of a house, she would not look at *every* available house and assess *every* facet of each one; she would spend more time and consider more options than she did when buying gas, but she will still satisfice to some degree. In a way, it's impossible to consider every possible choice, so we have to satisfice—we simply cannot make choices like the traditional model would have us do.[5]

### Where Psychology Meets Economics

Although Simon definitely brought psychological research to bear on the process of decision-making, it was a team of an economist and a psychologist that launched the modern behavioral economics revolution. The work of Daniel Kahneman (a psychologist who was awarded the Nobel Prize in economics in 2002) and Amos Tversky (an economist who was ineligible for the award due to his death in 1996), along with various other collaborators, laid the formal groundwork for behavioral economics by establishing the existence of a number of anomalies in human decision-making that deviate from the textbook model of rationality used by economists. These anomalies include *biases*, distortions in the way we understand and value of options, and *heuristics*, "rules of thumb" we unconsciously use to simplify complex choice problems, as well as well-known human frailties such as weakness of will.

Although their findings challenged the traditional economic model of choice, Kahneman and Tversky defended their arguments with copious experimental evidence that showed the systematic nature of these biases and heuristics. They also provided evolutionary reasoning for these "quirks" in decision-making, arguing that they enhanced reproductive success by (for example) focusing the brain on short-term survival over long-term flourishing. Today, the results of behavioral economics are widely accepted and are being integrated in economic models in a wide range of areas. It's not stepping out on a limb to imagine that in several decades, behavioral economics will be become simply "economics," and behavioral insights will be accepted as part of every economist's toolbox.

The cognitive biases and heuristics identified by Kahneman, Tversky, and their colleagues are too numerous to describe here at length, and the precise details of them are not important to the overall argument of this book. I'll just give some examples of common choice situations and suggest how the various decision-making anomalies play into them.[6]

- George waits in line for 36 hours to pay $100 for a choice ticket to see his favorite band—the most he would be willing to pay. As he walks back to his car, someone offers him $200 for his ticket, but George refuses. This implies that he values the ticket more than the $200, even though he was only willing to pay $100 for it in the first place. This reflects an *endowment effect*, wherein George places a higher value on the ticket once he has it than before he had it. His position also shows *loss aversion*, by which George seems to be more concerned with losing $200 out of his pocket than losing $200 he could have gained (but never had).
- Susan frequents her local diner, calling it "comfortable and familiar," even though the food is not fantastic and

the quality of the service varies from visit to visit. Her colleagues at work rave about another diner in her area, with wonderful food and first-class service, but nonetheless she hesitates to go. Susan may be subject to *status quo bias*, in which we have a strong preference for the way things are that leads us to discount alternatives which may be more preferred if we gave them a fair chance.

- Dave is shopping for a new cell phone. He dismisses the cheapest, most basic phone, turning his attention to more expensive phones with more features. He “oohs” and “ahhs” at the premium model, but considers it ostentatious and too expensive, and he instead chooses a mid-range model. Dave exhibits *extremeness aversion*, which biases us against options at the extreme ends of a range (in terms of size, quality, price, and so forth).
- Diane hates to exercise and can cite a number of articles about the dangers of excessive exercise. She cannot, however, recall any of the articles she’s read about how essential moderate exercise is to good health. Diane shows *confirmation bias* by unconsciously placing more emphasis on information that confirms her beliefs. And if she’s lucky and doesn’t suffer any ill effects from her refusal to exercise, Diane will likely say she “knew all along” that she didn’t have to exercise, even though her survival was a matter of luck rather than science. This shows her *hindsight bias*: she interprets her “20/20 hindsight” as confirming her beliefs and disguising the risks she took.
- Joe finds the rare baseball card he needs to complete a set in his collection, and the only way he can pay for it is by drawing from a savings account he set up for the trip to Europe he has planned for next year. Although he is confident he can replace the money in his savings account by the time he has to buy plane tickets, he is hesitant to, and passes on the baseball card out of reluctance to draw from his vacation account, even temporarily. Joe is engaging in *mental accounting*, by which he treats some money

differently than others, other than treating all his money as equivalent (or *fungible*, as economists would say). Mental accounting can be an effective tool for self-control, but can overly restrictive if adhered to rigidly (as in Joe's case).

- Jane watches the latest report of a jumbo jet crash that leaves no survivors, and she decides to drive to her next out-of-town business meeting rather than fly. She may even be aware of the statistics that show road travel to be much more dangerous than air travel, but just cannot imagine flying after watching the video of the flaming wreckage and the grieving relatives. This can be chalked up to the *availability heuristic*, which gives more recent and salient information a disproportionate impact on our decision-making, working in conjunction with basic *risk aversion* (based on the more visceral, although smaller, risk of flying compared to driving).
- Dave is a single man who regularly engages in unprotected sex. He is well aware of the risks, but his risk aversion is overwhelmed by *optimism bias*, which comforts him into thinking that bad things will "never happen to me." Diane may also be guilty of optimism bias in her refusal to exercise, especially if she never perceives any harm from her behavior. (Dave and Diane may be perfect for each other—what do you think?)
- Dave and Jane, while opposites in how they respond to risk, both serves as examples of the profound difficulties that human beings have in assessing and comparing situations of risk.[7] Another Nobel laureate in economics, Maurice Allais, is famous for demonstrating problems with decision-making under risk, of which the "Allais paradox" is a prime example. Simply put, people react differently to being told that a medical procedure carries a 10 percent chance of death than they do to being told there is a 90 percent chance of survival, even though they are equivalent statistics. This example, and the much more elaborate situations set up by Allais and his successors, involves a combination of risk aversion, loss aversion,

and *framing effects*, in which the way that facts and options are presented have a disproportionate effect on the choices made based on them.

I could go on—really, I could—but I hope these examples are sufficient, not only to suggest the numerous ways that we human beings deviate from the textbook models of rational choice, but also how common these phenomena are.[8] Furthermore, these are not mistakes that average out over time, such as missing a putt or free throw to the left half the time and to the right the other half. Instead, these are systematic and therefore predictable errors, but unlike pitching to the left consistently, they are not as easy to correct by adjusting your aim. According to the arguments of behavioral economists, these biases and heuristics are hardwired into our brains through thousands of years of natural selection. Risk averseness, in general, would have enhanced survival in the average person more than risk-seeking would have (although I would imagine that the risk-seekers probably got more dates!). However, just as our evolutionarily selected preference for sweet and fatty food served us well when food was scarce but works against us in a world of cheap and plenty, the deviations from textbook rationality listed above may have served us well in the wild, but less so in the modern world where mortgage terms need to be assessed, insurance options have be considered, and business proposals need to be ranked by their likely profitability.

It's hard to overstate the magnitude of the impact that these new findings had, both on popular and academic discussion. Because these observations are so germane to everyday decision-making, they were obvious candidates for presentation in popular trade books such as economist Dan Ariely's *Predictably Irrational*.[9] In the academic world, they led perhaps most directly to the field of *behavioral finance*, in which economists such as Richard Thaler explained that the world of finance—a field in which mathematics and the profit motive were presumed to rule out "odd" preferences, leaving "perfect" rationality to dominate—turns out to be influenced tremendously by cognitive biases and heuristics.[10] Likewise, legal scholars such as Cass Sunstein applied the insights from behavioral economics to the decisions of individuals involved with the law in various ways, from criminals to judges to jurors (whose ability to process information impartially is especially crucial to the operation of our trial system).[11] And together, Thaler and Sunstein wrote a book that, frankly, is the reason I wrote this one, and which we'll be talking a lot about in the next chapter. (All in good time!)

### In My Day, Preferences Were Stable!

One implication of behavioral economists' findings regarding cognitive biases and heuristics threatens the very core of the economic model of choice: its dependence on preferences. Many of the "quirks" surveyed in the last section deal with the way we perceive the costs or benefits of the options we have to choose from. In other words, they alter our preferences regarding those options. Jane's preference for air travel over car travel declines after news of a deadly crash, though the risk remains low; Joe's preference for the last baseball card he needs to finish his collection over the amount of money it costs declines because he can't buy it without drawing temporarily from his vacation savings; and Dan's preference for unprotected sex over protected sex (or no sex) is higher than it would be if he fully realized and appreciated the risks involved. What the three statements above have in common is that they all represent preferences changing or reversing based on the context under which they were formed.

The title of a 1995 paper by psychologist Paul Slovic, one of Kahneman and Tversky's frequent collaborators, sums up this interpretation: "The Construction of Preference."[12] Due to the distorting effects on cognitive biases and heuristics, behavioral economists consider preferences to be formed only when a decision has to be made using them. Furthermore, different decision contexts—such as those in the Allais paradox—can lead to different preferences being formed, expressed, and acted upon. As Slovic writes, "preferences appear to be remarkably labile, sensitive to the way a choice problem is described of 'framed' and to the mode of response used to express the preference."[13] Since preferences seem to depend on the particulars of each choice problem, according to behavioral economists, they are better understood as constructed when the decision calls for them, rather than being "ready" all the time, as if people "could look up their preferences in a book, and respond to situations accordingly."[14]

This is extremely troubling to mainstream economists, who depend on the assumption of fixed and stable preferences to provide a firm grounding from which they can explain choice and measure well-being. In studying behavior, economists normally change prices, income, and product selection and then observe or predict how people react—assuming their preferences are fixed. But if preferences can change at the same time, economists will find themselves without a foundation to start withor a basis upon which they can understand changes in behavior. For example, if the price of bananas rises, we would expect a person with stable preferences to buy less of them (and more of something else). But if the person sees the price of bananas rise and makes the conclusion that they must be better-tasting bananas, he or she might buy the same amount or even more of them because his or her preference for them increased due to the (imagined) higher quality! Not only is this behavior difficult to predict—who knows if a consumer will interpret an increase in price that way?—but also it's almost impossible for economists to explain, because this behavior would seem to violate the law of demand (people buy less of something at a higher price, all else the same). All else is not the same, of course—in the consumer's opinion, the bananas have increased in quality—but mainstream economists don't have the tools to explain this change in preference absent an independent reason for it.

However, because their emphasis remains on determining the precise nature and behavior of preferences (even as they adapt to difference decision-making situations), behavioral economists make the same mistake that mainstream economists make. They are correct to doubt the stability of preferences, but the larger question is why they focus on preferences so much at all? Both mainstream and behavioral economists choose to interpret all of a person's choices in terms of preferences, when they're actually based a wide range of influences (as discussed in the last chapter), and then they wonder why the resulting "observed" preferences behave strangely. The economist is forcing preferences to do all the explanatory work of the model, work that they simply aren't suited to do, because preferences can't explain human behavior based on principles, judgment, and willpower without changing or reversing to reflect these other influences.

Consider Jim, the consumer from the last chapter who preferred to buy fair-trade coffee and was willing to pay a premium for it based on a principle of supporting poor coffee growers. Let's suppose the price of fair-trade coffee falls to be almost equal to the price of regular coffee. Rather than buy more of it, let's say Jim stops buying fair-trade coffee altogether and starts buying regular coffee instead. An economist, interpreting Jim's behavior in terms of preferences only, would be flummoxed. "The price fell," the economist would exclaim, "but Jim bought less?" A mainstream economist would interpret this as a paradox—specifically, a *Giffen good*, one that violates the law of demand—whereas the behavioral economist would say that Jim's preferences reversed based on some cognitive bias or heuristic. However, Jim's "reversal" may have nothing to do with preferences at all, but a belief that fair-trade coffee can't possibly do what it's supposed to do—support poor coffee growers—while charging such a low price. And if buying fair-trade coffee no longer serves Jim's principles, and possibly offends them, it makes sense that he switches to buying regular coffee—which he preferred to buy all along based on the lower price but didn't because his principle drove him elsewhere.

In reality, people make choices based on their interests, including principles, ideals, and a myriad of other reasons, all competing with each other until judgment sorts them out—and if their willpower is strong enough to carry through with them. How can economists expect to take a choice motivated by so many different factors and interpret it in terms of one: preference? Economists are seeing real-world choices filtered through their "preferences-only" lens, but the preferences they "see" in decisions reflect complex and multifaceted interests, not just cognitive biases and heuristics. Just like forced-perspective photographs that "show" your Aunt Ethel holding up the Leaning Tower of Pisa because the third dimension is missing—a "surprising" role reversal!—observing real-world choices without the dimension of principles (and other reasons) gives economists a radically distorted view of real-world choice. When you look at it that way, it's surprising preference reversals don't show up more often!

## New Methods, Same Mistakes

In general, despite innovative research and provocative results—especially in recognizing the systematic deviations from textbook rationality that people make every day—behavioral economists do not confront the issues we described with economic models of choice in the last chapter. Although behavioral economics adds realism within the confines of the constrained preference-satisfaction model with regard to the nature of preferences and the way people process information, it does very little to expand the factors people use in choice to include such thing as principles, judgment, and the will. Essentially, behavioral economists still present people's choices as wholly determined by preferences, constraints, and information, just as mainstream economists do. In a tremendously influential law review article, legal scholar Christine Jolls, along with Richard Thaler and Cass Sunstein, list the foundational ideas of behavioral economics as bounded rationality, bounded willpower, and bounded self-interest.[15] The biases and heuristics we detailed above speak primarily to bounded rationality, but how do behavioral economists account for bounded self-interest and bounded willpower? As we'll see below, they do so by tweaking the standard model of constrained preference-satisfaction rather than expanding it, and therefore they perpetuate the shortcomings of mainstream economics (which lead to the ethically questionable policy recommendations that we'll start discussing in the next chapter).

Let's take bounded self-interest first. Adam Smith, known today as the father of economics but regarded in his time as a prominent moral philosopher, wrote extensively in his book *The Theory of Moral Sentiments* of the importance of sympathy and "fellow-feeling," an indispensable supplement to his seminal work on economics, *The Wealth of Nations*.[16] There is a sizable literature in economics that expands the typical self-interested preferences to include altruistic or other-regarding preferences (as we saw in the first chapter), or expanding the notion of utility and preferences to include multiple preference rankings, some based on self-interest and others on altruism or morality.[17] Behavioral economists have also constructed elaborate models to explain altruistic behavior, and have extended them to encompass principles such as fairness and justice, but they still work within the preference-constraints framework. For instance, Matthew Rabin, winner of the 2001 John Bates Clark Medal (the highest award in economics next to the Nobel Prize), developed a model that explains the cooperation and reciprocity that people often exhibit in laboratory experiments but that baffle mainstream economists who assume self-interested behavior.[18] Rabin suggests a "fairness function" that leads to reciprocity: a person will respond to kindness with kindness and selfishness with selfishness. (Think of the last time you dealt with a car salesperson and how you responded to his or her demeanor.) In this model, a person's level of preference-satisfaction or utility responds to the behavior of the other person (not just the effects of his or her action of the first person's outcome), and therefore a person reacts differently in a strategic situation depending on whether the other person plays fair or behaves opportunistically.

However, since a person's preferences for material resources (in other words, stuff) and preferences for kindness all get lumped into the same preference ordering, the person will trade off one for the other in the sense that we criticized in the previous chapter. In fact, Rabin writes that "the bigger the material payoffs, the less the players' behavior reflects their concern for fairness"; in other words, a person's inclination to "indulge" his or her taste for fairness is sensitive to opportunity cost.[19] When it becomes too costly to be fair, people stop being fair (in Rabin's model), which certainly describes *some* people's behavior, but not all. Only recognition of principles that are not sensitive to small changes in opportunity cost—particularly at relatively low levels—can explain why people often return lost wallets and purses, regardless of the amount of money in them (if they even look).

What about weakness of will? As we saw in the last chapter, mainstream economists have no concept of a will or willpower, and neither do behavioral economists, although they certainly recognize the problems that lack of self-control poses for individuals.[20] Behavioral economist George Ainslie, also a renowned psychiatrist, recommends that a person with self-control issues should adopt a *personal rule* to help prevent succumbing to temptation.[21] For instance, if Warren is trying to diet but finds himself at a birthday party for a colleague, he may be tempted to have a slice of cake, "just this once." But Warren works in a large office where every day is *someone's* birthday, so there are many opportunities for cake "just this once." What Warren should do, according to Ainslie, is adopt a personal rule of not eating cake, not even once in a while, based on a cultivated attitude about the consequences of such a choice: a lapse today will lead to a lapse tomorrow, which will lead to one the next day, and so on. If Warren sincerely believes this, he will pass up the cake today in the hopes that it will make it easier to pass up the cake in the future.

This is a brilliant strategy, and is used often by alcoholics who abstain from "just one drink, just this once" out of the fear that a slip today will make slips more likely in the future—and risk the sobriety they've worked so hard to regain. This works if the person loads up the consequences of a drink today with all the future drinks that it will lead to, making a drink today seem extremely costly when a person looks into the future. To put it a different way, a person restructures the costs and benefits of his or her action—or the preferences and outcomes—to make lapses less attractive now. But there is another way to think about this, based on the suggestions we presented in the last chapter, that doesn't require manipulation of costs and benefits: use *willpower* to stick to a *principle* of abstaining from alcohol (or dieting, or exercising, or whatever your goal is).

Admittedly, there is not much difference between my conception of a principle and Ainslie's personal rule; the difference is in how they affect a person's decision-making. A person can certainly think of the consequences of violating a principle or personal rule, just as Ainslie assumes, but he or she can also refuse to consider such things, putting compromise of principle above all thoughts of benefit or advantage. If people regard their principles as integral to their identity as persons, they may be less willing to compromise them than they would be if they considered and accepted the negative consequences of succumbing to temptation—or, more likely, rationalized these consequences away. ("One drink won't hurt," or "this is a special occasion, so I can let my guard down"—especially if the "special occasion" is celebrating a day that ends in "y.") Despite the insights into behavior it offers, as well as the sound practical guidance it provides to those who struggle with self-control, Ainslie's framework of personal rules, as an example of how behavioral economics deals with weakness of will, is based on the same constrained preference-satisfaction model as mainstream economics.

One type of behavior to which Ainslie applies his analysis of self-control is procrastination, a popular area of interest with other behavioral economists as well.[22] In fact, Ainslie is best known for his model of *hyperbolic discounting*, in which our preferences regarding future costs and rewards behave strangely as a person gets closer to them—a model perfectly suited for explaining procrastination, in which we put off boring, distasteful, or difficult tasks against our better judgment. For example, Judy may express a preference to have her root canal on May 10 rather than May 15 when she makes the appointment in February—maybe to get it over with sooner—but as May draws closer, she prefers to have it later, even if she has to live in pain for the extra five days. By the same token, Dave may be willing in February to wait a little longer to receive money in May, especially if it will earn interest while he waits, but grows impatient for it as May approaches and he gets itchy to get his hands on the money.[23] Other behavioral economists offer different explanations of procrastination, such as the salience or visceral nature of present costs ("I can't bear the thought of cleaning the bathroom *now*") or preferences biased toward present outcomes ("I really need to watch the reality dating show I recorded to find out who Damian chose *now*"). What all of these ways of framing the issue have in common is that they are all expressed in terms of "oddly" behaving preferences that lead to a person procrastinating—not anything having to do with willpower.[24]

And this is the problem: as I mentioned in the first chapter, these approaches to modeling weak-willed behavior such as procrastination may provide valuable insights into what factors make procrastination more likely, but they can't explain how or why people *refrain* from procrastination even when those factors are present. Except in cases of chronic procrastination, eventually all of us get around to doing what we have to do—sometimes because the costs of continued delay become larger than the benefits (usually due to impending deadlines), but other times because we simply get up off our ass…umptions and do what has to be done. Behavioral economists' models can't account for this spontaneous exertion of willpower by which we simply decide to bring our actions back into line with our best judgment. A more elaborate model of procrastination would contain a principle of timely performance of a task, which the person would follow if it were not for the temptation of putting the task off. Judy knows she should finish grading exams, but grading exams…well, it sucks. (I'm sorry, but it does. Hug a teacher today.) She'd much rather watch TV, or catch up on Twitter—or even clean her bathroom. Since her willpower is not perfect, she will occasionally lose the battle and put off her grading—and eventually she will return to it, even though grading still sucks and the TV is always there. The models of behavioral economics—just like mainstream economics—cannot explain this because preferences and constraints always pick out the "best action" for a person, which the person follows because he or she *has no choice.*[25]

### One Step Closer

To our benefit, behavioral economists have examined our actual decision-making processes and found systematic, predictable anomalies in the form of cognitive biases and heuristics that were successful evolutionary adaptations once but now hamper effective choice. However, these contributions, insightful as they are, represent intensive adjustments to the traditional model of constrained preference-satisfaction, rather than the extensive improvements necessary to represent the broader factors involved in decision-making. They are valuable contributions, no doubt, and they do add considerably to our understanding of choice, but their value is limited because they involve the same mistakes of omission that mainstream economists made.

Why do both mainstream and behavioral economists get choice so wrong? To be fair, any model that tries to capture some aspect of the real world, whether physical, social, or psychological, by necessity has to leave out some details. However, as these first two chapters explained, the details that economists leave out are not the least relevant ones to real choice, but rather the ones that are most difficult to incorporate in mathematical models—that is, principles whose importance and influence can't be easily measured, judgment between metaphorical apples and oranges, and the indeterminate effects of the will. In theory, you can assign numbers and values to wants and needs to model their effects on a particular choice situation or how much they matter to someone, but it's much harder to do so with concepts and ideals such as honor or love. As a result, economists—both mainstream and behavioral—put the cart before the horse, so to speak: rather than letting reality drive how they build their models, they let the limitations of their chosen modeling approach drive their understanding of reality. They treat choices as determined completely by preferences and constraints, and therefore they fail to consider choice options or processes that don't fit into the modeling framework they've chosen, including the principles and ideas we hold dear, the judgment we need to make hard choices, and the willpower necessary to follow through on them.

Most important for where we're heading in this book, neither mainstream nor behavioral economics consider the wide range of reasons that can motivate people's actions and behavior. Mainstream economists assume fixed and stable preferences, and behavioral economists loosen this up a bit to allow preferences that change or adapt, but neither considers principles or ideals, which make up an important but neglected portion of our overall *interests*. As they incorporate more features and details into the process of decision-making, mainstream economists and especially behavioral economists forget to expand the range of people's motivation at the same time. Instead, they usually restrict them, assuming that people make decisions only to increase their wealth, benefit their health, or some other single and simplistic goal.

It is this assumption regarding people's interests that leads economists to interpret people's behavior in strange ways (like with Jim and his decision to stop buying fair trade coffee after the price fell). If economists assume the only reason people buy lottery tickets is to maximize their wealth, then *of course* it seems irrational to buy them. If economists assume the only reason people eat food is for their health, then *of course* eating donuts seems irrational. But people's interests are much broader and complex than this, which leads to a bottleneck of sorts when economists—especially behavioral economists—try to squeeze all their advanced psychological insight about decision-making through their extremely narrow view of people's interests.

Oh well, at least we can be glad that no one is arguing for laws and regulations based on such problematic models of choice. *Oops.* Remember Richard Thaler and Cass Sunstein? Well, they wrote a book…

# Chapter 3

# How Behavioral Economics Met Law and Economics and Begat *Nudge*

*It didn't take long for the insights of behavioral economics to reach the field of law and economics, in which economic principles and analysis are applied to legal issues, and which is now considered a dominant approach to legal studies in law schools across the country and throughout the world. Law and economics uses economic models of choice to study how laws affect human behavior and societal outcomes (such as accident and crime rates), as well as to recommend changes in laws and legal procedures to influence behavior and improve outcomes. Behavioral law and economics was born when behavioral economists and legal scholars began to criticize the use of the shortsighted standard economic models of choice to study the law. When law and economics is enhanced with behavioral research, the resulting models of choice will predict different behavior when cognitive biases and dysfunctions are accounted for, which will change the models' recommendations for legal policymaking and regulation. For instance, mistaken perceptions of risk will affect people's decision-making in situations such as accident precaution, which affects their reactions to legal standards regarding negligence and liability, and in turn this may suggest changes in the standards themselves.*

*Before long, behavioral law and economics scholars realized that their work applying behavioral insights to laws and* *regulations could be applied, not only to broad societal outcomes, but to individuals' personal well-being as well. In 2008, behavioral economist Richard Thaler and law professor Cass Sunstein published their book* Nudge: Improving Decisions About Health, Wealth, and Happiness, *which popularized their previous academic work on libertarian paternalism and established "nudges" as a viable approach to influencing behavior. They recommend that policymakers take subtle and inexpensive steps toward influencing people's choices in their own interests: crafting default options on forms to prompt better choices, for instance, or more careful arrangement of options on a menu to encourage healthier eating. Ironically, nudges are designed to work with—in fact, they rely upon—the same cognitive biases and dysfunctions that motivated them; they use people's decision-making flaws, which normally work against their interests, to promote them instead. But there are serious informational, ethical, and practical problems with this, which are outlined in the next three chapters.*

* * *

As described in chapter 1, economists have applied their standard model of constrained preference-satisfaction to other fields, such as sociology and political science. But none has been as successful as the economic approach to law, or *law and economics*, which has become not only a large subfield within economics but also a dominant force in law schools and legal scholarship.[1] Law and economics focuses on the role of incentives in influencing the behavior of people in legal contexts, and then uses this information to design legal rules and procedures to maximize the total well-being or welfare of society. Economic reasoning has been applied to topics as varied as the design of liability rules for accidents, the punishments given for different crimes, settlement procedures, and contract remedies—all using the traditional model of choice to explain and predict the behavior of the various actors, such as citizens, police, lawyers, judges, and jurors.

I've long found it ironic that the traditional economic model of choice has been so successfully applied to the study of law, because behavior in legal situations—unlike most consumer choices—transcends mere preferences and constraints to incorporate principles (both personal and social), given the obligations most of us feel to follow the law (in most cases).[2] Of course, there will always be people who consider the law as something to be circumvented whenever it frustrates their goals. Many of us do this occasionally—we may fudge speed limits while driving or exaggerate our tax deductions—but the standard economic model of choice assumes that we always look for ways around the law whenever it stands in the way of satisfying our preferences. In his Nobel Prize acceptance speech, economist Gary Becker, who is responsible for laying the foundation for the economics of crime (as well as other fields), says that he got the idea for "rational criminal choice" after arriving late for a student's oral exam and having to weigh the costs and benefits of parking illegally.[3] The implication of modeling legal decisions this way is that court-mandated fines, damages, or imprisonments are simply the "price" of engaging in behavior that satisfies people's preferences, and the law itself represents no meaningful constraint for them. This is the mind-set of the hypothetical person whom the famous judge Oliver Wendell Holmes called the "bad man," who obeys the law only when it suits him—and who seems very closely related to *homo economicus*, the personification of the standard economic model of self-interested choice.[4]

But not all of us are "bad men" (or bad women)! Most of us, most of the time, obey the law not just when it suits us, but because it's the law. We aren't constantly on the lookout for ways around the law; most of us don't even go through red lights when there is no one around (and no red-light camera to be seen). The mainstream economist may say we have a preference for obeying the law, but as we saw in chapter 1, that leaves too much wiggle room for "acceptable" trade-offs in the interest of preference-satisfaction. Rather, most of the time we treat the law as a constraint, either by observing a general principle to obey the law or recognizing some moral principle at the heart of particular laws. There is no great moral cause behind driving on a particular side of the road, since either side will serve the cause of coordinating traffic and minimizing accidents, but we accept that law as making sense, so we follow it. But the reason most of us don't go around assaulting people isn't because it's illegal, but because it's *wrong*; a good person wouldn't hit people (without a good reason) even if it were legal (or if he or she could get away with it).

If choices that people make in legal contexts are based on principle as well as preferences—and therefore require judgment—the traditional economic choice will have a hard time explaining it. To make matters worse, willpower also plays a large role in such decisions, because many people will be tempted to break the law even when they know and believe it's wrong. It could be a person late for an appointment who considers double-parking to save time, a struggling writer considering omitting some freelancing income from her tax return, or a seasoned burglar trying to reform after being released from prison but having a hard time paying the rent. Cases such as these highlight the shortcoming of the economic model of choice: these people's preferences, based on wants and needs, pull them one way, but their principles pull them in another. The economist would have their preferences determine people's decisions, even though they may wish they didn't—each may like to rise above the pull of their preferences and act out of principle instead. This is difficult, as are all self-control problems, but at the same time all too common, which only emphasizes why recognizing the role of the will is essential in describing conflicting choices such as these.[5]

### Behavioral Economists Take a Shot at Improving Law and Economics . . .

As we saw in the last chapter, behavioral economists work to correct problems with the traditional economic model of decision-making, primarily by focusing on how people form preferences and process information given various cognitive biases and heuristics. They have applied their insights to topics conventionally regarded as "belonging" in economics proper, such as finance and consumer behavior, but also to topics falling outside the traditional purview of economics—including the law. Perhaps out of the same recognition of the ways that the traditional model of choice fails to explain and predict behavior in legal contexts, as well as the fact that the economic analysis of the law is still a relatively new development (and not as firmly established in its traditional form), many legal scholars have welcomed behavioral economics with open arms, leading to the field of *behavioral law and economics*. One legal scholar at the forefront of this movement, Russell Korobkin, proclaimed "victory for behavioral law and economics" in the title of a recent law review article, and asserted in it that "the behavioral economic analysis of law . . . has become *the* economic analysis of law."[6]

One of the foundational papers in behavioral law and economics is "A Behavioral Approach to Law and Economics" by Christine Jolls, Cass Sunstein, and Richard Thaler, which we saw briefly in the last chapter.[7] The paper gives a thorough account of a wide range of legal topics wherein behavioral law and economics explains observed behavior better than traditional law and economics does, including juries' assessment of defendants' negligence, mandatory contract terms, and laws prohibiting victimless crimes. Here's just one example of how behavioral insights can contribute to the understanding of a simple legal situation: the negotiations between private parties in a civil lawsuit after the judge decides for either the plaintiff or the defendant.

A central tenet of law and economics, considered by many to be the concept that launched the field itself, is the *Coase theorem*, based on the work on Nobel laureate economist Ronald H. Coase.[8] The Coase theorem states that under ideal conditions—clearly assigned rights and low costs of negotiation—people will bargain over transferrable rights regardless of a court's decision regarding who is entitled to them. Let's say Daisy really likes listening to heavy metal music in her apartment—as does your author—but her neighbor Jim in the adjacent apartment really likes it quiet so he can read fine books like this one. (What a quandary!) As it happens, Jim likes quiet so much that, after repeated polite requests that Daisy turn her music down, he takes her to court to force her hand. Both Daisy and Jim have preferences over how loudly Daisy plays her music, and they can even put a rough value on it: Daisy would pay up to $200 for the right to play her music loud, and Jim would pay up to $300 to have it quiet.

According to the Coase theorem, if Jim and Daisy can bargain, Jim will end up with the right to enjoy the quiet regardless of what the judge decides. Obviously, if the judge decides for Jim, Daisy will be forced to turn the music down and Jim will have his quiet. But what if the judge decides for Daisy? The judge would grant her the right to play her music, but that right is still worth more to Jim ($300) than to Daisy ($200). If they are both willing and able to bargain, Jim will offer Daisy between $200 and $300 to turn down her music; essentially, he will "buy" the right from her. In cases such as this, the judge's decision is irrelevant to the outcome—Jim will have the right either way—and its only effect is to determine if Jim will have to pay Daisy for it or not. (It works the other way as well, of course: if Daisy valued her right to listen to loud, righteous metal more than Jim did, then she would retain the right if the judge awarded it to her, or "purchase" it from Jim if the judge didn't. Up the irons, Daisy!)

As Coase emphasized, this result will occur only under the ideal conditions of ability and willingness to bargain (which he called "zero transaction costs").[9] There are many ways these ideal conditions could fail to obtain: for example, one or both parties could insist on retaining counsel for the purpose of negotiations, which makes negotiations foolish given the low sums of money they're bargaining over. Or they may be such animosity built up between them—perhaps intermingled with years of unresolved sexual tension!—that each one simply refuses to even *think* about talking to the other.

Behavioral economists might have another explanation, however, based on the endowment effect we summarized in the last chapter.[10] Let's say the judge awards Daisy the right to play loud music, and Jim approaches her with an offer to buy the right from her for between $200 and $300. She may be perfectly willing to listen to the offer, and holds nothing against Jim (even after he dragged her into court). But now that she "owns" the right to her music, her valuation of it may rise from $200 to (say) $400, which is more than Jim is willing to pay. This endowment effect may be enhanced by the fact that she had to go to court to defend it, and her sense of entitlement to it was verified by a judge, which in turn may be reinforced by hindsight bias, leading her to believe she had the right all along (even before the judge officially decided that she did). (This is also consistent with how it would look from the viewpoint of the theory presented in chapter 1: she is standing on the principle that she has the moral right to play loud music and refuses to sell that right for "mere" money.)

Throughout their paper, Jolls, Sunstein, and Thaler explain how insights from behavioral economics can improve the explanatory and predictive power of law and economics, crafting brilliant explanations for anomalies in legal behavior that baffled traditional law and economics scholars. There have been criticisms from within the legal academy, naturally, most of them criticizing the claims of behavioral economics more than the application of their results to legal topics. Richard

Posner, one of the pioneers of law and economics (and also a prominent US federal judge), argues that behavioral law and economics adds little to traditional law and economics scholarship. In a paper responding to Jolls, Sunstein, and Thaler (JST), he writes:

> Some of the insights they ascribe to behavioral economics are already a part of economic analysis of law, which long ago abandoned the model of hyperrational, emotionless, unsocial, supremely egoistic, nonstrategic man (or woman) that JST in places appear to ascribe to it. Other points they make are new labels for old challenges to the economic model of behavior that owe nothing to behavioral economics in any distinctive sense. Others are best explained by reference to evolutionary considerations that play no role in behavioral economics, at least as conceived by JST.[11]

Other critics focus on the disunified nature of behavioral economics, which they argue consists of a scattered set of experimental results that can augment traditional models of choice but cannot yet supplant them.[12] Yet others question the validity of behavioral research itself, arguing that the experimental settings are limited, experiments are contrived, and the subjects (typically college students) are not representative of the population as a whole.[13]

Although I don't want to dismiss these concerns, they aren't particularly relevant for the arguments I'll start making in the next chapter. The problems with behavioral economics—and therefore behavioral law and economics—that were laid out in the last chapter are of a different sort, and we don't need to rehash them here. You can rest assured that the same issues arise when behavioral economics is applied to the law as when it's applied to finance or consumer behavior, because they are foundational problems with the assumptions made by behavioral economics itself. Legal scholar Grant Hayden and philosopher Stephen Ellis, in criticizing the lack of a unified theory of behavioral economics, draw the same conclusion about behavioral and mainstream economics:

> A primary reason for the failure of behavioral economics to confront important questions is its continued unreflective reliance on the basic economic paradigm. Indeed, the usual behavioral methods for accommodating the empirical evidence take the basic economic account as canonical. Accept, reject, or tinker with the functional forms, most standard and behavioral economists confine themselves to thinking about the particular elements of common sense that originally inspired economic models.[14]

Our present concern with behavioral law and economics, one that will remain at the forefront for the rest of this book, is not with what its results are, but rather with what its proponents want to *do* with them.

### ...But Did They Overshoot Their Target?

At the end of their paper, Jolls, Sunstein, Thaler discuss a final implication of behavioral law and economics: *anti-antipaternalism*, "a skepticism about antipaternalism, but not an affirmative defense of paternalism."[15] Mainstream law and economics scholars, like economists in general, normally frown on paternalism—laws and regulations that aim at influencing people's behavior for their own good (as opposed to the good of others). They usually maintain that

1. people know their own preferences better than anyone else (including the government),
2. people make choices that maximize their level of preference-satisfaction, and
3. people's well-being is based on their level of preference-satisfaction.

Therefore, people will generally maximize their well-being by making their own choices based on the preferences that they know best. Economists refer to this position as *consumer sovereignty*, and traditionally it has restricted economists' analysis and recommendations regarding government policy to maximizing aggregate welfare or well-being (normally through fiscal and monetary policy and regulatory interventions). Similarly, law and economics scholars focus on laws that enable people to satisfy their own preferences with the law acting in either a positive capacity (such as enforcing contracts) or a negative capacity (such as protecting against harm from others).

But according to Jolls and her two colleagues—from whom we'll hear more very soon—the work of behavioral economists casts doubt on all three positions that mainstream and legal economists rely upon to oppose paternalism. As we saw in the last chapter:

1. Behavioral economists maintain that preferences are constructed at the time a choice has to be made, so there are no preexisting preferences to be "known."
2. Even if preferences were given and stable, the heuristics or cognitive shortcuts that people use to make decisions may lead to decisions that don't reliably maximize their level of preference-satisfaction.
3. Although we know that not all preferences increase people's well-being—such as self-destructive and other-regarding preferences—behavioral economists argue that people's cognitive biases may warp their "normal" preferences enough that their link to well-being is further weakened.[16]

For these reasons, behavioral economists argue that there is little reason to believe that people reliably make choices in their own well-being. They suggest that the government may be able to influence people's decision-making to help them make better choices in their own interests—in other words, paternalism may be justified by the findings of behavioral economists. (We'll have much more to say about paternalism in general in chapter 5.)

Jolls, Sunstein, and Thaler are not the only scholars to see the possible implications of behavioral research for the future of paternalism. As Ted O'Donoghue and Matthew Rabin, whom we heard from in the last chapter on the issue of procrastination, write,

> The classical economic approach to policy analysis assumes that people always respond optimally to the costs and benefits of their available choices. A great deal of evidence suggests, however, that in some contexts people make errors that lead them not to behave in their own best interests. Economic policy prescriptions might change once we recognize that humans are humanly rational rather than superhumanly rational, and in particular it may be fruitful for economists to study the possible advantages of paternalistic policies that help people make better choices.[17]

Their paper proposes changing the tax rates on different types of goods (while keeping the overall level of taxes constant), increasing taxes on goods that people overconsume due to self-control problems (such as potato chips) and lowering taxes on other goods (such as carrots). If successful, this would help consumers eat fewer potato chips and more carrots, as well as lower the tax burden on consumers already eating mostly carrots.

In a law review paper provocatively titled "Regulation for Conservatives," behavioral economist Colin Camerer, law professor Samuel Issacharoff, and economist/psychologist George Loewenstein, along with O'Donoghue and Rabin, make the same case for paternalism (with an important reservation):

> Paternalism treads on consumer sovereignty by forcing, or preventing, choices for the individual's own good, much as when parents limit their child's freedom to skip school or eat candy

> for dinner. Recent research in behavioral economics has identified a variety of decision-making errors that may expand the scope of paternalistic regulation. To the extent that the errors identified by behavioral research lead people not to behave in their own best interests, paternalism may prove useful. But, to the extent that paternalism prevents people from behaving in their own best interests, paternalism may prove costly.[18]

This reservation in particular will be echoed and expanded on later in this book; however, they have others as well:

> First, while research in behavioral economics documents common mistakes, those mistakes are typically far from universal, and we worry that paternalistic policies may impose undue burdens on those people who are behaving rationally in a particular situation. Second, behavioral economics is in an early stage of development, and therefore its findings should elicit more caution than those from more "mature" fields (which are by no means themselves invulnerable to revision).[19]

Although they aim to convince those "prone to rigid antipaternalism" of the implications of behavioral economics that make paternalism seem more reasonable, Camerer and his colleagues admit to "trepidations" about using behavioral economics to justify paternalism and want to caution those "prone to give unabashed support for paternalistic policies based on behavioral economics" to tread carefully.[20]

Jolls, Sunstein, and Thaler are also careful to qualify their endorsement of paternalism, focusing on the possibility of "behavioral bureaucrats" who are subject to the same cognitive dysfunctions as the people they claim to help make decisions.[21] People who work in government are no less subject to cognitive biases and heuristics than anyone else—no jokes, please—and those anomalies in decision-making are just as likely to influence their choices regarding law, policies, and regulations as they are to affect their decisions as consumers and investors.

Although regulators are distanced from the choice situations of ordinary citizens, and may be able to see the problems in their decision-making more clearly—like the bartender who can tell the lush he has had too much to drink—they are nonetheless subject to their own flawed thinking in crafting the paternalistic regulations themselves. The school of economics known as *public choice* models government decision-making as the result of the interaction of individual decision-makers, each acting in his or her own interests (whether or not they're aligned with the public's interests). Looking at paternalistic policy through the lens of public choice, and imagining the impact of cognitive biases and heuristics on regulator's own choices, may temper behavioral economists' enthusiasm for paternalistic laws and policies. As O'Donoghue and Rabin recognize, any paternalistic policies must be designed to craft the best overall results, rather than indulge private interests and enable "promiscuous paternalism by public or private entities."[22]

Overall, most considerations of paternalism on behalf of behavioral economists are cautious. In addition to Camerer and his colleagues' general skeptical approach quoted earlier, Jolls, Sunstein, and Thaler write that "from the perspective of behavioral economics, issues of paternalism are to a significant degree empirical questions, not questions to be answered on an a priori basis."[23] Similarly, O'Donoghue and Rabin write that "in some instances it will surely turn out that paternalistic policies do more harm than good (although even here our approach forces us to draw such conclusions through explicit analysis rather than a priori assumption)."[24] (Note, however, that I did say "most.").

## Speaking of Which...

Although their work with Christine Jolls that was largely responsible for starting the behavioral law and economics revolution was very cautious about endorsing paternalism based on its findings, Richard Thaler and Cass Sunstein seem to have fewer reservations when writing on their own. In several academic journal and law review articles, as well as their popular book *Nudge*, Thaler and Sunstein advocate forcefully for what they call *libertarian paternalism*.[25] (This would seem to be the very "affirmative defense of paternalism" that they shied away from when writing with Jolls.) As they explain in the book:

> It is legitimate . . . 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.[26]

They refer to their brand of paternalism as "libertarian" because "choices are not blocked, fenced off, or significantly burdened," as opposed to paternalistic laws that require certain behavior (such as using seat belts or wearing motorcycle helmets) or forbid it (such as recreational drugs), which are usually considered the antithesis of libertarianism (or classic liberalism).[27]

Instead, Thaler and Sunstein recommend that laws and regulations be designed to "nudge" people toward making certain decisions rather than others. They define a *nudge* as "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."[28] *Choice architecture* refers to the way options such as food choices in a cafeteria or names on a ballot are presented; due to cognitive biases and heuristics (such as the framing effect), the choice architecture can have a significant effect on the choices people make. In other words, Thaler and Sunstein argue that by simply thinking about the way options are presented and rearranging them, people can be guided into making better choices—choices people themselves would like to have made, were it not for their cognitive biases and heuristics—without taking away their freedom of choice.

Thaler and Sunstein spend about half of *Nudge* providing practical examples of libertarian paternalism.[29] For instance, Chapter 6 of their book describes nudges that private companies have implemented (sometimes in cooperation with the government) to increase savings among their employees. First, new employees are automatically enrolled in a 401(k) program with an option to decline, as opposed to the standard process in which you are given the option to enroll in a 401(k) plan but are not enrolled if you do nothing. Thaler and Sunstein consider passing up on a 401(k) plan to be "foolish without a doubt," a result of simple laziness or procrastination, so they recommend automatic enrollment in the hopes of increasing participation, which they assume new employees would prefer if they thought carefully about it.[30] Second, once enrolled, the "Save More Tomorrow" program corrects for failures to increase the contribution rate above the default when employees' wages or salaries go up, again out of laziness or procrastination ("I'll change my rate tomorrow, honest"). Under the program, contribution rates rise in a way that preserves their take-home pay (or lets it increase, albeit not as much as it would if savings had not increased), once again under the assumption that this is what employees would have chosen if they thought about it.

Another example that Sunstein and Thaler often provide deals with a cafeteria director, named Carolyn in some versions of the example, who (among other things) determines the arrangement of food and beverage items on display.[31] Carolyn can arrange the items any way she chooses, but since she has read all of the behavioral economics research, she knows that the way she arranges the items will have an effect on the choices her customers make. For instance, items that are better lit, more attractively presented, or simply easier to reach, will be chosen more often, regardless of her customers' "true" preferences. She also knows that consumers' true preferences may not be good for them anyway (if they have stable preferences at all), and she assumes that they would really like to eat better if only they could. With this in mind, she sees no other option but to arrange the items in her cafeteria so that her customers will be more likely to select healthy options; for instance, she puts the desserts on the top shelf in dim lighting and has the salads well-lit and placed right in front of them. Since she has to arrange the items in her cafeteria somehow, she reasons—as do Thaler and Sunstein—that she may as well arrange them in the interests of her customers' health. (In fact, she may regard this as a moral imperative, given that she has full authority over cafeteria organization and her customers' health is affected by the choices she makes.)

Despite some controversy in academic circles (which we'll talk about in the next few chapters), the idea of nudges has been widely influential among government policymakers on both sides of the pond. With help from Richard Thaler, English prime minister David Cameron instituted a Behavioral Insight Team—nicknamed the "nudge unit"—within his government to look into ways to influence behavior according to libertarian paternalist principles. The Behavioral Insight Team started with excessive food and drink consumption and then moved on to smoking, teenage pregnancy, charity donations, the environment, and personal banking.[32] American president Barack Obama is also a fan of *Nudge*: he appointed Cass Sunstein to head his Office of Information and Regulatory Affairs, where Sunstein implemented regulations consistent with libertarian paternalism in areas as diverse as pollution, automobile fuel economy, and yes, retirement accounts.[33] Especially in these times of debt crises, governments everywhere are looking for way to regulate behavior—paternalistically or not—that don't require much money, and nudges, almost by definition, fit that bill.

### The Problem with Nudges, in Three Acts

You probably noticed that my summary of libertarian paternalism and the idea of nudges here is brief, and you're right.

The idea is simple—and this is one of its virtues—but the issues surrounding it are not. I will deal with these issues in the rest of the book, citing not only Thaler and Sunstein themselves, but other academics on both sides of the debate. In this chapter, I meant to give a neutral account of libertarian paternalism, and there isn't much more to say without getting into issues that demand scrutiny.

In the following chapters, I argue that libertarian paternalism and nudges are ineffective, unethical, and counterproductive. At the most basic level, the problems with libertarian paternalism stem from the stance its proponents take toward people and the choices they make. For instance, arguing against their critics, Thaler and Sunstein write:

> 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 than the choices that would be made by someone else. We claim that this assumption is false—indeed, *obviously* false. In fact, we do not think that *anyone* believes it on reflection.[34]

Well, they're certainly wrong about the last part, because as it turns out, I believe firmly—"on reflection," even—that most people, most of the time, do make decisions in their best interests.[35] And even when they don't, there is no way for anyone else to know this—only the decision-makers themselves can know if their choices do not serve their own interests because *only they know their true interests.* This position is based on the more comprehensive understanding of choice and interests described in the first two chapters of this book, plus some basic ethical principles that I'll introduce soon, all of which I'll apply to the arguments for libertarian paternalism—as soon as you turn the page. (Nudge, nudge.)

# Chapter 4

# Why Nudges Can't Do What They Promise

*Countless times every day, people around the world make choices in their own interests, including their wants and needs, goals and dreams, and principles and ideals, all of which are incredibly complex and multifaceted. But behavioral economists question many choices made by ordinary people because they seem to contradict the simplistic interests they assume people have, such as wealth or health. For instance, people buy lottery tickets, even though the chance of winning is astronomically small. But no one buys lottery tickets because he or she thinks it is a prudent financial decision—they buy them because it's exciting to watch the balls drop and imagine the joy of winning. Likewise, people don't eat unhealthy foods just because they don't know better or have self-control issues, but they may have other reasons to eat them—reasons that may be suspect according to a health economist but not to the people making that choice. To them, it may have been a great choice, fully in their interests, regardless of what anyone else thinks of it.*

*Libertarian paternalists claim to nudge people to make the decisions people themselves would like to make, but this is impossible unless they know these people's true interests—and they can't. Instead, regulators point to the success of nudges in generating the choices they were designed to generate as evidence that they advance people's true interests. If people are nudged* *into eating more vegetables, this is taken as proof that the nudge worked, because regulators simply assume people wanted to eat more vegetables all along but somehow "couldn't" because of some cognitive defect. As this chapter explains, the only way to get even a glimpse into a person's true interest is to observe his or her choices absent any outside manipulation (or nudge). Even then, there are countless ways to explain any choice—desires, needs, principles, ideals, and so on—which reinforces the point that only a person can know his or her true interests. Nudges simply can't do what its proponents claim. If they claim nudges are effective, one has to wonder: effective at what?*

* * *

Every Sunday morning, Patrick, a heavyset man in his early 50s, takes a stroll to the local bakery and enjoys a muffin and coffee while he reads his newspaper. If a behavioral economist happened to walk into the same bakery on a Sunday morning and saw Patrick, he or she may think, "that man shouldn't be eating a muffin—he should have chosen the fruit cup instead." To be fair, that's not an uncommon reaction to seeing a person do something that seems harmful to himself or herself. I'm sure many of us, when we see people eating fast food, getting drunk, or smoking, think to ourselves, "they really shouldn't do that; it's not good for them." At best, this attitude shows concern for your fellow human beings; at worst, it's presumptuous ridicule, even if it goes unspoken.

But the behavioral economist would likely go one step farther, asking, "but why is he eating that muffin? He must be weak-willed, or perhaps he succumbed to crafty marketing or presentation on the part of the bakery or the food industry. Maybe he was confused about the nutritional content of the muffin compared to the fruit cup." (Of course, the behavioral economist chose the fruit cup.) The coup de grace, of course, is the final sentiment: "If only there were some way to help him make better choices in the future."

What the behavioral economists doesn't know (nor do we) is why Patrick is eating that muffin on a Sunday morning. When Patrick was a young boy, his grandfather would take him to that same bakery every Sunday morning and buy two muffins, one for each of them. Patrick would savor his treat while his grandfather told him stories of his experiences in World War II, how he met Patrick's grandmother (who passed when Patrick was eight), and what Patrick's father was like as a young boy. Patrick lost touch with his grandfather after he left for college, and was able to share a muffin with him just one last time, at that very bakery, before his grandfather died three years ago. So every Sunday since, he goes to that bakery, buys a muffin, and thinks about those mornings with his grandfather.

Or maybe not. Maybe there's another reason he's eating a muffin that morning. Maybe it's the first muffin he's had in a year, which he's using to celebrate his weight loss over the past 12 months during which he abstained completely from muffins and other pastries. (Fruit cups every day!) Maybe he just had a fight with his wife, who hates that he eats muffins, so he's enjoying it out of spite. Maybe he's sampling the muffin because he writes food reviews for the local paper (the same paper he's reading). Maybe he has a crush on one of the employees at the bakery and is eating the muffin to avoid looking conspicuous. Maybe he uses the muffin as incentive to walk in the morning (acting to counter his own weakness of will, however self-defeating the reward may seem). Or maybe he simply loves the taste—shocking, I know—and he felt like eating something that tastes good despite the fact that it's not a "part of this nutritious breakfast."

There are countless different reasons and motivations that could have led Patrick to make the choice to eat a muffin that Sunday morning. We might not agree that all of them are good reasons, and Patrick himself might not agree that all of them are good reasons. But we can safely presume that he had a reason to eat that muffin and that his choice serves his interest based on that reason. But all the behavioral economist sees is an overweight man eating a fattening treat, and based on this he or she concludes that "this man is acting against his interest in his health, and we should help him to improve his choices, to further his own interest."

This example highlights the two most significant problems with libertarian paternalism and nudges. First, there is no way for an outside observer to know what a person's interests are that drive his or her choices. A person's interests can only be inferred from his or her choices, and even then there is no way to determine what interests drove a particular decision. Therefore, if the observer wants to judge the person's choice, he or she has to use standards other than the person's own interests. Second, there is no justification for any policymaker to try to "nudge" a person's choice, even if the person's interests were known, but especially not when the interests promoted are imposed on the person by the policymaker. In this chapter we'll focus on the first problem, that of knowing a person's true interests, and we'll tackle the justification problem in the next chapter (although they're very closely related).

## Choices and Interests

As we saw in chapter 1, people have a wide range of interests—things and people and ideals that they care about—that motivate their choices. Notwithstanding the limitations of the models of decision-making used by both mainstream and behavioral economists, interests aren't based on preferences alone but also include principles and ideals, whether personal or general. Interests include everything that a person cares about (in a positive or negative sense) and all the reasons a person makes choices and takes action, and are an important part of the person himself or herself.

How can we determine a person's interests? To put it simply, we can't, if by determining them we mean knowing them with any degree of certainty. Each person's interests are internal to that person, as are his or her thoughts, beliefs, and desires; the only glimpse the outside world gets of them is given by the person's choices (unless the person expresses them outright). But any choice has countless possible explanations, as we saw with Patrick and his Sunday muffin. Even if some are more likely than others, we cannot know for certain which of the possible or probable interests was the one that motivated a particular choice. And there isn't necessarily *one* interest that led to a certain decision—there may be several. Nancy may stop for coffee at a certain coffee shop on the way home for work because (1) it was convenient, (2) it has the best coffee, (3) she often runs into friends there, and (4) she heard a movie producer stops by for coffee once in a while and if she meets him she may have a shot at her dream of becoming a Hollywood movie star and escape her humdrum life making copies of flyers for her neighbors' rent parties.[1] Each of these reasons represents an interest, and any one of them alone could explain her choice, but even she may not be able to identify one "controlling" interest because all of them may influenced her choice.[2]

But this only poses a problem *if* one has an interest in determining other people's interests, which economists do. Economists want to explain and predict people's behavior, but an essential part of that is determining a person's goal, what he or she strives for and make choices in pursuit of. Economists typically assume that people make choices in their self-interest—not because they think people always do act in their self-interest, but because it's the most reasonable assumption to make if you have to assume a single goal. This works fairly well for consumers trying to get the best deal for their money and businesses trying to earn a profit—especially corporations, wherein the directors have a fiduciary duty to maximize the stock value—but not so well for other roles people play, such as spouse, parent, teacher, doctor, or elected official. In these roles, people often pursue other goals, such as caring for others or performing a duty. But once you relax the self-interest assumption—one virtue of which is its simplicity—it's much more difficult to decide what to add to it to make explanations and predictions more realistic. Even if the economist adds simple altruism to the model, he or she must decide to whom people are altruistic, to what degree they're altruistic, and how they balance their altruism with their self-interest. This would be a more realistic model, to be sure, but much harder to specify. This is why self-interest is much easier to assume when modeling behavior, but it comes at the cost of poor explanations and predictions in situations in which altruism is obviously present.

As we explained at the end of chapter 2, behavioral economists have contributed immensely to describing our decision-making processes, but they have the same overly simplistic view of people's interests, taking them to be preferences—which they then doubt the validity and reliability of. Based on this view of people's interests, behavioral economists argue that people often make poor decisions because of their cognitive biases and heuristics. Furthermore, they claim that these decisions are poor in the judgment of the person making the decision. For instance, philosopher J. D. Trout, an advocate of what he calls "bias-harnessing" measures (similar to nudges), writes that "regulation can be permissible even when it runs counter to that person's spontaneous wishes, particularly when the regulation advances the agent's considered judgments or implicit long-term goals."[3] But there is no way regulators can draw that conclusion without knowing what the person's "implicit long-term goals" are, and the only way to tell anything about them is to observe his or her choices—the very choices that behavioral economists are assuming are not in people's true interests!

Not only do behavioral economists maintain that people do not make choices reliably in their best interests—even though there is no way to determine what people's true interests are—but they mock the position traditionally held by economists and legal scholars alike that people do. According to Ted O'Donoghue and Matthew Rabin:

> Economists will and should be ignored if we continue to insist that it is axiomatic that constantly trading stocks or accumulating consumer debt or becoming a heroin addict must be optimal for the people doing these things merely because they have chosen to do it.[4]

As we saw in the last chapter, Christine Jolls, Cass Sunstein, and Richard Thaler also use the term "axiom" to describe the assumption that people makes choices in their interests: "No axiom demonstrates that people make choices that serve their best interests; this is a question to be answered based on evidence."[5] Colin Camerer and his coauthors quote this passage from Jolls, Sunstein, and Thaler, and then add: "Of course, to the extent that faith-based antipaternalism practiced by some legal scholars rests on such an axiom, scientific debate will be unproductive."[6]

If there is such an "axiom"—a central assumption about people's choices in relation to their interests with which we start our "scientific debate"—it is based on both respect for persons (an *ethical* principle) and ignorance regarding other people's interests (an information-related or *epistemic* principle). Given that we have no way of knowing a person's true interests or the motivations behind a particular choice, we have two options: (1) we can either grant the person the benefit of the doubt and assume that person made the choice in his or her best interests (since we have no basis to say otherwise), or (2) we can question the person's choice, which requires that we make assumptions about his or her interests that are unrelated to the choices we observe.

To use the example of Patrick and his Sunday muffin, since we don't know for certain why he chose to eat the muffin, we can either assume he had a reason to eat the muffin that served his interests (unless he tells us differently), or we can assume it was a bad choice—but based on what? Jolls, Sunstein, and Thaler consider this "a question to be answered based on evidence," but what evidence do we have except *the choice itself*? Legal scholar Claire Hill states this point very well when she writes that proponents of libertarian paternalism

> sometimes speak as though they have access to the knowledge of what people really want apart from what they choose. This position is ultimately untenable.... As convenient and tempting as it may be to extrapolate from our own introspection that others want what we do, or should, want, we simply have no access to others' beliefs and desires.[7]

Interests are internal and subjective and therefore cannot be known by anyone else—unless a person reveals them through choice or expression. To maintain otherwise is sheer disrespect.

To their credit, behavioral economists and libertarian paternalists do recognize that they do not always—do not *always*—have enough information to judge people's decisions, but they don't draw the conclusions that follow from that ignorance and that are consistent with respect for choice and autonomy. For instance, Sunstein and Thaler write that, in cases of regulator ignorance, "the committed anti-paternalist might say...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."[8] Camerer and his colleagues grant that "we must carefully address whether patterns of apparently irrational behavior are mistakes or expressions of stable preference," and give the example of extended warranties, which are widely regarded to be a bad purchase motivated by risk aversion, the availability heuristic, and framing effects. Nonetheless, they admit, "if informed consumers continue to purchase the warranties, then it is quite possible that they have good reason to do so, however *unfathomable* that decision may seem to an economist."[9] I would suggest that, to the extent such a choice is regarded as "unfathomable," it reflects more on the economist than on the consumer. Is it really that difficult for a behavioral economist to imagine reasons why someone would make a different choice than he or she would?

It is one thing for the person to admit his or her decision was bad. We all make dumb choices from time to time—but *no one can possibly know our choices are bad except us.* Maybe Sally had a cigarette today, even though she quit three days ago, and she wishes she hadn't. Sally can admit she made a bad decision, a dumb choice, because she knows her own interests. She knows she wanted to quit, she knows having a cigarette would frustrate that goal, but for some reason she had a cigarette anyway. But unless behavioral economists *know* that about Sally, and unless they *know* that she had no other interest competing with her goal to stop smoking that could give her a reason to smoke again, then they *cannot know* that her choice was a bad one from her point of view. They simply can't.[10]

### Whose Interests Are They, Anyway?

But if behavioral economists can't base their judgment of bad decision-making on people's actual interests, what do they base them on? On what grounds do they say that Patrick's choice to enjoy a muffin on Sunday morning and Sally's choice to smoke a cigarette are bad ones? Absent any information about people's true interests (aside from the choices they make), behavioral economists have to impose *some* interests, compared to which they are able to judge people's choices. As philosopher Dan Brock writes:

> Paternalistic interference involves the claim of one person to know better what is good for another person than that other person him- or herself does. It involves the substitution by

> the paternalistic interferer of his or her conception of what is good for another for that other's own conception of his or her good. If this involves a claim to know the objectively correct conception of another's good—what ultimate values and aims define another competent individual's good, independent of whether that other accepts them—then it is ethically problematic.[11]

This process of *value substitution* negates any claim on the part of behavioral economists and policymakers that they are nudging people in their own interests. Instead, they are nudging people towards decisions to advance interests that were chosen for them, compounding the disrespect of ignoring people's true interests.[12]

Take the nudges designed to increase participation in 401(k) programs that we discussed in the last chapter. The regulators who designed these programs assumed that it is in the employees' interest—their sole interest—to have more retirement savings available when they stop working. On the surface this seems reasonable: anyone would rather have more retirement savings than less, all else the same. But all else is not the same, because increased retirement savings are not the only interest employees have. They may need money to put a down payment on a house or apartment, pay for a wedding, or support partners, children, or parents. New employees, who typically start at lower salaries, may have more pressing needs for this money than higher-paid employees with more seniority do. Or they may simply want to have some fun with the extra money. But those who would nudge employees into retirement programs consider only one interest, an interest in retirement savings, which is a reasonable assumption to make for the purpose of social science but is a gross generalization and oversimplification when it comes to policy.

Defenders of policies such as Save More Tomorrow argue that their nudges do not foreclose any options, leaving employees free to opt out of automatic enrollment or increases in savings rates if they choose. According to Camerer and his colleagues, a nudge (or, in their terms, asymmetric paternalism)

> creates large benefits for those who make errors, while imposing little or no harm on those who are fully rational. Such regulations are relatively harmless to those who reliably make decisions in their best interest, while at the same time advantageous to those making suboptimal choices.[13]

But behavioral economics and libertarian paternalism are based on people's supposed inability to make good choices in their own interests due to cognitive biases and heuristics. Nudges take advantage of these same cognitive defects to steer them into decisions that are judged to be better, regardless of any reasons they have that would lead to a different choice. Proponents of libertarian paternalism can't have it both ways, claiming that people will be nudged into better decisions when they would have made poor ones, but that they won't succumb to nudges when they have "good" reasons to make a different choice. If behavioral economists are correct and "choice architecture" has such a powerful effect on behavior, they will affect those who would have made the "bad" choice, no matter how good their reasons for it might have been.

But in the end this point is moot, because the success of nudges is measured by how much of the desired behavior they elicit, based on the assumption that this behavior is actually in people's best interests and represents the choices they would have made if only they could. The argument for automatic enrollment in 401(k) programs goes like this: (1) regulators assume retirement saving is the employees' sole interest, (2) regulators nudge them into enrolling, and (3) the high rate of enrollment is taken as evidence that employees ultimately made the "right choice" and furthered their true interest (retirement savings). As Sunstein and Thaler write:

> 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 . . . then by choosing automatic enrollment, they are acting paternalistically by our definition of the term . . . steer[ing] employees' choices in directions that will, in the view of employers, promote employees' welfare.[14]

This is what philosophers call "begging the question," assuming the results that you claim to have argued for. Libertarian paternalists treat the intended effect of nudges as an independent basis for judging their success, but that's rigging the game. All the higher enrollment rates tell us is that the nudges successfully steered employees' cognitive biases and heuristics toward enrollment rather than nonenrollment, *not* that this furthered their true interests, which were never a factor in the design of these nudges.

So Camerer and his colleagues' statement earlier is true if we understand the phrase "those who reliably make decisions in their best interest" to mean those who make the choices *we* want them to make. If people were already going to make the "right" choice, then the nudges have no effect—they only affect those who would have made the "wrong" choice, regardless of what reasons they had for making it. Sunstein and Thaler are much more blunt when they make the same point: "In its most cautious forms, libertarian paternalism imposes trivial costs on those who seek to depart from *the planner's preferred option*."[15] (The careful reader will no doubt have noticed also that the previous quote from Sunstein and Thaler mentions the viewpoint of *employers*, not employees, in judging employees' welfare.)

In addition to being imposed, the interests assumed by behavioral economists on behalf of the people being influenced are overly simplistic on several different levels. They only deal with one interest at a time, such as wealth or health, whereas we know that people have a myriad of interests, any number of which can affect even the most insignificant decisions (such as Patrick's Sunday muffin). Furthermore, these various interests must be balanced using judgment, which economic models are simply not equipped to deal with (as we saw in chapter 1). The models of behavioral economists are simplistic also in that they consider their single interest along just one dimension. This may not be a significant problem if you assume people want to maximize their wealth, since different forms of wealth can be measured in the same unit of money. However, health is more complicated, consisting of many dimensions: one person may be concerned about her strength but not her cardiovascular fitness while another has the opposite priorities, or one person smokes but carefully watches his weight and another vice versa. When regulators target health, they focus on the dimension of it affected by the choices in question: smoking most directly affect the lungs, overeating most directly affects weight, and so forth. No responsible physician would focus on just one dimension of his or her patient's health for fear of neglecting others. The simplistic interests assumed by behavioral economists are breeding grounds for unintended consequences: nudges that get people to smoke less may prompt them to eat more, "requiring" another nudge designed to change that![16]

But which interests or values are substituted for the person's actual ones? Are there obvious, neutral values that the behavioral economist or policymaker can just pull off the shelf and impose on the actual decision-maker? The range of interests a person can have and act upon is infinite, but somehow, behavioral economists can determine that retirement savings or health are a person's sole "real" interest. But how do they decide that? That determination is an act of judgment itself: the regulators are substituting *their* judgment regarding another person's best interests for the person's own. In defending their automatic 401(k) enrollment program, Sunstein and Thaler can't claim some objective basis for emphasizing retirement savings above all else. That's merely their opinion, as if they were to say: "You don't care enough about your retirement savings, but *we think* you should." Camerer and his colleagues make a similar point with respect to health: "Health and food regulations are heavily informed by scientific understanding . . . and by a widespread belief among professionals that average folks require information, prodding, and often regulation to improve their health and diet."[17] Value substitution does not simply involve imposition of interests, but an usurping of judgment about the way individuals structure and run their lives. When a regulator replaces someone's values and interests, that regulator is not replacing them with some obvious or neutral interests, but with his or her own.

Value substitution is a legacy of the way that decision-making is modeled in both behavioral economics and mainstream economics, where simple interests (such as self-interest or profit-maximization) are assumed and then choices are evaluated based on how well these interests are promoted. As I said earlier, this is problematic enough in purely theoretical work that attempts to explain and predict behavior (as we saw in the first two chapters), but it raises significant ethical issues when used in a policy context. It is one thing to build a model that assumes a single, simple interest to predict people's behavior, but another thing entirely to make the same assumption when designing ways to manipulate people's behavior in those interests—and then, adding insult to injury, to claim that they are people's actual interests.

### "Don't Worry, We Knew Your Interests Better Than You Do!"

As we saw earlier, behavioral economists and libertarian paternalists often make no pretense to nudge people in their own interests—and this is natural, given both their reliance on preferences to ground interests and their skepticism about preferences themselves! To an economist, remember, a person's interests are completely described by his or her preferences, since there is no room for including principles in either mainstream or behavioral economics. Although a core assumption of mainstream economists is the stability of preferences, behavioral economists question this, considering preferences to be constructed and malleable—or *endogenous*, which means they can be formed within a model of choice rather than being a fixed assumption. As Sunstein and Thaler write: "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."[18]

If you base your entire theory of choice on preferences, as mainstream and behavioral economists do, this poses a problem—and presents an opportunity. The problem is that there is literally no way to know a person's "true" preferences if his or her preferences can be changed or manipulated. This issues comes up often in economic studies of advertising, in which persuasive advertising can be seen as illegitimately strengthening a person's preferences for a product. Imagine that Jackie sees a commercial on TV for a new soda, and this "makes" her want the soda more than before. After she buys it and drinks it, her new, stronger preferences for the soda have been satisfied—but is she better off simply because her satisfied preferences, which were manipulated by the commercial, are now stronger than they used to be? Or should we measure her well-being according to her original preferences—should those be considered more authentic?[19] The same idea applies to anything that takes time to appreciate and develop a preference for, such as an unfamiliar type of music or a new sport, as well as addiction: obviously addiction creates stronger preferences for a good, but does satisfying these new, "enhanced" preferences make the person better off?[20]

These examples show the problem with endogenous preferences: without a foundation of stable preferences, economists have no basis on which to judge what makes people better off or worse off. Preferences are commonly used by economists to measure individuals' well-being, a practice that was already threatened by self-destructive and other-regarding preferences (as we saw in chapter 1). If we acknowledge the existence of preferences that change in response to each choice situation (as behavioral economists claim), there seems to be no way to preserve a useful link between preferences and well-being. But *this* is where the opportunity arises! Recognizing the difficulty with using people's own preferences to measure well-being, economists can—nay, they *must*—substitute another set of preferences, values, or interests for people's unstable and unreliable preferences, which can't be presumed to correspond to their well-being.[21] And that brings us back to value substitution, but without any expectation that economists should know people's preferences or respect them as shown by choice—because now they know how to make people's preferences *better*.

With a little help from philosophers—of course, *now* they listen to them—economists have an elegant solution that deals with the problems of preferences themselves. If preferences have been warped or distorted by cognitive biases, heuristics, or mistaken information so that they no longer "track" with a person's well-being, then they simply need to be fixed, adjusted, or reshaped to bring them in line with what a person *would* prefer *if* he or she weren't under their influence and could simply reflect on what is best for him or her. Economists and philosophers refer to these "corrected" preferences as a person's *rational preferences* (based on *informed desires*), which are designed to be more closely linked to a person's well-being than his or her actual preferences are.

Sunstein and Thaler invoke this concept when they "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."[22] As explained by Nobel-winning economist John Harsanyi:

> [A person's] manifest preferences are his actual preferences as manifested by his observed behavior, including preferences possibly based on erroneous factual beliefs, or on careless logical analysis, or on strong emotions that at the moment greatly hinder rational choice. In contrast, a person's true preferences are the preference he *would* have if he had all the relevant factual information, always reasoned with the greatest possible care, and were in a state of mind most conducive to rational choice.[23]

For example, our friend Patrick may experience a preference for muffins on Sunday morning and act on it. But the behavioral economist who sees him may conclude that he would prefer *not* to eat that muffin *if* he thought about it more, if he only took the time to reflect on how it will affect his long-term well-being, in the absence of any cognitive biases, mistaken information, or distracting sentimentality. Since Patrick is obviously not in a position to undergo this open-eyed, dispassionate, and rational reflection as he stands in front of the muffin display, the behavioral economist takes it upon himself or herself to do it for him.[24]

The problems with (so-called) rational preferences are many. Most basically, there is the logical impossibility of knowing what a person's preferences would be *if only* he or she had, well, different preferences. In philosophy this is called a *counterfactual*, a "what if" question, like "What if Germany had won World War II?" or "What if Sunstein and Thaler had written a book titled *Fudge*?"[25] Counterfactuals are notoriously difficult to determine or evaluate because they are literally "against facts" or reality. Also, why should we require that all preferences be formed, and decisions be made, under ideal conditions—conditions that someone else determined are rational (after deciding that our existing preferences are not)? As economist Robert Sugden asks, in response to Sunstein and Thaler's statement quoted earlier:

> How, without making normative judgments, do we determine what counts as complete information, unlimited cognition, or complete willpower? Even if we can specify what it would mean to have these supernatural powers, how do we discover how some ordinary human being would act if he were somehow to acquire them?[26]

What would Patrick have done if he hadn't decided to eat that muffin? There's no way to know, because he did eat the muffin—the counterfactual is anyone's guess.

But it's not just anyone's guess—it's the libertarian paternalist's guess, made to "correct" Patrick's preferences and then nudge him toward what he or she picks out as the "right" choice. In general, an outside observer decides that someone else's preferences are "irrational," and then tries to determine what preferences this person would have—or *should* have—if he or she were thinking "more rationally." But this involves a judgment on the part of an outsider in the same way that more blatant value substitution does. The behavioral economist still has to decide that a person's preferences (as revealed through his or her choices) are not "good" enough and then decide how to modify them, since it is impossible to know what the person's preferences would have been if they weren't what they are![27] A behavioral economist may decide that, were Patrick in his right mind—which "happens" to look a lot like the behavioral economist's mind—he would rather have a fruit cup than his Sunday muffin. But there is no way behavioral economists can sincerely claim to be stating Patrick's true interests: they are merely stating their own. In the end, rational preferences are nothing but value substitution conducted in a slightly different way—and with the same ethical implications.

### Nudges Do Something, But Not What They Promise

Although libertarian paternalism and nudges may work very well in terms of manipulating choices in the interests imagined by their designers, they do not—indeed, they *cannot*—do as they claim, which is to help people make decisions that *they* would like to make. Regulators simply do not have access to the kind of information that would allow them to do that—not even people themselves can articulate their interests that completely. Rather, they steer people into making decisions that behavioral economists and libertarian paternalists have chosen for them based on their own judgment.

Let's be careful, however, not to be overly suspicious or cynical. (Who, me?) There's no reason to believe, for instance, that Sunstein and Thaler have any personal interest in people's retirement savings. I presume they merely believe it's prudent to save for retirement—which it is, undoubtedly. But to maintain, even implicitly, that it's the only or most important interest for every person, even in the context of a narrow decision such as 401(k) enrollment, involves imposing the interests they choose for people in place of people's own.

It is not difficult to imagine, however, some policymakers nudging decisions for less benevolent reasons. Investment firms who administer 401(k) plans would happily support programs like Save More Tomorrow, and producers and sellers of health food would certainly endorse any nudges that got Patrick away from the bakery. We all know companies manipulate choices anyway, through advertising and promotion—to be sure, companies read behavioral economics books too—but you're free to ignore them, unlike choices you're forced to make by your employer or government. (More on this in chapter 6.)

Here's a more frightening prospect. Ballots for elections were criticized widely around the time of the 2000 presidential election for being confusing, and there's a case to be made for simplifying the layout and arrangement of the various candidates and referendum items. Now let's imagine that the members of a local election board decide that the candidate they personally prefer might not be supported as heavily by the rest of the voters in the area. From their point of view, the wrong person might be elected because voters have "misaligned" preferences, manipulated by vicious campaign ads, blatant appeals to emotion, and biased media reporting. If these voters only knew better—as the members of the election board do—they would obviously vote for the "right" person.[28] But they may need a little help, and seeing that ballots are so confusing and they need to be redesigned anyway, why not redesign them so people are "helped" to vote for the right person. It's what voters would do if they were thinking correctly anyway. It's in their best interests. Really!

We don't need to imagine threats to democracy to emphasize why choice is so important, so vital, and so essential to each and every person. (We'll talk more about that in the final chapter.) This doesn't mean people always make the best choices, or even good choices; as I said earlier, we all make dumb choices from time to time. But they can be judged as dumb only in comparison to our interests, and the only people able to make those judgments are *us*. Behavioral economists and libertarian paternalists claim to respect people's true interests whereas they influence people's behavior based on interests of their own choosing. I have no idea whose interests this serves, but it definitely isn't ours.

# Chapter 5

# Why Nudges Are Unethical

*In the last chapter, I discussed practical problems with libertarian paternalism based on information and interests, but I couldn't seem to avoid the closely related ethical issues regarding respect and autonomy that are raised by value substitution. In a way, value substitution has been in the background since the first chapter of this book, since any economic model used to explain and predict behavior has to assume some goals that are likely not the goals of any real-world person. In that case, value substitution is a problem with designing models and interpreting results—a* methodological *problem—but it becomes an* ethical *problem largely when policymakers use these models to influence behavior, especially when they presume to do it in people's interests. In other words, it becomes morally problematic when those in power act on the recommendations of behavioral economists, and the theoretical issues with behavioral economics are integrated into policy and start influencing people's lives.*

*Even if regulators did have information about people's true interests, it is nonetheless paternalistic of them to enact laws or regulations to benefit those interests on people's behalf. The term* paternalism *doesn't sit well with most people, invoking thoughts of Big Brother from George Orwell's novel* 1984 *or "nanny-states" in which the government presumes to take care of its citizens like a parent rightfully takes care of his or her children. This is for good reason: as this chapter explains, paternalism is inherently* *opposed to personal autonomy, no matter what form it may take, because it interferes with people's determination and pursuit of their own interests. And as we saw in the last chapter, even if policymakers are well-intentioned, they cannot possibly know people's true interests, and as a result any paternalistic laws or regulations are crafted in policymakers' idea of people's interests (usually overly simplistic and biased). Specifically, the "libertarian paternalism" of Cass Sunstein and Richard Thaler (as well as other sympathetic writers) falls victim to the same flaw as does any type of paternalism: rather than serve people's own interests, as it claims, it serves the interests that policymakers think people do or should have.*

*Sunstein and Thaler use the term* libertarian paternalism *in hopes of dispelling these concerns and convincing the reader that nudges are not coercive. But I argue that libertarian paternalism is very much coercive, and in some ways more insidious than "old school" paternalistic policies such as prohibiting or taxing behavior. Rather than telling people what to do or not to do, or influencing them explicitly with taxes or subsidies, nudges—such as changing default options or the arrangement of choices—have an intrinsically covert nature, designed as they are to piggyback on people's cognitive biases and dysfunctions to "guide" them into the "right" choices. Even if one is comfortable with some paternalism on the part of the government if done openly and transparently, it is unseemly for policymakers to use people's decision-making flaws to manipulate them, subtly and secretly, into making choices that policymakers want them to make, rather than the ones they would have otherwise made themselves.*

* * *

Paternalism is defined by philosopher Gerard Dworkin as "the interference of a state or an individual with another person, against their will, and defended or motivated by a claim that the person interfered with will be better off or protected from harm."[1] The two key elements of this definition are interference (with choices and actions) and purpose (to benefit the person interfered with). Our discussion in the previous chapter about value substitution criticized the stated purpose of libertarian paternalism: even if regulators sincerely want to help people with paternalistic laws and policies, they should realize that it is impossible to know other people's interests well enough to do this. This is both a practical problem—how can regulators nudge people in their own interests if they do not and cannot know what those interests are—and an ethical problem, deriving from the right of all persons to develop and pursue their own interests (provided this doesn't interfere with anybody else doing the same).

In this chapter we will focus on the coercive aspect of paternalism, especially libertarian paternalism. Let's be very clear from the start that coercion in and of itself is not always wrong. For example, criminal laws coerce us, whether or not we will ever be tempted to break them, and criminal penalties such as fines or imprisonment are coercive measures designed to enforce criminal laws. But we accept this coercion because (1) we are protected by such laws, and (2) criminal laws (for the most part) aim to prevent harm inflected by one person on another. If we consider theft and murder to be coercive acts on the part of our fellow citizens, then the criminal law can be seen as coercion to prevent coercion.[2] But paternalism is not protecting us from the deviant actions of each other—it's based on protecting us from our own "deviant" actions! It is precisely the purpose of paternalistic law and policies that makes their coercive nature so offensive.

The classic argument against paternalism was provided by the philosopher John Stuart Mill in his 1859 book *On Liberty*. In it, he articulated what has come to be known as the *harm principle*, which maintains that harm to others—not to oneself—is the only justification for government coercion:

> The only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others. His own good, either physical or moral, is not a sufficient warrant. He cannot rightfully

> be compelled to do or forbear because it will be better for him to do so, because it will make him happier, because, in the opinions of others, to do so would be wise, or even right. . . . The only part of the conduct of any one, for which he is amenable to society, is that which concerns others. In the part which merely concerns himself, his independence is, of right, absolute. Over himself, over his own body and mind, the individual is sovereign.[3]

Mill emphasized the right of all individuals to determine how they will run their lives, insofar as they do not interfere wrongfully with anybody else trying to do the same—in other words, as long as they don't cause harm to others.

This aspect of Mill's argument against paternalism is based on *autonomy*, the right to determine one's own interests and actions. Autonomy is valued by many in and of itself, but it also contributes to many other goods, such as human dignity and lifelong fulfillment. (We'll have a lot more to say about the value and importance of autonomy in chapter 7.) It can also be understood as a critical component of basic happiness and well-being—even well-being understood as preference-satisfaction—based on the understanding that a person will know his or her own interests, and therefore promote them, better than others would. (Sound familiar?)

Mill promotes this idea as well:

> The strongest of all the arguments against the interference of the public with purely personal conduct is that, when it does interfere, the odds are that it interferes wrongly, and in the wrong place. On questions of social morality, of duty to others, the opinion of the public, that is, of an overruling majority, though often wrong, is likely to be still oftener right; because on such questions they are only required to judge of their own interests; of the manner in which some mode of conduct, if allowed to be practiced, would affect themselves. But the opinion of a similar majority, imposed as a law on the minority, on questions of self-regarding conduct, is quite as

> likely to be wrong as right; for in these cases public opinion means, at the best, *some people's opinion of what is good or bad for other people.*[4]

The astute reader will recognize this as the reasoning behind consumer sovereignty (from chapter 2): since it is assumed that people know their personal preferences (or interests) better than anyone else, they will promote their well-being better by making their own choices than by having them made for them. So even if one were to deny that autonomy had any independent value, it nonetheless contributes to a person's well-being (or the pursuit of his or her interests).

Mill was no absolutist on paternalism, however. He allowed that the government may interfere with people's actions for their own good *if* their actions are judged to be involuntary. For instance, in *On Liberty* he argued that

> if either a public officer or any one else saw a person attempting to cross a bridge which had been ascertained to be unsafe, and there were no time to warn him of his danger, they might seize him and turn him back, without any real infringement of his liberty.[5]

Nowadays, paternalistic interference with involuntary action is known as *soft paternalism*, as opposed to *hard paternalism*, which describes interference with voluntary, fully aware action.[6] Hard paternalism obviously invokes the problem of value substitution, because the state is interfering with a person's behavior even when the person can be presumed to be acting voluntarily in his or her own interests. There is still, however, a hint of value substitution with soft paternalism: after all, it is not unimaginable that the person crossing the unsafe bridge in Mill's example is well aware of the risk but had good reason to cross it anyway. However, since actions like this are rare and situation-specific, there is less of a danger of co-opting people's interests on a regular basis with a law or regulation. (Plus, the fellow can always try to cross the bridge again later.)

## There's Voluntary and Then There's Voluntary

You're probably asking yourself (or screaming to the book) what "voluntary" means in this context—and well you should. We can safely presume that the man in Mill's example did not voluntarily walk onto an unsafe bridge, and we would say the same thing had he been pushed or tricked into crossing it. But what if he merely underestimated the risk, or didn't fully appreciate the risk—say, because of a cognitive dysfunction when assessing risk (such as optimism bias)? Philosopher Joel Feinberg, one the most important modern writers on paternalism, based his conception of voluntary action on what is considered necessary for moral or legal responsibility. In both law and morality, we normally hold people responsible for things they do voluntarily, but not if they were manipulated or coerced somehow, and this standard of voluntariness works well for filling out Mill's justification for paternalism also.

For example, if we see Brad pushing Carl into the path of an oncoming car, we would hold Brad responsible for any injuries Carl might suffer *unless* there is some reason to believe Brad did not act voluntarily. For instance, if Brad simply tripped on an uneven sidewalk and stumbled into Carl, or if we discover that Andy pushed Brad into Carl, we would not hold Brad responsible.[7] Philosophers would even say that Brad didn't "act" at all in these cases, but was acted upon by the laws of physics or Andy, who may have used Brad as a tool to hurt Carl. We would say the same thing if Andy held a gun to Brad's head and ordered him to give Carl a shove: Brad certainly did act in this case, but under duress (a threat to his life). Brad can be considered to have acted involuntarily in all of these cases—and according to Feinberg, we can consider his actions involuntary for the purpose of justifying paternalism as well (if it were Brad instead of Carl who was in danger of being hit by a car due to Brad's involuntary action).

In his 1986 book *Harm to Self*, Feinberg lays out a thorough model of what constitutes "perfectly voluntary choice," which he defines as "deficient in *none* of the ways that are *ever* taken into account for *any* moral or legal purpose in *any* context."[8] This is no simple matter: he dedicates the last 300 pages of the book drawing out the fine details! His model clearly represents an impossible, idealistic standard—which is made clear by the highlighted words in the quote above—but nonetheless provides a usual benchmark for judging deviations from it.

According to Feinberg, perfectly voluntary choice depends on these five conditions:

A. The chooser is "competent."
B. He does not choose under coercion or duress.
C. He does not choose because of more subtle manipulation.
D. He does not choose because of ignorance or mistaken belief.
E. He does not choose in circumstances that are temporarily distorting.[9]

Each of these is obviously vague, which is why he spends so many pages fleshing them out. In his initial presentation of these factors, he includes several elaborations or examples under each one; for instance, under competence, he lists "not an animal; not an infant; not insane" and so forth. Of particular interest to behavioral economists and libertarian paternalists are the last two categories, regarding ignorance, mistake, and distortionary circumstances, including impulse, fatigue, and overwhelming emotion. (In response, I would point out the third category: "subtle manipulation.")

Before our friends get too excited, however, they should note several things. Feinberg was clear, as we saw, that this is an impossible ideal of voluntariness. This level of perfection even exceeds the conditions for "rational preferences" we saw in the last chapter. As I said, Feinberg spends hundreds of pages teasing out the nuances of his description of voluntariness, exploring how each element contributes to a voluntary choice and what each implies about the justification of paternalism. When it is unclear whether an action was voluntary "enough" to be immune from paternalistic interference, the consequences of the action should be considered before anyone steps in to prevent it. Take the man crossing the unsafe bridge in Mill's example: the threat of great injury or even death would make interference worthwhile even if we could not be certain that he was acting involuntarily. But if the same man appears about to step into a shallow rain puddle—even a very, very wet one, perhaps containing a bit of mud which may stain his best trousers and embarrass him later that evening when he calls on Lady Pimbletonshire—the danger he risks would not be sufficient to justify interfering with his action (even keeping mind the well-known ficklemindedness of Lady Pimbletonshire).

But even if we consider the consequences of an imperfectly voluntary choice along the lines that behavioral economists would emphasize, there is yet another hurdle that will be difficult to climb.[10] It's impossible to assess the consequences of any decision to the person making it without knowing his or her interests. We can be fairly certain the interests of a man about to cross the unsafe bridge will be thwarted, but in the case of more mundane choices—such as planning one's retirement or next meal—we have no way of knowing if there are likely to be dire consequences from a certain choice unless we know why the person made it. The only other way to judge this choice is to substitute some interests for the person's own, as we described in the last chapter, at which point the choice is no longer being judged according to the person's own interests.

More important, in most ordinary choice situations, an outside observer cannot know if a choice is involuntary just by observing the results and then disapproving of them. The standards that Feinberg lists are formal or procedural, speaking to a person's decision-making *process* rather than its *outcome*. Soft paternalists don't judge the outcome of the action against their idea of the person's interests. They assess only the voluntariness of a person's decision-making according to the standards listed earlier: was the choice made under duress, was it manipulated, was it mistaken, and so forth. If a person satisfies Feinberg's conditions to a sufficient extent (whatever that may be), then his or her choice is judged to be voluntary and therefore made in the person's best interests—and no paternalist interference is necessary, no matter how "unfathomable" or "foolish" the behavioral economist or libertarian paternalist may think the choice is.

If libertarian paternalists, as they claim, are to be soft paternalists—acting to promote people's own interests rather than imposing their own—they need to know more than merely a person's actions, from which they normally "conclude" that the person made a bad choice. Instead, they need to evaluate the process of decision-making itself—which they've done in laboratory settings with carefully controlled studies, but which is much more difficult to assess in the real world. Take the example of the low enrollment rate in 401(k) programs that Sunstein and Thaler designed nudges to remedy. They didn't assess each employee's thought process when he or she chose whether to enroll in a 401(k) program. They simply judged the rate of enrollment to be low compared to what they regarded as rational, and from that they decided that people were not making sound decisions. If new employees made decisions in their best interests, the thinking goes, they would enroll, and if they didn't, they must be confused, or mistaken, or lazy—in other words, not completely voluntary in the sense of informed and rational decisions. Or, as Thaler and Sunstein put it:

> We can say for sure that *some* people in our society are definitely saving too little—namely, those employees who are not participating at all in their retirement plan, or are saving a low percentage of their income after having reached their forties (or older). These folks could clearly use a nudge.[11]

Thaler and Sunstein are evidently not concerned with discerning the plethora of other interests that may have led these employees to *rationally* and *voluntarily* defer enrollment in retirement programs. Their choice to defer is alone evidence enough of their irrationality, and justifies paternalistic intervention to correct their involuntary "bad" choices. This is hard paternalism disguised as soft paternalism: pretending to judge the voluntariness of people's decision-making processes but really questioning their choices and their interests. This should come as no surprise given the discussion in the last chapter, but the point holds even after we acknowledge the possibility of involuntary choice, which must be assessed at the level of decision-making itself, not just on its results.

### There's Coercion and Then There's Coercion

Paternalistic coercion can take many forms, most of them explicit. Some ban an activity outright, such as using illegal drugs. Other require certain precautions, such as seat belts or helmets, when engaging in an activity judged to be dangerous (riding in a car or on a motorcycle). Yet other paternalistic policies don't forbid an activity outright, but make it more difficult or costly to engage in, such as taxes on cigarettes that are meant to reduce smoking.[12] But nudges—by design—are not explicit and do not force any action or inaction; rather, they are subtle rearrangements of options meant to steer people toward a choice that is presumed to be better for them than the choice they would have otherwise made. For this reason, supporters of libertarian paternalism do not see nudges as being coercive at all.

Sunstein and Thaler's argument in their law review article that nudges are not coercive is succinct: "Libertarian paternalism is a relatively weak and nonintrusive type of paternalism, because choices are not blocked or fenced off."[13] Later in the paper they raise the example of rearranging items in the cafeteria, arguing that "the choice of the order in which to present the food items does not coerce anyone to do anything."[14]

To be sure, nudges are not comparable to a mugger holding a gun to your head or a police officer standing ready to arrest you. However, because they are designed with the express purpose of manipulating people's decision-making processes to change their behavior in pursuit of interests that are not their own, I argue that nudges *are* coercive—and, to make matters worse, the coercion operates through the very cognitive biases and heuristics that behavioral economists use to justify the interventions.

We must be careful here not to confuse paternalism and coercion. As Gerald Dworkin elaborates beyond the definition given at the beginning of this chapter, paternalism involves "a usurpation of decision-making, either by preventing people from doing what they have decided or by interfering with the way in which they arrive at their decisions."[15] Although both cases are paternalistic, the first is clearly coercive while the second is not obviously so, since nothing is being forced or threatened in the ordinary sense of the term. It is an accurate representation of libertarian paternalism, which interferes with the way people make decisions, but the case for coercion still needs to made.

Consider an example: Dr. Brown is treating Karen, a terminally ill patient who has requested an operation that, if successful, would add a couple months to her life, but has a significant chance of ending it altogether. Dr. Brown does not think the operation is worthwhile for someone in her condition and has urged Karen to reconsider the operation and spend her remaining days with her loved ones instead. But she is insistent, more than willing to pay any costs involved as long as she has a chance of surviving a little longer. So Dr. Brown decides to exaggerate the risks of the operation in hopes of "convincing" her to make the choice he thinks is best for her. (We can assume he's sincere in this.)

Let's say Karen agrees to forgo the operation. She made her decision freely in the sense that no one forced her to decline the operation. But it is more than fair to say Dr. Brown, with his embellishment backed by his authority and the trust she has in him, manipulated her decision. By exaggerating the negative outcome of the operation, he put a thumb on one side of the scale. Essentially, Dr. Brown lied to Karen to get her to make the decision he thought was in her interests, rather than the decision she would have made if he had told her the truth. He may have had her best interests—as he judged them—at heart, but he violated his patient's autonomy, one of the core tenets of medical ethics, as well as ethics in general.

As John Stuart Mill wrote, the ideal of autonomy—the ability of each person to decide his or her own interests and make choices in pursuit of those interests—is one aspect of paternalism that makes it so offensive. Paternalism, by its very nature, denies autonomy in two ways: by substituting someone else's idea of a person's interests for that person's own, and by blocking or manipulating choice to promote the interest imposed by the paternalist. Autonomy was a key concern of another moral philosopher, Immanuel Kant, who believed that deception and coercion were the two main ways that autonomy could be compromised, specifically by using people as "tools" in another person's plans. Both deception and coercion exclude people from full participation in what's going on and deny them the right to fully consent to it—both of which would compromise their autonomy.

Kant used the example of making a false promise to get money out of someone in his 1785 book *Grounding for the Metaphysics of Morals* to make his point about coercion and deception:

> The man whom I want to use for my own purposes by such a promise cannot possibly concur with my way of acting toward him and hence cannot himself hold the end of this action . . . [A] transgressor of the rights of men intends to make use of the persons of others merely as a means, without taking into consideration that, as rational beings, they should always be esteemed at the same time as ends, i.e., be esteemed only as beings who must themselves be able to hold the very same action as an end.[16]

As he wrote, coercion and deception both use a person as a means to another person's end without at the same time being treated as an end—as a rational, autonomous person.[17] Coercion and deception, which are closely related in this sense, both deny people the respect owed them, bypassing any sort of informed consent that would ensure us that they were included fully in the other person's plans. Ideally, a person who wants another's cooperation would try to persuade the other person, to appeal to his or her judgment. As philosophers Dan Hausman and Brynn Welch write: "Rational persuasion respects both individual liberty and the agent's control over her own decision-making, while, in contrast, deception, [by] limiting what choices are available or shaping choices, risks circumventing the individual's will."[18]

In our example, Dr. Brown lies to Karen about the risks of her operation, rendering her unable to consent to his true plan (to dissuade her from the operation) simply because she is unaware of it. She trusts him to give her unbiased information about the risks, and he relies on this trust to manipulate her choice toward his ends (however benevolent they may be). Simply put, he doesn't want her to have the operation, and he lies to her—uses her—to achieve this end. His actions are both deceptive and coercive in that he keeps his true end from her and then manipulates her decision-making toward that hidden end. To look at it from Karen's point of view, she has no opportunity to consent to Dr. Brown's goal because she is unaware of it and is subject to her doctor's subtle playing on her fears (regardless of his benevolent intent). To separate the explicit deception from the example of Karen and Dr. Brown, we can say that rather than exaggerating the risks of the operation, he merely overemphasizes them, giving Karen the worst-case scenario likelihood of death or coma. In that case, the coercion shines through based on his manipulative efforts alone and his hidden goals.

However, the coercive aspect of libertarian paternalism has to do less with the manipulation involved and more with the fact that nudges involve steering people toward making choices to promote ends or interests other than their own. (We'll see in the next chapter that manipulation based on behavioral research is not necessary coercive if the goals or ends of the manipulative party are well-known.) After all, if people were being nudged in their own interests, the nudges would be redundant. As Gerald Dworkin writes:

> The denial of autonomy is inconsistent with having others share the end of one's actions—for if they would share the end, it would not be necessary to usurp their decision-making powers. At one level, therefore, paternalism seems to treat others as means (with the important difference that it is a means to their ends, not ours).[19]

As we know, the ends or interests promoted by libertarian paternalists are not people's true interests, but rather the paternalists' idea of what those interests *should* be. When new employees are nudged toward enrolling in 401(k) plans, this is not done in their interests, but in the interests chosen for them. Libertarian paternalists claim, of course, that they are acting in people's true interests, but this is impossible—and even if it weren't, it couldn't justify the manipulation at the core of libertarian paternalism.

Consistent with behavioral economists' position regarding the unreliable nature of preferences, libertarian paternalists could use the concept of "rational preferences," as we discussed in the last chapter, to claim that people *hypothetically* consent to nudges since they serve their "real" interests. As the argument goes, if people were "rational"—fully informed, free of cognitive biases and heuristics, and calm of mind—they would make the choices they're being nudged into making anyway, so libertarian paternalists can simply assume that people *would* consent to the interventions *if* they were fully rational. But this approach is just as questionable as the concept of "rational preferences" itself based on the fact that the "rational person" I would be if I weren't who I actually am is a fiction. As Gerald Dworkin writes, hypothetical consent "is not actual consent that remains unexpressed. It is simply a judgment about what the agent would have agreed to under certain circumstances"—a judgment made by the libertarian paternalist based on his or her ideas of what people "should" prefer.[20] If consent is desired, it must be actual, not hypothetical; in the words of philosopher Onora O'Neill, "the morally significant aspect of treating others as persons may lie in making their consent or dissent *possible*, rather than in what they actually consent to or would hypothetically consent to if fully rational."[21] So even if libertarian paternalists had information about people's real interests and designed nudges to steer people toward them, this would still be coercive because the people would be unaware of the regulator's true ends (even if they coincide with their own) and have no chance to consent to them.

Another way to understand the coercive nature of nudges is that they are designed to change people's behavior in someone else's interests by relying on unconscious anomalies in decision-making. They do so not by force or threat, and not by persuasion to get voluntary compliance, but by relying on the same cognitive biases and heuristics that justified their use in the first place. The fact that nudges make use of defects in decision-making to alter people's decisions away from what they would have chosen otherwise—and toward those that support other interests—contributes to the case for coercion. Nudges do not take the more explicit forms of bans or taxes, which leave choices blocked or changed in obvious ways. They have more in common with tricking or fooling someone into a certain decision—"for their own good," of course—a decision that, at the point in time at which they actually decide, they are "free" to make. And there need be nothing deceptive about the nudge itself—changing a default choice and the arrangement of options don't disguise or hide anything from the decision-maker—but the goals or ends of the nudge are often hidden, and the cognitive biases and heuristics they rely on are rarely made clear.

If one is committed to paternalistic intervention, there remains little reason to favor nudges over more overt methods of paternalism given the coercive nature of both. If the government wants to discourage certain behavior, such as unhealthy eating or insufficient savings, they have the tools of bans, taxes, and subsidies to change the incentives to engage in them. These are no less paternalistic or coercive, but they do have a clear advantage: they are out in the open for all to see. Everyone can understand why they're being done and how it works, and most important, they leave people to make choices within the new parameters clearly determined by the policy. All smokers know why the price of cigarettes is so high—and they have every right to be offended by it—but they can make decisions, based on the high price of cigarettes, that further their interests. But nudges operate on a more subtle level, taking advantages of flaws in our decision-making processes, so we may not even know that we're being nudged. People are manipulated by libertarian paternalists into doing things that are literally against their will, making use of the same cognitive defects that so compromised their choices in the first place. The new employees who are too lazy to opt out of automatic 401(k) enrollment did not "choose" to enroll; they were tricked into it by a default option designed to elicit just such behavior. It's clever coercion, but coercion nonetheless, and all the more offensive due to its cleverness.[22]

### But It's Inevitable! We Can't Avoid It!

Libertarian paternalists also argue that no matter what you think of nudges, there's no way around them. After all, you have to arrange choices somehow; you need to designate some option as the default. As Sunstein and Thaler write: "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."[23] What's a choice architect to do?

Note that Sunstein and Thaler confuse making "a choice" with making a choice *paternalistically*. Obviously one arrangement or default option out of many must be chosen, but that choice is only paternalistic if it is chosen specifically for the decision-makers' own good. To Sunstein and Thaler, however, there simply is no other choice. In their law review article, they propose the alternatives available to the person charged with the presentation of food items in a cafeteria:

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.[24]

"Option 1 appears to be paternalistic," they write, and ask, "Would anyone advocate options 2 or 3?" Option 4 is "much harder to implement than it may seem," because people's preferences are not well-formed, instead depending, in part, on the presentation of the items themselves.[25] Because preferences are unreliable guides to people's well-being, they feel it better for the cafeteria director to choose option 1, arranging the items based on what she feels are her customer's best interests. ("All things considered," of course—oh, the irony.)

Let's look at these options a different way, and suggest some more. We need not discuss option 1—that's what this book is about! Option 2 is ridiculous—why would anyone even think of randomizing the assortment, unless a underpaid and resentful cafeteria worker simply threw the food willy-nilly on the shelves—and option 3 is simply absurd. Option 4, besides being difficult to implement, is completely redundant, unless the cafeteria director thinks she has tremendous sway over customer's choices, in which case option 4 is indeed impossible to implement since people's choices would be dependent entirely on how she chose to arrange the items.

Why is it the only two plausible options they suggest require the cafeteria director to gauge her customers' preferences or interests? Here are some options Sunstein and Thaler did not suggest:

5. She could arrange options by some commonsense ordering.
6. She could arrange options by weight or fragility, so the more difficult items to handle are closer to the customer.
7. She could arrange them in an aesthetically pleasing manner so as to make the cafeteria a lovely haven from the stresses of the 9–5 (or, increasingly, 9–9) workday.

(Feel free to use the inside back cover of this book to suggest your own—try it, it's fun!) As we see from the three alternatives offered above, there are different ways to think about designing a cafeteria than what Sunstein and Thaler suggest. Option 5 recognizes that there are natural ways to arrange food items based on cultural norms: for instance, soups and salads first, followed by entrées, and then dessert at the end. (Weird, right?) Option 6 could be seen as paternalistic, I suppose, but I prefer to think of it as being considerate and practical; if bowls of soup are offered, I'd much rather they be placed near my tray rather than several feet above, risking a nasty spill on my person. Option 7 is perfectly reasonable as well; the director may want to organize the food items by size, shape, or color, in a way that pleases her as well as her customers (not to mention her underpaid and resentful workers). The important thing to note is that *none* of these plans takes into account any influence they may have on customers' choices. That isn't even a concern, much less an "inevitable" one. But if the only tool you have is a nudge, every choice looks like a bad one—and every cafeteria looks like "choice architecture."

Sunstein and Thaler also point to the designation of a default choice as inevitable, and it is—but again, it need not be paternalistic. This is clear in the case of enrollment in 401(k) programs: either the employee is not enrolled by default with the option to join (an "opt-in") or the employee is enrolled by default with the option to refuse (an "opt-out"). As they explain it:

> 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 their earning 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.[26]

Indeed, there is no law of nature that says an employee must make an active choice to be enrolled—but there would seem to be an ethical principle that points us in the same direction. It is more consistent with respect for employees' autonomy that they not be enrolled in *any* programs incidental to the requirements of their job without their express consent. Will money be deducted from the new employee's paycheck to pay for a health club membership unless he or she opts out? Will money be deducted to contribute to a political candidate the employer feels would promote the employee's interests unless he or she opts out? Such plans are not unknown to Sunstein and Thaler, who suggest that "if private or public planners would like to increase charitable donations, they could easily do that simply by creating automatic deductions for charity. Even if workers are allowed to opt out, clever planners should easily be able to ensure a much higher level of donations."[27] (Very clever indeed.)

An employee agrees to perform certain services for a certain package of salary and benefits, and any discretion he or she has regarding said benefits should not be exercised on his or her behalf with a nudge, even with an option to opt out. (See how I can sound like a lawyer when I try? Expect my bill.) Remember, after all, that this nudge is based on cognitive biases that will make the employee unlikely to change the default "setting." Instead of reinforcing that irrational tendency by using it for paternalistic purposes, employers (and any government that encourages such initiatives among businesses) should respect their employees' right to choose without any manipulation of their part. They should acknowledge, as argued in the last chapter, that what appears to them to be an irrational "failure to choose" may very well be a rational choice on the part of an employee who has many other interests besides retirement savings to consider at this time in his or her life.

Sunstein and Thaler anticipate a proposal that would seem to avoid the issue of setting a default option at all: forcing a response one way or the other. In the 401(k) case, the employee can be forced to check "yes" or "no" when asked if he or she wants to enroll. This would avoid the issue of having to select either option as a default—and therefore any accusation of paternalism—and would account for cognitive biases such as the status quo effect and simple laziness (or, more likely, human-resources form fatigue!). But Sunstein and Thaler downplay the benefits of forced choice, arguing that "the very requirement that employees make a choice has a strong paternalistic element."[28] This is difficult to understand, since a new employee is forced to make many choices, out of which a retirement plan is just one. It may be coercive—just as coercive as having to pose for an ID badge picture—but it is not paternalistic because it does not privilege one choice over another in the employee's own interests. As Sunstein and Thaler point out, a choice has to be made somehow (enroll or not enroll), but the company is not behaving paternalistically as long as it doesn't nudge the employee toward one choice or the other. The company is merely requiring the employee to make what is, at the point of the decision, a choice free of any paternalistic manipulation.[29]

Keep in mind, however, that Sunstein and Thaler are not concerned with being characterized as paternalistic, so maybe it is the explicit coercion involved with being forced into a choice that bothers them. But ultimately, for them, it is "an empirical question," one which they presumably answer by citing a study showing that forced choice results in lower enrollment than unforced choice with default enrollment.[30] As always, the real concern of libertarian paternalists is generating the behavior they want to see, not helping people make the choices they want to make. In the end, forced choice regarding enrollment, which respects employees' interests, is judged inferior because it doesn't further the regulators' idea of those interests. That's the point of paternalism, and it's much more troubling than coercion by itself.

This reversal of the stated goal of libertarian paternalism reveals a disturbing attitude on their part toward people (including themselves, as they often note). To them, people are not responsible, autonomous decision-makers with complex interests and goals who pursue them (albeit imperfectly). Instead, they are machines that are broken, responding in the wrong ways to their sensory inputs and not fulfilling their true function. As a result, their responses need to be adjusted, fine-tuned, so that they behave as they "should," according to the goals the regulators prefer—for their own good, of course. This mind-set goes a long way toward explaining the jump from "sometimes people make bad choices" to "we can and should help them make better ones." If your toaster burns your toast, it isn't operating correctly, and needs an adjustment. And if your employee isn't saving enough, "it" too needs an adjustment. So simple—but so, so wrong.

## But Wait—There's More!

Although we've discussed nudges in terms of libertarian paternalism, there are other uses for behavioral insights, within both government and business. Governments use nudges to get more people to recycle and donate blood, for instance, and businesses play on our cognitive biases and heuristics to sell more products and make more profit. In the next chapter we'll turn our attention to these other uses of behavioral research, and ask if they're just as ethically questionable as paternalistic nudges. You'll have to read the next chapter to see what that answer is, of course, but here's a hint: we'll find it in the purpose of such tactics. Exploring this idea will bring us back, full circle, to paternalism and the issue of whether it is legitimate for the government to take an interest in our interests. And even if it is—especially if it is—we'll see why nudges are ultimately counterproductive, particularly in terms of our decision-making abilities. It turns out that nudges are a lot like the adage about giving a man a fish versus teaching him to fish: even if nudges lead to better choices, they don't help us to *learn* to make better choices in the future.

# Chapter 6

# All Nudges Are Not Created Equal

*Some of you reading this book may be wondering why I focus almost exclusively on the use of behavioral economics on the part of government and not private business. Certainly, private companies use the same behavioral research on cognitive biases and heuristics to get consumers to buy their products. Why don't I criticize private companies for this as well, especially since businesses are using these tools to increase their own profit, while the government, even if it's not effective and ethically problematic, is at least trying to help? And what about nudges used by the government for nonpaternalistic reasons, such as recycling? Are these also problematic?*

*In this chapter, I address these questions by pointing out key differences between governments and private businesses, mainly in terms of the attitudes they have toward citizens or customers. The issue, once again, hinges on both interests and respect and serves to emphasize why paternalism is so disrespectful, especially when implemented with nudges. To look at it another way, while the most offensive aspect of nudges may seem to be their manipulative nature, it is actually their paternalistic and presumptive aspects that should worry us more. Paternalism also reflects an inappropriate attitude of a government toward its citizens that leads us to question its intention of trying to help in the way it does. We'll see that although governments may be understood to have many duties and responsibilities to its citizens, these do not include "caring" for them, like friends or family do.*

*The disrespect of paternalistic nudges for all these reasons is further compounded by the way in which they rely on the very cognitive biases and heuristics that motivated them in the first place. In the final section of this chapter, we'll see that nudges serve to reinforce these cognitive defects, creating a cycle of "bad" choices that perpetuates the "need" for nudges (and contributing to concerns that nudges will lead to a "slippery slope"). If this is how the government looks out for the interests of its citizens, then libertarian paternalism is even worse than we thought.*

* * *

Long before the work of behavioral economists became widely known, retailers, marketers, and advertisers used their insights into consumer behavior—born out of experience more than research—to fine-tune their pitches to consumers. Behavioral psychology has been used, formally or informally, by companies not only to develop products but also to design packaging, advertising, and promotions, and retailers use these concepts to lay out their stores and their shelves to maximize the sales of the most profitable products. Books such as Martin Lindstrom's *Brandwashed: Tricks Companies Use to Manipulate Our Minds and Persuade Us to Buy* detail many of these techniques, both for marketers to increase sales and for consumers to avoid the common lures.[1]

The question facing us as we begin this chapter is not whether businesses use their knowledge about consumers' cognitive quirks to influence buying behavior—they do, clearly. What we want to know is whether the use of these tools by business is as problematic as their use by government. If we're going to criticize governments for nudging us this way and that, shouldn't we criticize business just as harshly? The simple answer is no, we shouldn't, but not because "government is bad" and "business is good," nothing that simplistic. Like much of our discussion so far in this book, the difference in how we look at businesses' and governments' use of behavioral research comes down to purpose: specifically, their motivation with respect to our interests.

Let's take an example: after a long day at work, Jennifer stops by her favorite coffee shop for a hazelnut latte. She is about to order her usual, a medium-sized latte—or whatever fancy Italian word this particular coffee shop uses for medium (assuming they don't just call the medium-sized drinks "large"). Her usual choice of the medium size may be the result of extremeness aversion, which we discussed in chapter 2: on most days, she may feel silly buying the smallest drink but indulgent buying the largest.

But earlier today she finished a major project and feels like rewarding herself, so she considers the larger size, which costs only 25 cents more. "What a bargain!" she says to herself, but we know better. The large size only seems like a bargain compared to the medium size, which is overpriced to convince consumers to upgrade to the large. Nonetheless, this triggers Jennifer's impulse to grab a bargain, and she orders the large hazelnut latte (half-caff, low-fat milk, please). The benefits of this strategy to the coffee shop are well-known to economists who study pricing of different sizes or qualities of the same product, whether boxes of cereal or seats on an airplane. Realizing that some people simply can't purchase the better variety, the business can lower the quality or size of the lesser one and make extra profit from it, while they drive those who can afford the better item to buy it, increasing the profit made there as well. (Consider that on your next flight as you shoehorn yourself into a coach seat. Those seats are made smaller and less comfortable to make first class look even better and worth the high price—if you can afford it. If not, you're . . . well, you know.)

We could go on: the coffee shop may have a combo deal wherein Jennifer can save money on a pastry or sandwich if she purchases it with her large hazelnut latte (half-caff, low-fat). They may also offer little trinkets, snacks, or CDs on the counter next to the register that she may be enticed to buy. Of course, on her way to get to the counter at all, Jennifer has to pass dozens of items for sale, all in hopes of triggering one impulse or another. ("Yes, I certainly need another coffee mug that serves as advertising for the shop from which I bought it. In fact, I'll take two!")

So Jennifer's favorite coffee shop uses all these tools to get her to buy larger sizes and more items than perhaps she had planned to buy when she first walked in. What makes this behavior on the part of the coffee shop different from government nudges? Remember that the main problem with nudges, as we explained in the last two chapters, is that they manipulate people into making choices in what are supposed to be their interests but in truth are not. This combination of value substitution and coercion is what makes libertarian paternalism so offensive. Are these aspects of paternalistic nudges present in the ways business use behavioral research to get us to buy more and buy big?

Is there manipulation in these business practices? Of course there is. Every single decision a business makes regarding the design, pricing, and promotion of a product is calculated to maximize its appeal to the consumer and, in turn, maximize the profit of the company. Some of these actions are explicit, such as when an auto manufacturer adds a feature to one of its cars, a tech company improves a gadget, or a clothing store marks everything at 20 percent off. And some of these actions are not as explicit, such as when the manager of a supermarket arranges shelves so that the items with the highest profit margin is placed at eye level, or designs the floor plan to steer you toward the highest margin items. (But hey, those shelves have be arranged somehow, right? It's unavoidable!)

But is this *coercive*? In the sense that any manipulation is coercive, it would seem so, yes. But then every salesperson is coercive, every political candidate is coercive—okay, bad example—and every person who has an agenda of any kind is coercive. We realize anyone with a goal is going to try to manipulate us in some way to help them achieve it, but we don't think of these actions as coercive. But the less explicit tactics used by business to sell their products, the kind based on the same cognitive biases and heuristics that both motivate libertarian paternalism and implicate it as sneaky, are no less sneaky when done by a business. Then again, political candidates dress impeccably, pose for pictures with their families and random babies, and always have plenty of flags in their TV commercials, all subtle ploys to grab our unconscious support. We may dislike this behavior, whether on the part of business, political candidates, or the slick player at the bar, but we don't regard them as coercive.

Why not? There is one very simple reason we usually don't think of these as coercive: these people or companies have no power over us. If we don't like their behavior, we can simply walk away, vote for someone else, or go to another store. If Jennifer gets sick of her coffee shop's manipulative promotions, she can try to find another coffee shop with practices she likes better. But libertarian paternalism on the part of the government often cannot be avoided because it's a matter of law, policy, or regulation with the force of the state behind it. You can walk out of the government cafeteria and avoid the salad that is made more appealing than the pudding, but it is not as easy to avoid choices such as automatic enrollment in retirement programs or charitable giving. Even though you have the option to decline, these choices were designed so that you wouldn't—and in any case, as we saw in the last chapter, you can't avoid making a choice in these situations (whether forced or not). But in most interactions with business, you have the choice not to decide and take your business elsewhere. (This assumes, of course, that you are aware of these people's and companies' sneaky tactics, but we'll talk more about this later in the chapter.) Nonetheless, our ability to refuse to deal with people and companies that use manipulative tactics weakens any case for labeling those tactics coercive, whereas our government's laws and policies are much harder to avoid.

### It's Not about You—and That's a Good Thing!

But coercion is not the real issue here, and it never was. As we discussed at the beginning of the last chapter, legal coercion is necessary at times, particularly where people need protection from other people. Governments cannot exist without exercising some degree of coercion, so except for those who believe that government is bad in and of itself (as some people do), most of us are willing to accept some coercion as long as it's used for the right reasons. And paternalism is not one of those reasons. Libertarian paternalists emphasize that their policies are not coercive (or very mildly so) to make the paternalism sound harmless, but we know better. The paternalism is the more offensive part, and is also the main reason that manipulative, psychologically informed business practices are not dangerous in the same way that nudges are.

To emphasize this point, we have focused throughout this book on value substitution, arguing that because there is no way someone else can know your interests (other than inferring them from your choices), anyone who claims to make choices for you in your interests is actually making them in pursuit of some other set of interests that are imposed on you (usually single interests such as wealth and health). But we can even set that aside, and pretend for the time being that behavioral economists and libertarian paternalists have you nailed, knowing precisely what your various interests are and how you balance them in different situations. Nonetheless, we can still distinguish government nudges from business manipulation based on their purposes: promoting your interests versus promoting their own. Although the government, when acting paternalistically, presumes to promote your interests, businesses have their own interests—primarily, to maximize profit. Businesses are interested in your interests only insofar as those interests lead you to buy their products, not for your good. Whatever you think of the profit motive, the advantage of the single-minded purpose behind much business behavior is that they're not presuming to make decisions for their customers in their own interests.

In the last chapter we borrowed the words of Immanuel Kant describing coercion and deception as using other people "merely as a means, without taking into consideration that...they should always be esteemed at the same time as ends, i.e., be esteemed only as beings *who must themselves be able to hold the very same action as an end*." Earlier in the same paragraph, he describes the victim of his false promise as a man who "cannot possibly concur with my way of acting toward him and hence *cannot himself hold the end of this action*."[2] What does he mean by "hold the same action as an end" and "hold the end of this action"? He means that the person must be able to agree to the ends of the other person, which the first person can't do if he or she is deceived about those ends (such as when offered a false promise) or coerced into serving those ends (such as if the money were simply stolen at gunpoint). What Kant did *not* say was that the first person actually had to be *willing* to agree to the other person's ends—just that there was the opportunity. Given the opportunity, the first person can either accept them or reject them, neither of which is an option under deception or coercion.

This understanding of coercion helps to clear manipulative business practices of that label, as well as highlight their nonpaternalistic nature. Whether or not we're completely aware of every aspect of a business' "means"—their various manipulative tactics—we are perfectly clear as to their "ends"—to make money. This is no secret, and it is an end we can either agree or disagree with when we engage in commerce with them (or walk away). We know businesses do whatever they can—hopefully within the law and other ethical norms of their industry—to make money, and we fully *expect* them to do whatever they can to make money, even if we're not aware of what exactly they do. It doesn't take an extreme belief in caveat emptor ("let the buyer beware") to realize that consumers should expect businesses to manipulate their behavior to some degree and to guard themselves against it to whatever extent they can.

This argument also extends to nudges used to promote other government ends that are not seen as paternalistic. Sunstein and Thaler devote entire chapters to describing how their approach can help increase the rate of organ donations and contribute to environmental causes.[3] Whatever your political views regarding these causes and programs (and other government efforts), they are not paternalistic in the sense we've been discussing here. They do not involve government regulators claiming to serve our true interests while necessarily imposing their own. Instead, these nudges are designed to change behavior in some public interest that is ideally endorsed, directly or indirectly, by the electorate. We can agree or disagree with these interests, of course, but that's the point—the goals of these programs are usually made public and subject to democratic scrutiny. If we don't like them, we can vote their supporters out or appeal to them directly. There is still some manipulation based on behavioral research, but it is in service of a legitimate public interest—the government maintains respect for personal interests and does not presume to know them and influence our behavior for our own good.[4]

This discussion leads us to the core difference between profit-motivated manipulation by businesses and paternalistic manipulation by government: we expect businesses to do it, *but we expect more from our government*. We know businesses do not share our interests and instead pursue their own ends using whatever tools possible (and allowed), but we're prepared for it. It's a process of adaptive give-and-take: they offer certain products at certain prices promoted in certain ways, and we buy some and don't buy others. They learn what works and what doesn't and they change their offerings, and we respond to that.[5] To a certain extent it's us against them, both sides struggling to get the better value or return, but in a larger sense we're dependent on each other, consumers for goods and services and businesses for money to provide income to owners, suppliers, and workers—who then buy the goods and services, and so forth.[6] Businesses and consumers don't have to like each other, but they each acknowledge the other's role in the game and what the other is after.

But we don't like to think ourselves as being at odds with our government. We want to believe our government treats us with respect, not subtle manipulation. As economist Edward Glaeser writes regarding libertarian paternalism, "persuasion lies at the heart of much of soft paternalism, and it is not obvious that we want governments to become more adept at persuading voters or for governments to invest in infrastructure that will support persuasion."[7] We're on guard against manipulative business practices, but we don't want to feel that we have to be on guard against manipulative government practices as well—particularly those taken, presumably, for our own good. It's one thing for the government to play on our cognitive biases and heuristics to get us to recycle more and register for organ donation—both legitimate candidates for government action—but another thing altogether to do the same in the presumption of benefiting our personal interests.

Behavioral economists claim, however, that business don't cater to consumers' existing interests or preferences, but instead play a significant role in creating them. This is based on the concept of constructed preferences that we saw in chapter 2, which grounds Sunstein and Thaler's argument regarding the importance of context in decision-making, such as the argument that consumers in a cafeteria don't have fully formed preferences until they observe the ordering of food items. When focused on businesses, constructed preferences refer to the role that producers and sellers play in determining our tastes and desires. After all, you can't say that any consumer had an interest in drinking Coca-Cola before it was invented; instead, the Coca-Cola Company invented Coke and then convinced consumers that they wanted it. (Who wouldn't want The Real Thing once it was available?) According to behavioral economists and other scholars, businesses don't provide products to satisfy existing preferences—they create preferences for the products they make.[8]

To a certain extent this is obvious; the only people who wanted Coke before Coca-Cola was invented were only interested in one of its ingredients in particular! But people must have had an interest or a preference in drinking something that tasted like Coke, even if they couldn't articulate it. If Jerry doesn't like soda at all, no variety of Coke that the Coca-Cola Company tries to sell him will somehow cause a desire to sprout in him to drink it. But he may be a soda drinker who would like a Coke—the flavor of which is largely based on vanilla—with even more vanilla, but would never have realized that until the Coca-Cola Company came out with Vanilla Coke (or, as I call it, Coke Squared).

Of course, we can never know if Jerry had the preference for Vanilla Coke before it was introduced, but neither can we know that the Coca-Cola Company "created" his preference for it either. As I've stressed since chapter 1, we can never know anyone's true interests, and many of us have interests we're completely unaware of. But the claim that businesses create preferences or interests strikes me as odd for several reasons. First, what accounted for people's interests in less commercial times? One could certainly answer that they had less commercial interests—I almost walked right into that one—and one would certainly have a point. But one could also argue that people's interests in commercial goods were not realized until they were available—present but unrealized, rather than absent and then created. I'm sure people would have liked to have had iPads ten years ago; Apple created the product but didn't create the preference for it out of whole cloth. Second, it denies people any role in reflecting on their preferences and interests and either affirming them or rejecting them, making them seem like dogs who follow an attractive scent by instinct alone.[9]

It is a very cynical view of consumers that paints them as mindless automatons being manipulated into buying whatever corporations offer them. (If that were the case, we'd all be drinking New Coke now.) Certainly consumers can be manipulated to some extent, but products still have to fulfill some basic interest to capture consumers' attention (and dollars). It is those basic interests, not momentary preferences, that correspond to what matters to a person. To succeed, businesses respect these basic interests implicitly—and libertarian paternalists should do the same. (In the next chapter, we'll suggest some constructive ways in which they can do just that.)

### Government Could Care Less—and Should!

As I said above, we expect businesses to advance their own interests through whatever means they can (within legal and ethical bounds). We may criticize those means, or even lobby our elected representatives to regulate them or ban them, but we understand why they do them—to maximize their profit. Businesses know that consumers have their own interests, and they try to provide products that will serve those interests. Businesses don't care about advancing their customer's interests except insofar as they can make money helping this happen—and we wouldn't expect anything more. The idea of an employer manipulating its new employees into enrolling in 401(k) plans, or a company steering its customers into healthier products out of a concern for their health, is nearly as disturbing as the government encouraging it—because it's not an appropriate attitude for business, or government, to take toward individuals as customers or citizens.[10]

The attitude that is implied by paternalism in its ideal form is *care*, a sincere concern for someone's well-being. As individuals, we care for our loved ones: our family, friends, and significant others. But we care about them in a different way than we care for strangers (if we do at all). Although we should be kind to strangers, offering help when needed, those actions don't spring out of care except in a minimal sense: concern for a fellow human being. Perhaps Immanuel Kant, who maintained that we have a duty of kindness (or *beneficence*) to all, said it best: "When I say that I take an interest in this human being's well-being only out of my love for all human beings, the interest I take is as slight as an interest can be. *I am only not indifferent with regard to him*."[11] And the selective nature of our care extends past our own species as well: for example, most of us care for our pets to a greater degree than random animals we come across in the park.

What accounts for the difference in how we care about loved ones and strangers? In general, it is the emotional closeness between us that reflects how much people matter to us. Evolutionary psychologists would explain this in terms of group survival and kin selection, which is undoubtedly true, but on a conscious level these people's well-being and interests are important to us—and because of this they become part of *our* interests as well. Adam Smith, the father of economics (who was better known in his day as a moral philosopher), recognized that the sympathy we feel for each other naturally declines with social distance; and as we read earlier, Immanuel Kant acknowledged that our duty of beneficence is naturally expressed more intensely with regard to those closest to us.[12] More specifically, we *know* the people close to us well enough to have some idea what their well-being consists of and what their interests are. When our partners are upset, we have a good idea what will make them feel better. When our children succeed in school, we have a good idea what restaurant they'll want to go to for a celebration. When our best friends lose their jobs, we have a good idea what they'll need for comfort and consolation.

Note that I said we have a "good idea" about what our loved ones' interests will be. As well as we think we know our family and friends—certainly better than we know anyone else—we can never know them perfectly.[13] If Julie's best friend loses her job, she isn't going to rush in and take over her friend's life for her. She'll take some steps—such as calling or visiting, bringing her friend's favorite junk food—but then she's going to ask her friend what she can do for her, what she needs, what will make her feel better. In other words, Julie's going to ask her friend what *she* wants her to do and then (presumably) respect her wishes. And if Julie disagrees with something her friend asks for (such as helping her plot revenge against her ex-boss), she will discuss it rather than disregard it out of hand and do what she thinks will make her friend feel better instead. That's what friends are for![14]

In this way, Julie is not only caring for her friend but *respecting* her as well. She cares for her friend because she is close to her, and she knows her well enough to have a good idea about her well-being and interests. But above all, Julie respects what her friend wants her to do and does not impose her own ideas about what would be good for her. At most, if she disagrees with her friend, Julie tries to persuade her otherwise—but failing that, she falls back on respecting her friend's wishes. Once again, Immanuel Kant put it very well when he wrote that "I cannot do good to anyone in accordance with *my* concepts of happiness (except to young children and the insane), thinking to benefit him by forcing a gift upon him; rather, I can benefit him only in accordance with *his* concepts of happiness."[15] What Kant was saying is that care must be combined with respect, or care becomes condescending—and paternalistic.

Respect is an essential attitude—some would say *the* essential attitude—between individuals. Respect arises from the recognition that, however much we bond together into groups and rely on each other for help and support, deep down we are individuals with our own thoughts, feelings, and interests. In the ideal, we think of each other as important, as valuable, as worthy, possessing a dignity regardless of each other's relative social and economic status. A CEO may have little in common with the janitor in his building, but they should have a mutual respect for each other based on the fact that they're both human beings with worth and dignity.

Do the CEO and janitor care about each other? We can imagine circumstances in which they might—imagine that the CEO and janitor both started work at the company on the same day and formed a bond based on this. As the months and years pass, whenever they cross paths in the hallway they stop to chat. Maybe they find out they grew up in the same town, maybe even went to the same high school. And through these bonds and conversation they grow to be friends and care about each other. The CEO drops by with flowers when he hears that the janitor's wife is in the hospital, and the janitor may be the only person with whom the CEO can relax his professional facade when his mother dies.

Although this is an inspiring story—I see Tom Hanks in the role of the stern but sensitive CEO—somehow I don't see it being the norm. And that's OK. Most of us don't have the chance to get to know each other well enough to get close and truly care for one another, but we should respect each other all the same—and that respect implies that we shouldn't pretend we're close enough to care when we aren't. Most CEOs and janitors aren't close enough to each other socially to care about each other in this sense, but they don't need to be close to respect each other—their shared humanity is enough for that. Furthermore, they should recognize that they aren't close enough to care for each other, and rely on respect to govern their interactions. If the CEO hears that a janitor's wife is in the hospital, sending flowers is a kind and respectful thing to do—it doesn't assume too much familiarity, but expresses a universal concern for a fellow human being. It would awkward, however, if the CEO offered to house-sit for the janitor while he was in the hospital with his wife, watering the couple's plants, letting out the cat, and helping their kids with their homework. Maybe

Tom Hanks could pull it off, but we wouldn't expect this in the real world.

If we understand respect to be the attitude required of everybody based on our shared humanity, and care to be an appropriate attitude only for people who are close to each other and have some idea of each other's interests, then we can see the problem with paternalism.[16] The government, whether in the form of our elected representatives or appointed regulators, is in no position to express an attitude of care toward its citizens. We expect government to do certain things for us, which may include (depending on a person's political persuasions) enforcing the laws, building the roads, teaching the children, running the hospitals, and defending the country. Some of us want government to do more, others of us want government to do less, but we all want the government to do only those things that we want them to do. We don't give our government free rein to do whatever they think is in our interests. Few of us want our elected representatives to care about us in the sense that our mothers or grandfathers or aunts or partners do. That's what we have mothers and grandfathers and aunts and partners for. These people care about us—our government should respect us.

Not only are our elected representatives and their appointed bureaucrats not personally close to us—unless you're related to one of them, in which case your secret is safe with me—but more important, they don't know our interests. We express our interests through the ballot box, citizen referendums or initiatives, protests, and through personal communications. We expect them to respect those interests, which means considering them and balancing them with other interests of other citizens. That job is difficult enough! So why would the government need to or want to promote interests that we *haven't* expressed, interests that they presume on our behalf and then work to promote? Julie knows her best friend very well and cares for her very much, but still respects her wishes—and the government should do the same.

In an article about paternalism, philosopher Dan Brock describes the conflict between respect and care in terms of autonomy and well-being:

> One value is that of autonomy or self-determination, the interest persons have in making significant choices about their lives for themselves and in pursuing the courses chosen without interference from others. The other value is individual well-being or good which paternalistic action seeks to protect or promote. These values will be in conflict when the action chosen by the subject appears to be contrary to his well-being or good. The potential paternalist's alternatives then are either to respect the subject's autonomy and not protect his well-being or to infringe his autonomy by interfering to protect his well-being. So understood, the issue of paternalism requires a determination in any particular case of which value—autonomy or well-being—is more important or weighty.[17]

Any claim of a conflict between respecting a person's autonomy and caring for his or her well-being assumes that paternalist can know what contributes to a person's well-being. But we know that the paternalist has no way to know what constitutes a person's well-being (or interests), so the conflict disappears: assuming the government has benevolent intentions, there is no better way to promote people's interests then to allow them to make their own choices (as John Stuart Mill wrote in the quotation from chapter 4).

Ironically, value substitution denies individuals both respect *and* care. The denial of respect is more obvious, since the paternalist substitutes his or her idea of a person's interests for that person's own true interests. But also, since the interests promoted by paternalism are not a person's own, that person is not being cared for—some imaginary person with those interests is, the hypothetical perfectly rational person of myth! You don't care for someone by acting in your own idea of their interests; as Kant wrote, you can help them only by acting according to their interests as they express or otherwise reveal to you. Julie asks her friend what she needs, but paternalists don't do this. Their desire to help may be well-intentioned, but their methods are counterproductive at best and offensive at worst.

### It's Not (Just) What You Do, It's (Also) the Way That You Do It

Traditional paternalism contradicts its presumption of care through value substitution, but libertarian paternalism takes this counterproductive tendency to a different level entirely due to the way that nudges work. In addition to guiding people into decisions they would not necessarily have made themselves, nudges not only rely upon but also reinforce the very cognitive biases and dysfunctions they're supposed to correct for (as opposed to using them for other purposes, such as profit or public policy). In the process, they further diminish people's decision-making skills—which, ironically, will likely make nudges seem more "necessary" in the future.

How do nudges reinforce cognitive biases and dysfunctions? For the time being, let's give libertarian paternalists the benefit of the doubt and assume that the bad decisions that nudges prevent are actually judged bad by the decision-makers themselves. (That's a big assumption, but we don't need the disrespect of value substitution to make the points of this section.) However, by correcting people's mistakes before they can make them, nudges block the beneficial aspect of bad decisions: that we *learn* from them. Just because they result from cognitive biases and heuristics doesn't mean that such mistakes are inevitable, although some may be harder to correct than others. If people are allowed to make mistakes—which they can recognize as mistakes only if they actually stand in the way of their goals—they can then take actions to correct them, whether that means addressing their cognitive faults directly or finding ways to work around them.

Even Sunstein and Thaler realize this, writing in *Nudge* that the best way to help people "improve their performance is to provide feedback. Well-designed systems tell people when they are doing well and when they are making mistakes."[18] Some of the examples they cite as giving feedback just provide simple information, such as the "click" sound that digital cameras make to assure the user that the crucial umpteenth picture of little Billy in the sandbox has been preserved for eternity. But the concept of nudges itself betrays this intent, subverting the essential feedback that you get from bad decisions.

In a terrifically rich paper, legal scholars Jonathan Klick and Gregory Mitchell summarize this aspect of libertarian paternalism, drawing on psychologist James Byrne's model of self-regulation and learning-by-doing:

1. Paternalistic policies that restrict choice options restrict learning opportunities.
2. The noisier the learning environment, the more difficult to learn, and paternalistic policies introduce noise into, or mute feedback signals in, the learning environment.
3. The more extensive the paternalism imposed on citizens, the greater the cognitive hazard, due to restricted learning opportunities and more noise in learning environments.[19]

Klick and Mitchell recognize that learning from mistakes is not always easy and never guaranteed, but is made all that much harder by the "noise" added by yet another influence on decision-making, that of the nudge itself. Furthermore, not only can people not learn from their mistakes if they're not allowed to make them, but also they can't invest the time and effort in making their decision-making better if they don't even know they're doing it wrong.

One way we discover that we've make bad decisions is through their consequences. This is why responsibility and accountability are so important—not just in terms of morality and our obligations toward our fellow human beings, but in making our decision-making better. As Klick and Mitchell explain (quoting psychologists Jennifer Lerner and Philip Tetlock):

> Holding people accountable for their judgments and decisions can . . . move behavior toward the rational norm. "Predecisional accountability to an unknown audience will attenuate biases that arise from lack of self-critical attention to one's decision processes and failure to use all relevant cues." Thus, expecting to have to account for a choice may have positive effects on decision-making quality.[20]

Preventing people from making mistakes also prevents any immediate negative consequences from these decisions—which is a good thing, no doubt, as far as those consequences go, but it also blocks an important signal that the decisions are mistakes and the process of making those decisions was somehow flawed. We must also keep in mind that people use these same decision-making processes in many areas of their lives, not just the ones being monitored by libertarian paternalists. If people do not have the chance to improve their decision-making with regard to their diet or retirement savings, this compromised process of choice will also be used to make many other decisions in their lives. As philosopher Luc Bovens warns, "the cost of *Nudge* may be that we forgo the chance to gain the virtue of self-command."[21]

Furthermore, if people cannot learn from their mistakes, they will just keep making them—which, in the eyes of regulators, only justifies more paternalistic intervention! In this sense, the need for nudges becomes a self-fulfilling prophecy: regulators manipulate people into making better decisions, but since people never have the chance to learn how to make better decisions for themselves, they are locked in a cycle of dependency on the nudges themselves. As Klick and Mitchell write:

> Research on self-fulfilling prophecies warns that regulated parties are likely to become the weak decision-makers envisioned by paternalistic policy makers, as paternalistic regulations undercut personal incentives to invest in cognitive capital and the regulated parties conform to the expectancies of the paternalist.[22]

Furthermore, nudges can be potentially self-defeating. For example, automatic enrollment in retirement plans may lead to less voluntary savings or increased financial responsibility, and misunderstood nutritional information may lead to even more unhealthy eating (or excessive under-eating, exacerbating eating disorders). The fact is, even with the advances of behavioral research, we don't know if nudges will lead to the desired effect, much less the unintended consequences on other areas of decision-making and behavior.

Ironically, the self-fulfilling prophecy is not only to be expected on the part of the decision-maker who is blocked by nudges from learning to make better decisions. It also applies to the libertarian paternalistic regulators themselves, who are hardly immune from confirmation bias themselves:

> The perception of irrationality in the general public likewise leads to self-fulfilling behaviors within the paternalist himself, such as interpreting ambiguous evidence as evidence of irrational consumer behavior or engaging in strict review of disconfirming data and lax review of confirmatory data.[23]

In chapter 3, we mentioned that government regulators are no less prone to cognitive biases and dysfunctions than the rest of us, and in chapter 4 we emphasized how the success of nudges is measured by how well they generate the desired behavior, but in the above quotation Klick and Mitchell point out a specific way in which such factors may affect how they craft paternalistic regulations. Sunstein and Thaler's response to this is simple: What else are they supposed to do? Although they admit that bureaucrats are human, bounded rational, and subject to cognitive flaws, "nevertheless, as we have stressed, these human planners are sometimes forced to make choices, and it is surely better to have them trying to improve people's welfare rather than the opposite."[24] The error here, of course, lies in thinking there are only two options: paternalism or randomness, which ignores the wide range of neutral possibilities for choice architects to consider.

These cognitive errors on the part of regulators also contribute to the threat of libertarian paternalistic policies falling down a "slippery slope": policies are perceived to be successful, which prompts calls for more interventions, influencing more areas of choice, with greater degrees of manipulation, coercion, and value substitution.[25] Sunstein and Thaler anticipate this objection as well, offering three responses:

> First, in many cases there is simply no viable alternative to paternalism in the weak sense, and hence planners are forced to take at least a few tiny steps down that slope. Recall that paternalism, in the form of effects on behavior, is frequently inevitable. In such cases, the slope cannot be avoided. Second, the libertarian condition, requiring opt-out rights, sharply limits the steepness of the slope. So long as paternalistic interventions can be easily avoided by those who seek to adopt a course of their own, the risks emphasized by anti-paternalists are minimal. Third, those who make the slippery slope argument are acknowledging the existence of a self-control problem, at least for planners. But if planners, including bureaucrats and human resource managers, suffer from self-control problems, then it is highly likely that other people do too.[26]

The first two points were refuted earlier in this book: paternalistic intervention is not unavoidable, and opt-out rights are a red herring because nudges are designed to steer people away from such behavior, regardless of whether or not they "seek to adopt a course of their own." The third point is a shameless dodge: yes, everyone has self-control problems, but not everyone has the power to implement policies to guide other people's behavior. If Sally can't control her craving for cigarettes, that's unfortunate, but if her senator can't control his or her impulse to nudge people like Sally away from what the senator feels are bad decisions, that's more than unfortunate—that's scary. If anything, the potential of self-control problems among policymakers and bureaucrats should be regarded as an argument *against* paternalist intervention on their part, not for it.

As I wrote earlier, if the only tool you have is a nudge, every decision looks like a result of cognitive defect. Because libertarian paternalists have too narrow a view of human interests and too little respect for autonomy and dignity, they are far too eager to judge people's decisions to be mistaken and use nudges to "correct" them—and then use the success of their manipulations as confirmation that they were right all along. This twist is similar to the sayings, "the first sign that you have a problem is denial," and "if you're so innocent, why won't you admit you're guilty?" The foregone conclusion of irrationality further reinforces the point that libertarian paternalists are not out to make decision-making processes better, but rather to generate the choices they think are best.

The point of this section is to suggest another reason—besides respect for autonomy with regard to determining interests and making choices—to let people make choices without interference from regulators: it contributes to better decision-making through learning-by-doing enabled by reflection on mistakes and reinforced by accountability. If regulators were truly interested in helping people make better decisions, then they would take actions to help people counter their cognitive biases and heuristics (as we'll see in the next chapter). Instead, they take them as given and permanent, and then design nudges to exploit them—often using one cognitive defect to offset the result of another without considering the greater consequences for people's decision-making.[27] As legal scholar and federal judge Richard Posner argues, libertarian paternalists "treat the irrationalities that form the subject matter of behavioral economics as unalterable constituents of human personality. All their suggestions for legal reform are of devices for getting around, rather than dispelling, our irrational tendencies."[28] And as we saw earlier, leveraging people's cognitive biases and heuristics perpetuates them and ensures that nudges will always be in demand. Instead, if policymakers do anything, it should be to try to counter cognitive biases to improve decision-making. As legal scholar Jeffrey Rachlinski writes, "the role of individual learning and adaptation…cannot be ignored in assessing the need for paternalism. Simple experience might, in some contexts, be a much better cure for cognitive missteps than adopting a paternalist intervention."[29] If libertarian paternalists truly wanted to improve people's decision-making so they can make decisions that better serve their interests—whether or not the paternalists find their choices "unfathomable"—then they should do it.[30]

If libertarian paternalists want to see how paternalism should be done, look to those who deserve the name: parents. No responsible parent prevents his or her child from making mistakes. As much as it hurts, parents know that they have to let their kids make mistakes because that's the best way to learn. You can tell them not to jump off the porch or touch the hot stove or eat that entire 18-inch meat-lovers stuffed-crust pan pizza, but until they do it they won't fully understand why they shouldn't. The irony with government paternalism is that not only it is inappropriate—they don't even do it right!

### So How Can We Do It Better?

Let's step back for a minute. (Make sure you have room first.) We spent the first three chapters of this book setting up the background for libertarian paternalism, seeing how value substitution is a natural product of mainstream economics, behavioral economics, and behavioral law and economics. Then we spent the next three chapters (ending with this one) presenting various critiques of libertarian paternalism and nudges from informational, ethical, and practical perspectives. But what we have not yet done—but will do in the next and final chapter—is offer alternatives, better ways to tackle decision-making problems, including the option of not tackling them at all! We'll also see how these alternatives can improve on real-world "solutions" based on paternalistic policies and nudges, such as cigarette labeling, mortgages and credit cards, and food restrictions. In short, we'll see if we can find a way to help people make better choices, not just the choices someone else wants them to make—in other words, a respectful *and* caring alternative to paternalism of any kind.

# Chapter 7

## Why Choice Matters So Much—and What Can Be Done to Preserve It

*When you dig beneath the surface of the previous chapters in this book, everything comes down to autonomy: determining your own interests and making choices in pursuit of them. Not only do libertarian paternalism and nudges manipulate our choices, but more importantly, they claim to do so in our interests while furthering others. We've also seen that this disregard for people's true interests is a natural legacy of the way that both mainstream and behavioral economists think about decision-making: a deliberative process, however complex, guided by an overly simplistic goal. These simplistic goals allow economists to build complicated models of decision-making, but economists neglect to question whether the goals and interests assumed in their models correspond to what real people value. They focus on the process more than the goal, and they end up missing the forest for the trees. In the end, they presume to know what people's interests are and to act to promote those interests—which is the most distressing problem with libertarian paternalism and nudges.*

*In this final chapter we will explore why autonomy is so important, drawing on a range of philosophical traditions, and why libertarian paternalism is a particular threat to it, even more so than traditional paternalism. Then we'll suggest ways to achieve the goals libertarian paternalists claim to pursue—improving decision-making in people's own interests—without engaging in* *the value substitution and manipulation that we've criticized in this book. After all, people* do *make bad choices according to their own judgment, and it is worthwhile to think of ways to help them make better ones—without crossing the line and making people's choices for them.*

* * *

*Autonomy* is a word with many related meanings, each slightly different from the rest but all sharing a common thread: self-governance. We talk about countries and governments being autonomous or sovereign, determining their own laws and policies without taking orders from any other official body. Much of the controversy over the proper reach and jurisdiction of the United Nations and the International Criminal Court deals with the autonomy of nation-states who would be subject to their laws, rules, and regulations. This understanding of autonomy corresponds very closely to how we used it in this book, describing the right of individuals to govern their own lives, setting their own agendas, goals, and interests, and being free to make choices as they choose (provided those choices don't interfere with anyone else doing the same).

In philosophy, autonomy is most closely associated with the eighteenth-century German philosopher Immanuel Kant, from whom we've heard from time to time throughout this book. Kant's moral philosophy is based on the position that rational beings (such as us) can make choices without undue influence from others *and* independent of our own desires and preferences. We can choose to do the right thing even when we don't want to and even when other people pressure us to do otherwise. Resisting peer pressure, standing up to your boss, and taking a stand for a cause you believe in are all expressions of your autonomy with regard to external influence and authority, or what Kant called "outer freedom." Passing up dessert when you're on a diet, refusing to lie to your professor about why you missed an exam, and devoting your time to helping others are all expressions of your autonomy with respect to your own desires, which Kant called "inner freedom." Both inner and outer freedom reflect the general idea of self-governance: outer freedom means self-governance in the face of *external* pressure, whereas inner freedom means self-governance in opposition to *internal* pressure (self-governance in the sense of self-control).[1]

If you've read this book from the beginning, this should sound familiar: it was the basis for the expanded model of economic choice presented in chapter 1 as an alternative to the standard economic model. In my book *Kantian Ethics and Economics*, I explained how a focus on autonomy and the possibility of principled choice could enrich economic models of choice by recognizing that decision-making depends on not only preferences but principles as well.[2] But you don't have to agree with all of the implications of Kantian ethics to accept that choice can be based on principles, obligations, and ideals in addition to preferences, desires, and drives. That's why I didn't "reveal" the Kantian background to this model when I first introduced it in the first chapter on this book; I wanted the idea to stand on its own without the baggage that Kantian ethics brings with it. I think it does, and I hope you agree!

But we don't need even this much Kantian baggage to grasp the understanding of autonomy behind the critiques of libertarian paternalism in this book. Although the whole picture is useful when developing a more inclusive model of economic choice (and showing the shortcoming of the model used by mainstream and behavioral economists), we really only need to appreciate autonomy as self-governance and self-determination. But these concepts should be applied not only in terms of choice, but also in terms of the interests that motivate your choices in the first place. Although others can justifiably interfere in choices and actions—such as those that wrongfully interfere with other people's choices and actions—we are much more protective of our autonomy when it comes to determining how we each choose to live, the values we hold dear, and things we care about.[3]

One of the most basic values of modern democracies is *liberal neutrality*, the right of each person to decide for himself or herself what constitutes "the good life" with minimal interference from the government. (As we use the term here, "liberal" should be read not in the modern sense, as contrasted to "conservative," but rather in the traditional sense of "free"—which means vastly different things to today's liberals and conservatives!) As John Stuart Mill, another philosopher we've heard from throughout this book, wrote:

> As it is useful that while mankind are imperfect there should be different opinions, so it is that there should be different experiments in living; that free scope should be given to varieties of character, short of injury to others; and that the worth of different modes of life should be proved practically, when any one thinks fit to try them. It is desirable, in short, that in things which do not primarily concern others, individuality should assert itself.[4]

Of course, the government may interfere with actions that people take in pursuit of their individual views of the good life when those actions interfere with other people doing the same, but the government should remain neutral with respect to those views themselves. Liberal neutrality is an ideal, to be sure, and much more complicated to maintain in the real world, especially when a society attempts to be tolerant of views that question its foundational principles. We needn't invoke any controversial international conflicts to illustrate this; any question of constitutional interpretation raises issues about the nature or character of a country, and those on one side will inevitably accuse the others of trying to "destroy" it.

As you can imagine, the idea of character also ties in closely to autonomy and its importance. Just as the founding principles, traditions, and laws of a country contribute to its character, our personal principles, interests, and choice are all parts of our characters as persons. I am who I am largely because of the principles I adhere to, the things and people I care about, and the choices I make based on them. When I behave strangely, those who know me may say I seem "out of character," since my friends have come to know me by my actions and the principles and preferences they reflect.[5] In this sense, to respect people's autonomy is to acknowledge them as the people they are—and to deny people their autonomy is to strike at the core of who they are as individuals.

The word *character* also invokes a participant in a story, which suggests the metaphor of people as "authors" of their lives, an approach that many philosophers embrace. For example, as Christine Korsgaard writes, "from the practical point of view our relationship to our actions and choices is essentially *authorial*: from it, we view them as *our own*."[6] Another philosopher, Thomas Hill, writes of the value of being the author of your own life:

> The value of the life as author, when the story is one's own, is not seen as entirely derivative from the final content of the story, once finished, nor from the feelings experienced in living through it. Rather, living as the author, making the crucial choices, deciding what to count meaningful and what trivial, these are valued for their own sakes. This is not simply to say that one enjoys being the author and so values living derivatively as a necessary condition of such enjoyment; that puts the focus in the wrong place. Even if it is true that one *enjoys* living as author, one enjoys it partly because this expresses what one is and wants to be.[7]

Considering a person to be the author of his or her life also connects well with existentialist ideas about self-creation and authenticity (that have their roots in Kant's idea of autonomy). Existentialist philosophers such as Jean-Paul Sartre emphasize that as human beings we are responsible for constructing and maintaining our own characters in the face of social pressures based on how others see us (which Sartre calls "facticity").[8] Ultimately, each of us is responsible for remaining true to the person we make ourselves into.

No one says this is a simple task, of course; determining what kind of person you are going to be is an issue that we all must deal with, and is one with which many of us struggle. One thing that makes it particularly difficult is figuring out how to integrate our various roles and responsibilities, all of which are integral parts of our character. Drawing from both Kant and Plato in her book *Self-Constitution*, Christine Korsgaard describes how the roles and responsibility a person chooses to undertake, as well as how he or she balances them, determines who that person is:

> The task of self-constitution involves finding some roles and fulfilling them with integrity and dedication. It also involves integrating those roles into a single identity, into a coherent life. People are more or less successful at constituting their identities as unified agents, and a good action is one that does this well. It is one that both achieves and springs from the integrity of the person who performs it.[9]

This sounds circular, and to some extent it is—your character determines your actions, which in turn determine your character, and so forth. None of us does this perfectly, and we're all familiar with the feeling we get when we fail at keeping ourselves integrated, when our various roles and responsibilities pull us in different directions and we feel like completely different people at different times. But making decisions based on our interests, our core principles and preferences, helps us to stay integrated—to behave consistently in different situations—and authentic to the true selves that we've built up over time.

This section was a bit of a sprint through a lot of philosophy, but I hope the basic idea is clear: through our autonomy, *we are who we are.* Only by determining our interests and making our own choices to balance, reconcile, and pursue them, can we be true, complete persons. Each of us has the responsibility for who we are—even if you decide to conform to popular opinion, that is still a decision that you make.[10] No one else can determine who you're going to be—that is our ultimate freedom as human beings as well as our greatest responsibility. Furthermore, these decisions affect not only you but everyone around you, through your actions which reflect the principles that constitute your character and your identity. (Yes, just like the house that Jack built.)[11]

### Back to Nudge Once More

Of course, we did spend three entire chapters of this book critiquing the concepts of libertarian paternalism and nudges, but please indulge me just a bit more, now that we've spelled out the true value of autonomy. As I've tried to emphasize since chapter 1, the value substitution inherent in libertarian paternalism is its most ethically problematic aspect. The fact that regulators do not have access to information about people's true interests, and therefore must impose their own idea about them when designing paternalistic rules and policies, is the greatest threat to autonomy posed by libertarian paternalism. Finally, the way that nudges work—by piggybacking on the same cognitive biases and heuristics that motivated them—means that their effects on autonomy often go unnoticed, with strong implications regarding our identities and characters.

It is the combination of its subtle influence on our choices and value substitution that makes libertarian paternalism uniquely problematic. An example may help to show why traditional paternalism is less of a concern on these grounds. Remember Sally, our friend from chapter 4 who is trying to stop smoking? When she does buy cigarettes, she is certain to notice the warnings on the package advising her of the dangers to her health from smoking, and she is even more certain to notice the pinch she feels when paying for them, largely due to taxes meant (in part) to discourage smoking. These are both paternalistic policies, but they are hardly nudges. They don't play on any cognitive flaws in decision-making; rather, they influence through either persuasion or incentive.

To put it another way, both health warnings and taxes work by *engaging* Sally's decision-making process. She is free to make use of this information any way she chooses in pursuing her own interests, whatever they may be. The standard arguments against paternalism still apply; the government has no place influencing her decision-making by any means, including health warnings and taxes. But although these methods were designed with particular interests in mind—promoting health by discouraging smoking—they do not usurp Sally's interests since her decision-making processes are left virtually unaffected. She has more information to deliberate on, information that may or may not affect her final decision, but she is still free to make her own choice in her own interests without the government abusing any cognitive defects in her decision-making process.

Nudges, however, do exactly that—and their success, as we saw, depends precisely on how well they do that. Let's imagine that the city in which Sally lives passes a law that cigarettes can only be sold on Wednesdays. This does not prevent people from buying cigarettes, but it does play on self-control issues that may make it hard for some smokers to keep to a schedule of buying them every Wednesday. Libertarian paternalists would argue that they are not keeping people from buying cigarettes; they are simply using people's natural weakness of will or procrastination to help them quit (or at least cut down). If people *really* want to smoke, paternalists would argue, then they'll just buy them on Wednesdays. (People who want to smoke wouldn't have self-control problems with time management, I guess. Funny how that works.)

This is the particularly insidious way that nudges work: rather than engaging Sally's decision-making process like warnings or taxes do, the Wednesday-only rule circumvents it altogether, making use of her cognitive dysfunction to bypass conscious deliberation and generate the desired result. The irony is that nudges *won't work* if people think about them too much—if she does, Sally may decide not to play along and will instead make regular plans to buy cigarettes on Wednesday. Only if nudges are allowed to work covertly through flaws in people's decision-making processes can they produce the results sought by regulators: not better decision-making but "better" choices. But this depends critically on subverting the normal processes of decision-making that enable us to express our autonomy, the interests we hold dear and make choices in pursuit of. It is, as Kant said, to use people as mere means or tools; it is to treat them as faulty machines that need to be adjusted to obtain the results that paternalists want from them. (*For* them, of course, for them. My apologies.)

Furthermore, to the extent that our choices affect our character, nudges can have a frightening effect on who we are. In the last section, we described how choices and character can be considered to determine and reinforce each other in a constant process of self-creation and maintenance. But this assumes that all of our choices are indeed our own—and nudges force us to question this assumption and consider the implications of that. When we first met Sally in chapter 4, she was trying to quit. If she is better able to do that thanks to the incessant health warnings and high cigarette taxes, then they have simply enabled her choice, made in her own interests, to quit smoking. She can resent the government's paternalism on principle and at the same time be grateful that she was able to quit sooner because of it. However, if she had been nudged into quitting by being limited to buying cigarettes on Wednesdays—because, as the paternalists hoped, she kept forgetting to buy them on Wednesday—the choice to quit was not entirely her own, but was a result of the law working through her forgetfulness or procrastination.

Why is this a problem—she was able to quit either way, right? Yes, the result is good, and it is what she wanted, so her interests were promoted. But if she quit with the help of the unconscious effect of nudges, this choice does not reflect her character—but it will affect it. Choice and character are both aspects of our autonomy, feeding back on each other in our continuous circle. But the nudge introduced an outside element, one that did not arise autonomously from either her choice or her character—and now it's in the system, affecting the choice first, and then her character, which will go on to make more choices, further changing her character, and so forth. It is the very opposite of autonomy: outside influences are subtly introduced into her decision-making process, which then becomes part of her character. She will not likely be aware of this influence, so she will never have the chance to reflect critically upon it and either accept it or reject it depending on how it fits into the person she wants to be.

Such effects may also be seen in how well we integrate the various roles and responsibilities that are parts of our identities, as Christine Korsgaard identified in the quotation in the last section. If the Wednesday-only nudge helps Sally quit smoking—by reinforcing her weakness of will rather than reducing it—she may be surprised to find out she doesn't have the same resolve when it comes to dieting, doing her work in a timely fashion, or keeping up with her piano lessons. Philosopher Luc Bovens warns that a nudge potentially "leaves us with a fragmented self. We become incomprehensible to ourselves—why did we not act in line with our overall preferences or why is this kind of agency not resilient under non-*Nudge* conditions?"[12] Even if nudges do help us reach our goals, as they did with Sally, in a sense it is not we who are doing it—and that may have a strong impact the next time we make a similar effort without a nudge helping us. In times of trouble, a helpful push is always appreciated, but a controlling nudge may only make the problem worse.

### So What Should We Do Instead?

It is all well and good to criticize libertarian paternalism and the nudge concept, but it would be an empty critique that didn't suggest alternatives to address the problem. But wait a minute—what is the problem nudges claim to solve? This may be a case of a question with a false premise, like the classic trap, "Have you stopped cheating on your wife?" If in fact there is no problem to be solved, while there *are* problems with nudges themselves (as we've argued throughout this book), then the solution would simply be to stop nudging and rethink our approach to paternalistic regulation—or reconsider paternalism overall.

The problem, as Sunstein and Thaler explain, is "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," on which basis they "argue for self-conscious efforts, by private and public institutions, to steer people's choices in directions that will improve the choosers' own welfare."[13] As we've discussed at length, the solution they propose does not address the problem they identify, and instead can make the problem worse (along with other ethical concerns). We've acknowledged all along, however, that people *do* make choices that fail to promote their own interests—but only the people making the choices know if they were bad choices, and only they know what better choices would be.

So there are two approaches to this problem, depending on how you see the proper role of government. First, we could say simply that the government should do nothing—not because there is no problem, but because it's not the government's responsibility to solve every problem facing its citizens, especially problems that it may have little ability to address. Unlike Sunstein and Thaler's "libertarian" paternalism, this is a truly libertarian option that leaves the government free to handle what few responsibilities are relegated to it (such as national defense, a judicial system, and so on) and trust individual choices to individuals, no matter how sensible or foolish they may seem to anyone else. The basic point of this approach is that the government, above all, should respect the choices based by their citizens and not interfere with them except in cases of clearly involuntary choice (such as Mill's example of a man about to cross an unsafe bridge) or wrongful harm to others.

The second approach would appeal to those who see a more positive role for government, while acknowledging all the informational and ethical problems with nudges that we've discussed to this point. Here, we take on an idealized version of what libertarian paternalists claim to do: help people make better decisions, rather than help them make the choices libertarian paternalists want them to make. These are suggestions to improve people's decision-making processes and to enable people to make choices more reliably in line with their interests, *whatever* they may be and *whatever* the eventual choices turns out to be. In this sense, it shares with the "hands-off" approach above the respect for people's choices, the only difference being that the government takes steps to help improve its citizens' decision-making processes. What makes this different from what Sunstein and Thaler propose is that there is no concern whatsoever for the choices made. There is no value substitution or external judgment, nor is there any subtle manipulation that reinforces existing cognitive biases and heuristics. Instead, this approach sees the cognitive defects themselves as problems, not opportunities, and attempts to lessen their influence so people can make better choices in their own interests.

Here are just a few ideas for improving people's decision-making processes without steering their choices away from their own interests. Many of these naturally incorporate aspects of the critiques made in the last three chapters, but here we see how they apply to real-world examples such as cigarette and nutritional labeling, financial regulation, and food restrictions.

*Provide Information Neutrally.* One cause of bad decisions is insufficient or poor information, and the government often engages in campaigns and initiatives to provide people with more and better information. Health warnings on cigarettes, nutritional information on food packaging and restaurant menus, disclosure of financial terms of credit cards and home mortgages are examples of government mandates that allow people to make more informed decisions.

These interventions are supported by libertarian paternalists; as Dan Hausman and Brynn Welch note, Sunstein and Thaler "regard the following as nudges: educational campaigns, warning labels on cigarettes, requirements that firms notify employees of hazards, and signs warning people on a hot day to drink more water." But these are nudges in the ideal sense: "Unlike constraining someone or substituting your judgment for theirs, providing information and giving advice treats individuals as fully competent decision makers."[14] If these were the typical nudges, I would not have had to write this book! As we know, however, most of the nudges emphasized by Sunstein and Thaler (and others) embrace value substitution and encourage that certain choices be made, not simply better decision-making.

Even with laws and regulations meant to provide better and more complete information, there is still a danger of steering or guiding choice. There is a clear difference between the staid and clinical Surgeon General's warning found on cigarettes in the United States and the starkly emotional "Tobacco Kills!" prominent on cigarette packaging in the United Kingdom.[15] While the US labels provide information regarding the negative health effects of smoking—who knew?—the UK version hopes to prompt a visceral reaction from present and potential smokers. Obviously governments on both sides of the pond want to reduce smoking—for a variety of reasons, not all of them paternalistic—but the US approach is more respectful to the individual's choice.

We also see the potentially manipulative aspect of informational campaigns in the case of nutritional labeling. Traditionally, nutritional labeling, like the United States Surgeon General's cigarette warnings, are plain, boring, black-and-white boxes listing the calories in each serving as well as the amount of fat, protein, carbohydrates, sodium, and various vitamins and minerals. Serving sizes do not often reflect how much we eat in one sitting, but if we eat an entire bag of potato chips that supposedly contains two servings, it is not that difficult to double the calories, grams of fat, and so on. The no-frills presentation of nutritional information leaves the labels admirably neutral, giving consumers the information they want—or, if they don't, they're free to ignore it—while not steering them toward any particular dangers of concern to regulators.

But critics argue that nutritional labels *should* guide consumers' choices, because under the existing label design people are still making choices that the critics judge to be unwise. As bioethicist Ezekiel Emanuel wrote in the *New York Times*:

> What we need are simple, standardized icons that can be understood by a shopper in a second or less, located in a consistent place on all packages. No higher math or advanced nutrition knowledge should be required to grasp the icons' meaning. The information should reflect real serving sizes (canned soup labels regularly give the amount of salt for just half the can, trying to disguise that a whole can contain almost an entire day's intake of salt). And we should have interpretive symbols telling shoppers simply whether an item is healthy or unhealthy.[16]

What he may be suggesting—and what the F.D.A. has considered—is a "red light green light" labeling system like the one used in the United Kingdom: foods regarded by regulators as particularly high in fat or sodium, for instance, have these aspects highlighted in red on the labels to indicate their danger.[17]

The issue here is the same as with cigarette labeling: it is one thing to give the consumer information with which he or she can make better decisions, but another to indicate which information is more important for the consumer to consider. Even calorie counts on restaurant menus can imply a exaggerated importance to calories over more detailed nutritional information, providing an example of how the choice of *which* information the government chooses to provide can also give a nudge to consumers' choices. In this way, value substitution can infect efforts to provide better and more complete information—which is of particular concern with respect to nutritional labeling because health and food science is still rife with controversy, with different camps touting different health benefits of various aspects of foods. To the consumer, it can seem that experts' guidelines for best dietary practices change every day.[18] Even though the experts cannot agree on what we should eat, how much we should eat, and when we should eat it, they nonetheless presume to impose their judgment on consumers under the guise of merely providing better information.

This is why, if the government decides to help people make better choices by giving them information, it must be very careful to do so as neutrally as possible. Paternalistic manipulation is not inevitable, as Sunstein and Thaler claim, but it can be difficult to avoid even if the government's intentions are sound—and modest.

*Helping People Overcome Their Cognitive Biases and Heuristics.* As we've noted repeatedly, one ironic aspect of libertarian paternalism is that it relies on the same cognitive biases and heuristics that motivate it in the first place. Advocates of nudges argue that these deviations from textbook rationality make it uncertain that people will make choices in their own interests, but then use these same decision-making flaws to change their behavior rather than addressing the flaws themselves. But if policymakers are sincere about improving people's decision-making, they should focus on the processes by which they make choices rather than the choices themselves.[19]

It may sound simplistic, but in many cases it can help to encourage people to take more time making a decision. Many bad choices may stem from hurried, impulsive decision-making, and simply slowing down and thinking things through may result in choices that better promote well-being. This is the reasoning behind the "cooling-down periods" that some cities and states place on significant decisions such as buying a house or car and getting married or divorced. In their law review article "Regulation for Conservatives," behavioral economist Colin Camerer and his colleagues recommend cooling-off periods in cases when decisions are often made in emotionally "hot" situations and are difficult to reverse later.[20] To combat the compromised state of people's decision-making in such cases, they can either be required to wait a certain amount of time before finalizing a choice (often used in family law), or they can be given a certain amount of time after making a choice during which they can reverse it (such as with major purchases). Camerer and his colleagues consider this to be "asymmetric paternalism" since people who made sound, rational choices will not be affected significantly by the cooling-off period, but it may help those who did make rash, emotional decisions.[21] (I'd hate to think of a decision to get married that wasn't somewhat emotional, though!)

As with providing information, cooling-off periods can potentially be taken too far and involve value substitution. A government could easily use cooling-off periods selectively, not to encourage better decision-making but to discourage certain behavior altogether, especially with politically charged decisions such as divorce or abortion. The mere suggestion that a decision should be reconsidered could be interpreted by some as a nudge toward making a different one, and measures should be taken to reinforce that this is not the case (if indeed it isn't). Ironically, Camerer and his colleagues regard cooling-off periods to "appear more intrusive than our earlier policies" such as automatic enrollment in retirement plans; while they may be more overtly coercive, they are less likely to involve value substitution (if done right) and therefore less paternalistic.

Emotionally charged decisions often resemble those made in times of weakness of will, such as continued smoking, drinking, or overeating while trying to quit them. People are often quite willing to admit self-control problems in these contexts, and many voluntarily sign up for programs like Weight Watchers or Alcoholics Anonymous. This, of course, acknowledges that decision-makers are the best judges of their own good or bad choices and governments may provide or subsidize such programs, provided they don't go so far as to nudge people into them. Again, the most important thing is to respect people's choices while providing opportunities to make better ones for those who judge their choices to be poor, not judge their decisions to be poor for them.

Careful deliberation can also help combat some cognitive biases and heuristics that distort our valuations of options and our processing of information (as summarized in chapter 2). Legal scholar and psychologist Gregory Mitchell suggests that, if the framing of options is affecting the way people make decisions, they can be encouraged to reframe them in a way that makes it easier for them to make an unbiased choice. Asking decision-makers to state their reasons for making a particular choice can also help prevent choices made from unconscious biases.[22] Consumers can critically assess offers from merchants and lenders, demand that complex terms be spelled out in a more straightforward fashion, and refuse to deal with companies that use manipulative practices. Calls for increased government regulation are common in such cases, but it is all too easy to forget the power that consumers have to vote with their dollars, as shown by the wider availability of healthier and more environmentally friendly products in response to consumer demand. This avoids simplistic and overbroad "solutions" to problems that foreclose too much choice and make too many decisions for consumers based on imposed interests and value substitution.

We've seen this done the wrong way in the case of financial regulation in the United States, particularly with regard to mortgages and credit cards in response to the crisis that began in 2007. In an influential law review article, legal scholars Oren Bar-Gill and Elizabeth Warren argue for limitations on loan terms based on the belief that "consumers make systematic mistakes in their choice of credit products and in their use of these products. These observed mistakes indicate the existence of deficits in either information or rationality—or both."[23] But these "mistakes" can be "observed" only if the observer assumes a simplistic goal of financial optimization among all consumers. For instance, they cite studies that show that many consumers who receive a credit card with a low introductory "teaser rate" do not switch to another card as soon as the rate rises, and many consumers also pay relatively high credit card interest rates while holding cash in savings accounts earning lower interest.[24] They also criticize consumers for taking advantage of low introductory teaser rates in the first place—so consumers are regarded as irrational for accepting the low-rate credit cards and then irrational *again* for not dumping them soon enough.[25] This is a fine example of regulators' willingness to see irrationality wherever they can, as we discussed in the last chapter.

But such behavior on the part of financial consumers is as easy to explain as our friend Patrick's choice to eat a muffin on Sunday mornings. There are many reasons to sign up for a credit card with a low introductory rate, most obviously reasons to do with financial hardship or short-term needs such as paying to relocate for a new job. The refusal of consumers to bounce from one low-teaser-rate card to another may be an admirable exercise of self-control, if they know they will overspend once they switch to a cheaper card, or may arise out the simple recognition that they don't need the cheap credit any longer. And economists have long recognized and endorsed consumers' preference to have ready funds on hand for emergencies, even when they could use it to pay expensive credit cards off a little faster. (They even have a term for it: *liquidity demand*.)

Bar-Gill and Warren, however, see these choices as failures of rationality and propose rules and regulations that will limit the credit products available to everybody because of the "mistakes" of some—mistakes that they *choose* to interpret as such. Along the same lines, legal scholar Michael Barr, economist Sendhil Mullainathan, and psychologist Eldar Shafir suggest that lenders "be required to offer eligible borrowers a standard mortgage (or set of mortgages), such as a fixed rate, self-amortizing 30-year mortgage loan, according to reasonable underwriting standards. The precise contours of the standard set of mortgages would be set by regulation," with an opt-out option for borrowers.[26] (Sound familiar?) Rather than encouraging more thoughtful deliberation and critical examination of financial products, these scholars and regulators prefer to make choices for people, relieving them of the burden of making difficult choices—and responsibility for the consequences of them.

*Holding People Accountable and Responsible for Their Choices.* As we saw in the last chapter, accountability and responsibility are important not just for moral reasons, but they also contribute to better decision-making through learning-by-doing. In other words, if people reasonably expect not to be held accountable and responsible for their actions, they will spend less time and effort trying to make the best decision, including being less careful to consider the various cognitive biases and heuristics they may be operating under. Economists have recognized a similar effect for years: *moral hazard*, wherein people rationally take more risks if they know any losses will be covered by insurance or government bailouts. A related effect, *cognitive hazard*, occurs when people have less incentive to carefully assess the risks they take—and the decisions they make. In the case of risky behavior encouraged by insurance and bailouts, we don't have to assume that people opportunistically "take advantage" of these safety nets (though this definitely contributed to our various financial crises over the years, especially the savings-and-loan failures of the 1980s). Rather, people could merely—and possibly unconsciously—spend less time carefully assessing the risks of various options because they know that any losses will be covered. But if people knew they would held responsible for any negative consequences of their risky behavior—by ruling out bailouts, limiting insurance payouts, or increasing insurance premiums—they will have a direct incentive to spend more time making decisions carefully and more deliberately.

This effect is seen in the financial industry, which we discussed in the last section, on behalf of both the lenders—who can combine and sell risky mortgages to other parties so their failure will not reflect directly on the original lender—and the borrowers—who can apply for bankruptcy or government-sponsored refinancing. Both of these "outs," one enabled by government negligence and the other through government action, represent failure to hold people accountable for their decisions and therefore reduce the incentive to make careful ones in the first place. If the original lenders were held to a greater degree of responsibility for excessively risky loans (through a more careful investigation of the trail of securitization), and borrowers were similarly held responsible for their behavior (through tighter restrictions on bankruptcy and stricter guidelines on refinancing), those on both side of the mortgage transaction would have incentive to think more carefully about their decisions and would likely make better ones for it. Unfortunately, the policies recommended by the scholars cited in the last section move in the opposite direction, choosing instead to make decisions for consumers while doing little to address responsibility.

Moral and cognitive hazards are also seen in a different form in the health care area, wherein the degree of responsibility people must take for the costs of their health care has little to do with their behavior. As often mandated by law, individuals' health insurance premiums depend very little on behavior such as proper eating and exercise, and normally only take smoking into account. Insurance programs, whether private or public, with low co-pays or deductibles significantly cut the link between behavior and costs, leaving very little incentives for people to think about the financial consequences of their health decisions. Instead of holding people more accountable for their health-related behavior, regulators faced with growing health care costs take the opposite approach: limiting choice. They ban foods seen as particularly unhealthy (such as trans fats), limit smoking in most public places (also justified, of course, by the effects of secondhand smoke), and nudging people into healthier behavior—none of which encourages better decision-making.

For example, New York City mayor Michael Bloomberg has won accolades and criticism for his initiatives to promote healthier lifestyles in his city, including all of the examples listed in the previous sentence. In the spring of 2012, he proposed a 16-ounce limit on sodas, lattes, and other sugary drinks sold in certain venues, such as delis, restaurants, sporting events, and movie theaters (but not grocery stores and convenience stores). (This is not as arbitrary as it sounds: the covered venues are those regulated by the city, and the rest are regulated by the state.) Although this may be a bit heavy-handed to be considered a true nudge, it does not prohibit people from buying a second drink or taking advantage of free refills. It discourages certain choices without foreclosing them altogether, relying on a little extra burden and the people's natural laziness to reduce overall soda consumption—*that* sounds like a nudge. As reported in the *New York Times*, Mayor Bloomberg went so far as to mock the very inconvenience that drives the plan itself:

> "Your argument, I guess, could be that it's a little less convenient to have to carry two 16-ounce drinks to your seat in the movie theater rather than one 32 ounce," Mr. Bloomberg said in a sarcastic tone. "I don't think you can make the case that we're taking things away."[27]

The extremely limited effectiveness of this ban, given the many other venues from which people can buy large sodas—including 7–11 stores, home of the infamous Big Gulp—and the likely adaptations on the part of both consumers and affected businesses (such as we described in the last chapter) to get around the ban, make this attempt at promoting healthier behavior even more insulting. Not only is choice being manipulated in a laughably clumsy manner, but it will most likely fail in terms of its stated goal.

It is reasonable to assume that the mayor's intent was not to make his citizens healthier but to reduce public medical costs. Although understandable given the state of medical care in the United States, this reinforces our larger point: accountability and responsibility are much more effective and respectful ways to generate good decision-making. Let's say that Paula, a proud lifelong citizen of New York City, agrees that drinking too much soda is bad for her, but she finds herself unable to resist the offer to order the jumbo drink at her local pizza place. Which plan is likely to be more effective to helping her curtail her addiction to high-fructose corn syrup in fizzy liquid form: a ban on large-sized cups that is easily circumvented, or assuring her that she will be held responsible for any medical costs linked to her soda habit, including an immediate increase in her medical insurance premiums or exclusion from certain public health programs? I think the second option—especially if the additional costs are realized *now*—will be much more likely to prompt some serious thinking on Paula's part regarding how much she values soda. And while most behavioral economics and libertarian paternalists will see this as an empirical question—which, admittedly, is how I stated it—more importantly it is an ethical question focused on which option is more respectful of choice. Not surprisingly, the option based on accountability and responsibility wins hand down.

### Choice Must Be Respected and Sound Decision-Making Should Be Encouraged

That's the bottom line, isn't it? Given the informational, ethical, and practical problems with nudges, the best things for economists and policymakers to do are: (1) respect the choices made by people and (2) hold them accountable and responsible for the consequences of those choices. In addition, they can (3) make sure they are not doing anything that would discourage good decision-making, and (4) possibly take careful and neutral actions to promote it.

Although this focus on individual choice and responsibility sounds very libertarian—unlike the "libertarian" paternalism behind nudges—you do not have to be libertarian in your politics to agree with the points made in this book. You can support very strong government involvement in many aspects of life, including the economy, health care, education, or the environment, all of which involve personal and political decisions that affect us all in significant ways. The only government programs that I take issue with in this book are paternalistic policies in general and nudges in particular. I think that most people of any political persuasion, regardless of their views concerning the appropriate scale and scope of government, agree that their choices, insofar as they don't have any direct and wrongful effect on anybody else, are their own business. Choices like these are of no legitimate concern of economists or policymakers, and they have no right to try to change them, especially by using our natural cognitive biases and heuristics.

In the end, there are many more pressing problems in the world than what we choose to eat for dessert or how we choose to spend or invest our money. Economists and policymakers should focus on the big issues, and leave our lives to us. If they really want to help people—and I trust they do—that's the best way to do it.

# Notes

## 1 The Problems with Traditional Economic Models of Choice

1. For the basics of this model, see any microeconomics textbook, and for a more thorough, critical discussion, see Shaun Hargreaves Heap, Martin Hollis, Bruce Lyons, Robert Sugden, and Albert Weale, *The Theory of Choice: A Critical Guide* (Oxford: Blackwell, 1992). There are also many popular books, such as Steven Levitt and Stephen Dubner's *Freakonomics* (New York: Harper, 2005) and Steven Landsburg's *The Armchair Economist* (rev. ed.; New York: Free Press, 2012), which masterfully illustrate the insights of this model when applied to many puzzles of everyday life.
2. Commitment was introduced into the economics discussion by economist and philosopher Amartya Sen, who won the Nobel Prize for his integration of philosophical concepts into economic analysis. See his paper "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory" in his collection *Choice, Welfare and Measurement* (Cambridge, MA: Harvard University Press, 1982), pp. 84–106. I also strongly recommend his short but powerful book *On Ethics & Economics* (Oxford: Blackwell, 1987) for an introduction to incorporating ethics into economics and the challenges it presents. His work was a tremendous influence and inspiration for my own efforts in the field of ethics-and-economics, including the book you're holding now (and this chapter in particular).
3. On the issue of modeling moral principles as preferences or constraints, see Robert S. Goldfarb and William B. Griffith chapters "Amending the Economist's 'Rational Egoist' Model to Include Moral Values and Norms, Part 1: The Problem" and "Amending the Economist's 'Rational Egoist' Model to Include Moral Values and Norms, Part 2: Alternative

Solutions" in Kenneth J. Koford and Jeffrey B. Miller (eds.), *Social Norms & Economic Institutions* (Ann Arbor: University of Michigan Press, 1991), pp. 39–84. Some economists go even further and claim that principles are *lexicographic preferences*, which must be satisfied before any others. In the example of lying, the "preference" not to lie is the highest one, and must be satisfied before any others can considered. This eliminates the problem of trade-offs by ruling them out, but then the question becomes: Why call it a preference if it doesn't allow for trade-offs (one of the chief characteristics of preferences) and doesn't "act" like the commonsense idea of a preference? Just call it what it is: a constraint.

4. We'll see in later chapters that these self-destructive preferences contribute to the calls for paternalism to change people's choices.
5. For more on the problems with preferences with respect to well-being, see philosopher Daniel Hausman's recent book *Preference, Value, Choice, and Welfare* (Cambridge: Cambridge University Press, 2011), especially chap. 7. Together with economist Michael McPherson, Hausman also wrote a marvelous primer for ethics and economics titled *Economic Analysis, Moral Philosophy, and Public Policy*, 2nd ed. (Cambridge: Cambridge University Press, 2006).
6. As Amartya Sen wrote critically, a person's set of preferences "is supposed to reflect his interests, represent his welfare, summarize his idea of what should be done, and describe his actual choices and behavior" ("Rational Fools," p. 99).
7. On the several meanings of utility in economics, see John Broome, "Utility," in his collection *Ethics Out of Economics* (Cambridge: Cambridge University Press, 1999), pp. 19–28.
8. Never give up on a good metaphor, I always say.
9. For one particularly good discussion of moral dilemmas from the viewpoint of virtue ethics, see philosopher Rosalind Hursthouse's book *On Virtue Ethics* (Oxford: Oxford University Press, 1999), chap. 3 ("Irresolvable and Tragic Dilemmas").
10. For more on economist's attempts to model weakness of will, particularly with respect to procrastination, see my chapter "Resisting Procrastination: Kantian Autonomy and the Role of the Will" in Chrisoula Andreou and Mark D. White (eds.), *The Thief of Time: Philosophical Essays on Procrastination* (Oxford: Oxford University Press, 2010), pp. 216–232.

11. This could very well be me, if I were named Fred, but I'm not.
12. In the philosophical lingo, Searle is questioning *psychological* determinism, not *physical* determinism. See his *Rationality in Action* (Cambridge, MA: MIT Press, 2003), or my chapter on procrastination cited in note 10.
13. Searle, *Rationality in Action*, p. 17.
14. Philosopher Thomas E. Hill, Jr., draws the connection between weakness of will and self-respect in his paper "Weakness of Will and Character," included in his collection *Autonomy and Self-Respect* (Cambridge: Cambridge University Press, 1991), pp. 118–137.

### 2 How Behavioral Economics Makes the Same Mistakes

1. The latter are explored in more detail in my book *Kantian Ethics and Economics: Autonomy, Dignity, and Character* (Stanford, CA: Stanford University Press, 2011), especially Chapters 1 and 2.
2. Thorstein Veblen, "Why Economics Is Not an Evolutionary Science," *Quarterly Journal of Economics* 12 (1898): 373–397, at p. 389. Veblen's eloquent characterization of the traditional model has echoes in contemporary philosophy: take R. Jay Wallace, for instance, who calls the standard model of choice (in philosophy as well as economics) the "hydraulic conception," which "pictures desires as vectors of force to which persons are subject, where the force of such desires in turn determines causally the actions the persons perform" (*Normativity & the Will*, Oxford: Oxford University Press, 2006, p. 172).
3. Herbert Simon, "A Behavioral Model of Rational Choice," *Quarterly Journal of Economics* 69 (1955): 99–118, at p. 99.
4. If you like homework—and who doesn't like homework—you can look at Chris and Pat's search for a car from chapter 1 in the same light. (Bonus points for including city *and* highway mileage.)
5. Social theorist Barry Schwartz makes much the same point about the complexity of decision-making when choices are plentiful, describing how the burden of decision-making in such contexts can be self-defeating in terms of happiness and well-being, in his book *The Paradox of Choice: Why More Is Less* (New York: Harper, 2004). The dangers of "too much

choice" are also a recurring theme in the work of psychologist Sheena Iyengar, in particular her book *The Art of Choosing* (New York: Grand Central, 2010).

6. The original academic sources for many cognitive biases and heuristics are collected in Daniel Kahneman, Amos Tversky, and Paul Slovic (eds.), *Judgment under Uncertainty: Heuristics and Biases* (Cambridge: Cambridge University Press, 1982). Other useful academic collections include Kahneman and Tversky (eds.), *Choices, Values, and Frames* (Cambridge: Cambridge University Press, 2000) and Thomas Gilovich, Dale Griffin, and Daniel Kahneman (eds.), *Heuristics and Biases: The Psychology of Intuitive Judgment* (Cambridge: Cambridge University Press, 2002). Kahneman's latest book, *Thinking, Fast and Slow* (New York: Farrar, Straus and Giroux, 2011), gives a terrific overview of his work with Tversky and others.
7. On second thought, opposite do attract . . . hmm.
8. Dan Ariely's book *Predictably Irrational: The Hidden Forces That Shape Our Decisions* (New York: Harper Perennial, 2010) masterfully presents many more examples of cognitive biases and heuristics, many drawn from Ariely's own life.
9. Other contribution to the popular behavioral economics field (some by journalists or popular science writers rather than academics) include Jonah Lehrer, *How We Decide* (Mariner, 2010); Ori Brafman and Rom Brafman, *Sway: The Irresistible Pull of Irrational Behavior* (Broadway Books, 2009); and Ariely's follow-up, *The Upside of Irrationality: The Unexpected Benefits of Defying Logic* (New York: Harper Perennial, 2011).
10. Thaler also made important contributions to consumer behavior using the same approach: for his important work in these fields, see his books *The Winner's Curse: Paradoxes and Anomalies of Economic Life* (Princeton, NJ: Princeton University Press, 1994) and *Quasi-Rational Economics* (New York: Russell Sage Foundation, 1994).
11. Cass Sunstein (ed.), *Behavioral Law & Economics* (Cambridge: Cambridge University Press, 2000).
12. Paul Slovic, "The Construction of Preference," *American Psychologist* 50, no. 5 (1995): 364–371. For another overview of some of the studies contributing to this view of preferences, see Amos Tversky and Richard Thaler, "Anomalies: Preference Reversals," *Journal of Economic Perspectives* 4, no. 2 (1990): 201–211.
13. Slovic, "Construction of Preference," p. 365.

14. Tversky and Thaler, “Anomalies: Preference Reversals,” p. 209.
15. Christine Jolls, Cass R. Sunstein, and Richard H. Thaler, “A Behavioral Approach to Law and Economics,” in Sunstein (ed.), *Behavioral Law & Economics*, pp. 13–58, at pp. 14–16. We’ll have reason to revisit this paper in the next chapter.
16. The 2010 edition of *The Theory of Moral Sentiments* from Penguin (New York) includes an introduction by Amartya Sen, who reintroduced the concept of sympathy into the economic discussion in his 1977 paper “Rational Fools: A Critique of the Behavioral Foundations of Economic Theory” in his collection *Choice, Welfare and Measurement* (Cambridge, MA: Harvard University Press, 1982). (This is the same paper cited in the last chapter; I’d read it if I were you. I’ll wait.)
17. Examples of this include Howard Margolis, *Selfishness, Altruism, and Rationality* (Chicago: University of Chicago Press, 1984) and Amitai Etzioni, *The Moral Dimension: Toward a New Economics* (New York: Free Press, 1990). Would you believe me if I told you the idea of multiple preference rankings was also introduced by Amartya Sen in his “Rational Fools” paper? I told you that you should read it. (Maybe next time you’ll listen.)
18. Matthew Rabin, “Incorporating Fairness into Game Theory and Economics,” *American Economic Review* 83 (1993): 1281–1302.
19. Ibid., p. 1287.
20. What about judgment, you ask? As with the traditional economic model of choice, there is no need for judgment in behavioral economics model because preferences—although more elaborate than in mainstream models—always provide “the answer.”
21. George Ainslie, *Breakdown of Will* (Cambridge: Cambridge University Press, 2001).
22. For Ainslie’s contribution to procrastination, see his chapter “Procrastination: The Basic Impulse,” in Chrisoula Andreou and Mark D. White (eds.), *The Thief of Time: Philosophical Essays on Procrastination* (Oxford: Oxford University Press, 2010), pp. 11–27.
23. See Ainslie, *Breakdown of Will*, chap. 3.
24. On salience of present costs, see (Nobel laureate) George Akerlof, “Procrastination and Obedience,” in his collection *Explorations in Pragmatic Economics* (Oxford: Oxford University Press,

2005), pp. 209–231; on present-biased preferences, see Ted O'Donoghue and Matthew Rabin, "Doing It Now or Later," *American Economic Review* 89 (1999): 103–124 (the first of several papers by O'Donoghue and Rabin on procrastination).

25. For a more formal explanation of this approach to modeling procrastination, see my chapter "Resisting Procrastination: Kantian Autonomy and the Role of the Will" in Andreou and White (eds.), *Thief of Time*, pp. 216–232.

### 3 How Behavioral Economics Met Law and Economics and Begat *Nudge*

1. For an excellent overview of the field, see Nicholas Mercuro and Steven Medema, *Economics and the Law: From Posner to Post-Modernism and Beyond*, 2nd ed. (Princeton, NJ: Princeton University Press, 1997).
2. I discuss this at more length in "Social Law and Economics and the Quest for Dignity and Rights," in John B. Davis and Wilfred Dolfsma (eds.), *The Elgar Companion to Social Economics* (Cheltenham, UK: Edward Elgar, 2008), pp. 575–594.
3. Gary Becker, "Economic Way of Looking at Life," Nobel Lecture, December 9, 1992, available at http://home.uchicago.edu/gbecker/Nobel/nobellecture.pdf, p. 41.
4. Oliver Wendell Holmes, "The Path of Law," *Harvard Law Review* 10 (1897): 457–478, reprinted in Richard Posner (ed.), *The Essential Holmes* (Chicago: University of Chicago Press, 1992), pp. 160–177. For a look at what Holmes's "bad man" implies about law and economics, see Robert D. Cooter, "Models of Morality in Law and Economics: Self-Control and Self-Improvement for the 'Bad Man' of Holmes," *Boston University Law Review* 78 (1998): 903–930.
5. This way of describing the burglar's dilemma—and the various examples of weakness of will from chapter 1—bear some resemblance to philosopher Harry Frankfurt's work on desires and the will. According to Frankfurt, a person has *first-order desires*, based on simple preferences, and *second-order desires* regarding his or her first-order desires, or "preferences on preferences." For example, the burglar wants to rob a house, but he doesn't *want* to want to rob a house. If he just gives up and succumbs to his first-order desires and steals, he is a "wanton," but if he can align his action with his second-order desires and not steal, he is a "person." To put it simply, a person can control himself or herself, and is not subject to

his or her desires; in the terms I used in chapter, a person can put principle above preferences when choosing an action. See Frankfurt, "Freedom of the Will and the Concept of a Person," in his collection *The Importance of What We Care About: Philosophical Essays* (Cambridge: Cambridge University Press, 1988), pp. 11–25. As you probably guessed, traditional economic models apply to wantons (such as Holmes's bad man). Some economists, however, have dealt with different levels of preferences or desires, such as Amartya Sen in "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory," in *Choice, Welfare and Measurement* (Cambridge, MA: Harvard University Press, 1982), pp. 84–106, and David George in his book *Preference Pollution: How Markets Create the Desires We Dislike* (Ann Arbor: University of Michigan Press, 2004).

6. Russell Korobkin, "What Comes after Victory for Behavioral Law and Economics?", *University of Illinois Law Review* 2011 (2011): 1653–1674, at p. 1655.
7. This paper was originally published in the *Stanford Law Review* 50 (1998): 1471–1550, and was soon reprinted as the lead chapter in Sunstein's influential collection *Behavioral Law & Economics* (Cambridge: Cambridge University Press, 2000), pp. 13–58. (All citations of the paper will refer to the page range of the book version.) It is the most frequently cited law review article between 1995 and 2010 by a wide margin; the second most frequently cited article is also a foundational piece of behavioral law and economics scholarship, Russell B. Korobkin and Thomas S. Ulen, "Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics," *California Law Review* 88 (2000): 1051–1144. (The rankings are reported in Korobkin, "What Comes after Victory," pp. 1653–1655, during one of his victory laps.)
8. Coase's original work that led to the Coase theorem (which was formalized and named by another Nobel-winning economist, George Stigler) appeared in his 1960 paper "The Problem of Social Cost," reprinted in his collection *The Firm, the Market, and the Law* (Chicago: University of Chicago Press, 1990), pp. 95–156. Incidentally, a study published in 2012 found "The Problem of Social Cost" to be *the* most cited law-related paper ever. (Holmes's "The Path of the Law" was number 3, and Jolls, Sunstein, and Thaler's 1998 paper was number 57.) See Fred R. Shapiro and Michelle Pearse, "The Most-Cited

Law Review Articles of All Time," *Michigan Law Review* 110 (2012): 1483–1520.

9. Transaction costs were key not only to the Coase theorem but also to Coase's theory of the firm, as described in his 1937 paper "The Theory of the Firm," also included in *The Firm, the Market, and the Law* (pp. 33–55); he was awarded the Nobel Prize in recognition of both ideas.
10. See Jolls, Sunstein, and Thaler, "Behavioral Approach," pp. 27–29.
11. Richard A. Posner, "Rational Choice, Behavioral Economics, and the Law," *Stanford Law Review* 50 (1998): 1551–1575, at p. 1552. For similar arguments in defense of assumptions of rational behavior in economics, see Richard B. McKenzie, *Predictably Rational? In Search of Defenses for Rational Behavior in Economics* (Dordrecht: Springer, 2010).
12. For a survey of such criticism of behavioral law and economics, see Grant M. Hayden and Stephen E. Ellis, "Law and Economics after Behavioral Economics," *Kansas Law Review* 55 (2007): 629–675, at pp. 653–657.
13. Legal scholar and psychologist Gregory Mitchell is well-known for these criticisms; see his papers "Taking Behavioralism Too Seriously? Unwarranted Pessimism of the New Behavioral Analysis of Law," *William and Mary Law Review* 43 (2002): 1907–2021, and "Why Law and Economics' Perfect Rationality Should Not Be Traded for Behavioral Law and Economics' Equal Incompetence," *Georgetown Law Journal* 91 (2002): 67–167. On the biases introduced by using college students as test subjects—because they're literally WEIRD (Western, Educated, Industrialized, Rich, and Democratic)—see the papers in the journal *Behavioral and Brain Sciences* 33 no. 2–3 (2010).
14. Hayden and Ellis, "Law and Economics after Behavioral Economics," p. 632.
15. Jolls, Sunstein, and Thaler, "Behavioral Approach," pp. 46–47. As my nine-year-old daughter might put it, they like paternalism, but they don't *like* like paternalism.
16. A decade earlier, Cass Sunstein wrote about these problems with preferences in his law review article "Legal Interference with Private Preferences," *University of Chicago Law Review* 53 (1986): 1129–1174.
17. Ted O'Donoghue and Matthew Rabin, "Studying Optimal Paternalism, Illustrated by a Model of Sin Taxes," *American Economic Review Papers and Proceedings* 93 (2003): 186–191, at p. 186.

18. Colin Camerer, Samuel Issacharoff, George Loewenstein, Ted O'Donoghue, and Matthew Rabin, "Regulation for Conservatives: Behavioral Economics and the Case for 'Asymmetric Paternalism,'" *University of Pennsylvania Law Review* 151 (2003): 1211–1254, at pp. 1211–1212. Camerer is also well-known for his work in *neuroeconomics*, an even more recent approach to economics, which investigates the neurological foundations of economic behavior. For an introduction, see Colin Camerer, George Loewenstein, and Drazen Prelec, "Neuroeconomics: How Neuroscience Can Inform Economics," *Journal of Economic Literature* 43 (2005): 9–64.
19. Camerer et al., "Regulation for Conservatives," p. 1214.
20. Ibid., pp. 1212–1214.
21. Jolls, Sunstein, and Thaler, "Behavioral Approach," pp. 48–49.
22. O'Donoghue and Rabin, "Studying Optimal Paternalism," p. 186. For an introduction to the field of public choice, I'd recommend Dennis Mueller (ed.), *Perspectives on Public Choice: A Handbook* (Cambridge: Cambridge University Press, 1996).
23. Jolls, Sunstein, and Thaler, "Behavioral Approach," p. 49.
24. O'Donoghue and Rabin, "Studying Optimal Paternalism," p. 186. Camerer et al., too, write that a "crucial assumption in our approach is that the bounds on rationality—their range and implications, as well as which policies help—are empirical questions subject to systematic analysis, and thus cost-benefit judgments can be made" ("Regulation for Conservatives," p. 1222).
25. Richard H. Thaler and Cass R. Sunstein, *Nudge: Improving Decisions about Health, Wealth, and Happiness* (New Haven, CT: Yale University Press, 2008). Their main academic work on the topic is Sunstein and Thaler, "Libertarian Paternalism Is Not an Oxymoron," *University of Chicago Law Review* 70 (2001): 1159–1202. (A condensed version, presumably designed paternalistically to account the irrational impatience of your average economist, is Thaler and Sunstein, "Libertarian Paternalism," *American Economic Review Papers and Proceedings* 93 (2003): 175–179.)
26. Thaler and Sunstein, *Nudge*, p. 5.
27. Ibid. While this book doesn't focus on this point, many people have criticized their use of the word "libertarian." See, for instance, Daniel B. Klein, "Statist Quo Bias," *Econ Journal Watch* 1 (2004): 260–271; Gregory Mitchell, "Libertarian Paternalism Is an Oxymoron," *Northwestern University Law*

*Review* 99 (2005): 1245–1277; On Amir and Orly Lobel, "Stumble, Predict, Nudge: How Behavioral Economics Informs Law and Policy," *Columbia Law Review* 108 (2008): 2098–2139, at pp. 2117–2127; and Daniel M. Hausman and Brynn Welch, "Debate: To Nudge or Not to Nudge," *Journal of Political Philosophy* 18 (2010): 123–136, at pp. 124–126.

28. Thaler and Sunstein, *Nudge*, p. 6.
29. Camerer et al. give a similar list of examples in "Regulation for Conservatives," pp. 1224–1250.
30. Thaler and Sunstein, *Nudge*, p. 108.
31. In their book *Nudge* (pp. 1–3) Carolyn runs a school cafeteria system, which changes the implications regarding paternalism since children are often considered legitimate subjects for such interventions. (She also has a degree in nutrition, natch.) In their law review paper "Libertarian Paternalism Is Not an Oxymoron" (p. 1164), however, the unnamed female director runs a cafeteria in an "organization" that presumably caters to adults, which is more controversial—and more relevant to debates over paternalism.
32. Felicity Lawrence, "First goal of David Cameron's 'nudge unit' is to encourage healthy living," *The Guardian*, November 12, 2010, available at http://www.guardian.co.uk/politics/2010/nov/12/david-cameron-nudge-unit.
33. Benjamin Wallace-Wells, "Cass Sunstein Wants to Nudge Us," *The New York Times Magzine*, May 13, 2010, available at http://www.nytimes.com/2010/05/16/magazine/16Sunstein-t.html.
34. Thaler and Sunstein, *Nudge*, p. 9 (emphasis added).
35. I guess I qualify as what Camerer et al. call a "faith-based antipaternalist" ("Regulation for Conservatives," p. 1222), though my position has little to do with faith and a whole lot to do with respect.

### 4 Why Nudges Can't Do What They Promise

1. They never even invite Nancy to these parties, although the roommate with the tattoo winked at her once, but it was while she was working and it was a pretty heavy day for pollen, so he may have just been blinking, but she's pretty sure he was looking at her when he did it . . .
2. Do we really know our interests that well at all? Perhaps not as well as we think. Psychologists such as Dr. Timothy D. Wilson (*Strangers to Ourselves*, Cambridge, MA: Harvard University

Press, 2002) argue that most of what constitutes our personalities or selves is hidden in our unconscious, and we only imagine a deep self-awareness. He cites experiments that suggest our friends and family actually know us better than we know ourselves—OK, Mom, you were right—because they can observe what choices we make without the distorting effects of our self-image. Although I acknowledge this and am fascinated by it, it has little relevance for the argument I'm making—unless the behavioral economist designing nudges for you is your mother.

3. J. D. Trout, "Paternalism and Cognitive Bias," *Law and Philosophy* 24 (2005): 393–434, at p. 394. We'll see more about what "considered judgments" means later in this chapter.
4. Ted O'Donoghue and Matthew Rabin, "Studying Optimal Paternalism, Illustrated by a Model of Sin Taxes," *American Economic Review Papers and Proceedings* 93 (2003): 186–191, at p. 186.
5. Christine Jolls, Cass Sunstein, and Richard Thaler, "A Behavioral Approach to Law and Economics," reprinted in Cass Sunstein (ed.), *Behavioral Law & Economics* (Cambridge: Cambridge University Press, 2000), pp. 13–58, at p. 49.
6. Colin Camerer, Samuel Issacharoff, George Loewenstein, Ted O'Donoghue, and Matthew Rabin, "Regulation for Conservatives: Behavioral Economics and the Case for 'Asymmetric Paternalism,'" *University of Pennsylvania Law Review* 151 (2003): 1211–1254, at p. 1222.
7. Claire A. Hill, "Anti-Anti-Anti-Paternalism," *NYU Journal of Law & Liberty* 2 (2007): 444–454, at p. 448.
8. Cass Sunstein and Richard Thaler, "Libertarian Paternalism Is Not an Oxymoron," *University of Chicago Law Review* 70 (2001): 1159–1202, at p. 1193.
9. Camerer *et al*, "Regulation for Conservatives," pp. 1253–1254 (emphasis added).
10. This argument has strong parallels with economist and Nobel laureate Friedrich Hayek's arguments against the possibility of economy-wide central planning: planners do not and cannot have all of the information about preferences, resources, and technology necessary to determine how resources should be allocated in order to maximize well-being. (See his 1945 paper "The Use of Knowledge in Society," reprinted in his collection *Individualism and Economic Order*, Chicago: University of Chicago Press, 1948, pp. 77–91.) Economists Mario Rizzo

and Glen Whitman extend Hayek's point to libertarian paternalism in their paper "The Knowledge Problem of New Paternalism" (*Brigham Young University Law Review* 2009 (2009): 905–968, at p. 910):

> We argue that the new paternalism spawned by behavioral economics faces a very similar knowledge problem and for similar reasons. *If* well-meaning policymakers possess all the relevant information about individuals' true preferences, their cognitive biases, and the choice contexts in which they manifest themselves, *then* policymakers could potentially implement paternalist policies that improve the welfare of individuals by their own standards. But lacking such information, we cannot conclude that actual paternalism will make their decisions better; under a wide range of circumstances, it will even make them worse.

11. Dan W. Brock, "Paternalism and Autonomy," *Ethics* 98 (1988): 550–565, at p. 559.
12. Philosopher and legal scholar Seana Valentine Shiffrin's characterization of paternalism similarly focuses on value substitution; see her article "Paternalism, Unconscionability Doctrine, and Accommodation," *Philosophy and Public Affairs* 29 (2000): 205–250, at pp. 218–219.
13. Camerer et al., "Regulations for Conservatives," p. 1212.
14. Sunstein and Thaler, "Libertarian Paternalism Is Not an Oxymoron," pp. 1172–1173.
15. Ibid., p. 1162.
16. We'll talk more about the interactions between different paternalistic interventions in chapter 6.
17. Camerer et al., "Regulation for Conservatives," p. 1222. They go to say that they "envision a vaguely similar system in which substituting the phrase 'economic judgment' for the phrase 'health and diet' in the preceding analogy leads to a similar mix of information, persuasion, and regulation" (pp. 1222–1223).
18. Sunstein and Thaler, "Libertarian Paternalism Is Not an Oxymoron," p. 1164.
19. For an early influential paper in this area, see Avinash Dixit and Victor Norman, "Advertising and Welfare," *Bell Journal of Economics* 9 (1978): 1–17. On the issue of the effect of advertising on preferences, see Len M. Nichols, "Advertising and Economic Welfare," *American Economic Review* 75 (1985): 213–218.

20. For a unique approach to endogenous preferences in the context of things like music and drugs—not that those two have anything to do with each other—see Nobel laureate Gary Becker's *Accounting for Tastes* (Cambridge, MA: Harvard University Press, 1998). (We met Becker briefly in chapter 1 as he struggled with the decision to park illegally to get to a student's exam—and then founded the economics of crime based on the experience.)
21. To get the full effect, assertively thrust a finger in the air when you read the word "must."
22. Sunstein and Thaler, "Libertarian Paternalism Is Not an Oxymoron," p. 1162.
23. John Harsanyi, "Morality and the Theory of Rational Behavior," reprinted in Amartya Sen and Bernard Williams (eds.), *Utilitarianism and Beyond* (Cambridge: Cambridge University Press, 1982), pp. 39–62, at p. 55. (See also Sen and Williams' introduction to the volume, pp. 9–11, for more on the problems with "idealization" of preferences.)
24. Interestingly—to me, at least—this issue also arises in contract law when the judge decides that a contract as written is "unconscionable" and cannot be enforced, because its terms are, in the words of Seana Valentine Shiffrin, "seriously one-sided, overreaching, exploitative, or otherwise manifestly unfair" ("Paternalism, Unconscionability Doctrine, and Accommodation," p. 205). In such cases, the judge either voids the contract or modifies it to reflect what the parties "would have" agreed to under ideal conditions of equal bargaining power, perfect foresight, and so forth—in other words, the same counterfactual as in the construction of rational preferences. As implied by the title of Shiffrin's paper, this is widely regarded as a clear example of paternalism (though Shiffrin's view on the issue is more nuanced).
25. Mmm . . . fudge.
26. Robert Sugden, "Why Incoherent Preferences Do Not Justify Paternalism," *Constitutional Political Economy* 19 (2008): 226–248, at p. 232.
27. Imagine Winnie the Pooh saying that last bit. (Right?)
28. Think of Thomas Frank's book *What's Wrong with Kansas?* (New York: Metropolitan, 2004), which argued that Kansas voters were "duped" into voting for Republicans whereas their economic interests were supported by Democrats—assuming their true interests, the ones that should have influenced

their votes, were purely economic. Forget what's wrong with Kansas—what's wrong with Mr. Frank?

### 5 Why Nudges Are Unethical

1. Gerald Dworkin, "Paternalism," in *Stanford Encyclopedia of Philosophy* (http://plato.stanford.edu/entries/paternalism/), 2010.
2. Philosopher Immanuel Kant made this argument in his 1797 book *The Metaphysics of Morals*, ed. Mary Gregor (Cambridge: Cambridge University Press, 1996), p. 231. (This is standard pagination based on the Prussian Academy edition of his complete works. All citations of Kant's work herein will give this pagination, which is used in all reputable editions of his work.)
3. John Stuart Mill, *On Liberty*, ed. D. Spitz (New York: W. W. Norton, 1975), pp. 10–11.
4. Ibid., p. 78 (emphasis added).
5. Ibid., p. 89.
6. See Joel Feinberg, *Harm to Self* (Oxford: Oxford University Press, 1986), pp. 12–16.
7. As you can probably guess, the car was driven by Dave, with Ed in the passenger seat and their adopted children Frank, Greg, and Hillary in the back. They're a wonderful family from Indiana, and they feel just horrible about what happened to Carl.
8. Feinberg, *Harm to Self*, p. 115 (emphasis in original).
9. Ibid. (adapted from Diagram 20–5).
10. On second thought, you don't climb a hurdle, you jump it. Hmm.
11. Richard H. Thaler and Cass R. Sunstein, *Nudge: Improving Decisions about Health, Wealth, and Happiness* (New Haven, CT: Yale University Press, 2008), pp. 106–107.
12. Keep in mind that there may be other, nonpaternalistic motives for such coercive measures; for instance, bans on illegal drugs are supported by public safety concerns as well as concern for users, and cigarette taxes that decrease smoking also benefit nonsmokers who may be exposed to secondhand smoke.
13. Cass Sunstein and Richard Thaler, "Libertarian Paternalism Is Not an Oxymoron," *University of Chicago Law Review* 70 (2001): 1159–1202, at p. 1162.

14. Ibid., p. 1165.
15. Gerald Dworkin, *The Theory and Practice of Autonomy* (Cambridge: Cambridge University Press, 1988), p. 123.
16. Immanuel Kant, *Grounding for the Metaphysics of Morals*, trans. James W. Ellington (Indianapolis: Hackett, 1993), pp. 429–430.
17. This is almost identical to the wording of one of the formulae of Kant's famous *categorical imperative*: "Act in such a way that you treat humanity, whether in your own person or in the person of another, always at the same time as an end and never simply as a means" (*Grounding*, p. 429).
18. Daniel M. Hausman and Brynn Welch, "Debate: To Nudge or Not to Nudge," *Journal of Political Philosophy* 18 (2010): 123–136, at p. 130.
19. Dworkin, *Theory and Practice of Autonomy*, pp. 123–124.
20. Ibid., p. 88.
21. Onora O'Neill, "Between Consenting Adults," in *Constructions of Reason: Explorations of Kant's Practical Philosophy* (Cambridge: Cambridge University Press, 1989), pp. 105–125, at p. 110.
22. To take this even further, Hausman and Welch raise the specter of subliminal messaging as the natural extension of the libertarian paternalist philosophy and argue against Sunstein and Thaler's dismissal of the idea. (See "Debate: To Nudge or Not to Nudge," pp. 130–132, and also Luc Bovens, "The Ethics of Nudge," in Till Grüne-Yanoof and Sven Ove Hansson (eds.), *Preference Change: Approaches from Philosophy, Economics and Psychology*, (Dordrecht: Springer, 2009), pp. 207–219, at pp. 216–217.)
23. Sunstein and Thaler, "Libertarian Paternalism Is Not an Oxymoron," p. 1164.
24. Ibid. We shall assume that this cafeteria is embedded in some institution—a school, a hospital, the Behavioral Insight Team headquarters—and not an independent concern that must turn a profit to survive.
25. Ibid.
26. Ibid., p. 1173.
27. Ibid., p. 1193, n. 132.
28. Ibid., p. 1173.
29. As Rush—the band, not the radio personality—reminds us, if you choose not to decide, you still have made a choice.
30. A study like this one: James J. Choi, David Laibson, Brigitte C. Madrian, and Andrew Metrick, "Defined Contribution

Pensions: Plan Rules, Participant Choices, 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, at p. 86, cited in Sunstein and Thaler, "Libertarian Paternalism Is Not an Oxymoron," p. 1173.

## 6 All Nudges Are Not Created Equal

1. Martin Lindstrom, *Brandwashed: Tricks Companies Use to Manipulate Our Minds and Persuade Us to Buy* (New York: Crown, 2011). Other books in this area include Lindstrom's *Buyology: Truth and Lies About Why We Buy* (New York: Crown, 2008); Ellen Ruppel Shell, *Cheap: The High Cost of Discount Culture* (New York: Penguin, 2009); William Poundstone, *Priceless: The Myth of Fair Value (and How to Take Advantage of It)* (New York: Hill and Wang, 2010); and Eduardo Porter, *The Price of Everything: Solving the Mystery of Why We Pay What We Do* (New York: Portfolio, 2011).
2. Immanuel Kant, *Grounding for the Metaphysics of Morals*, trans. James W. Ellington (Indianapolis: Hackett, 1993), pp. 429–430 (emphasis added).
3. Richard H. Thaler and Cass R. Sunstein, *Nudge: Improving Decisions about Health, Wealth, and Happiness* (New Haven, CT: Yale University Press, 2008), chapters 11 ("How to Increase Organ Donations") and 12 ("Saving the Planet").
4. For an excellent summary of research and efforts in the United Kingdom to pursue government policy using the Sunstein-Thaler approach, see Peter John, Sarah Cotterill, Alice Moseley, Liz Richardson, Graham Smith, Gerry Stoker, and Carinne Wales, *Nudge, Nudge, Think, Think: Experimenting with Ways to Change Civic Behaviour* (London: Bloomsbury Academic, 2011). Philosophers Evan Selinger and Kyle Powys Whyte have commented critically on this use of nudges: on the effectiveness of nudges in policy contexts in general, see Selinger and Whyte, "Nudging Cannot Solve Complex Policy Problems," *European Journal of Risk Regulation* 1 (2012): 26–31, and on their applicability in promoting organ donation, see Kyle Powys Whyte, Evan Selinger, Arthur L. Caplan, and Jathan Sadowski, "Nudge, Nudge or Shove, Shove—The Right Way for Nudges to Increase the Supply of Donated Cadaver Organs," *American Journal of Bioethics* 12 (2012): 32–39.

5. See note 23 in this chapter regarding a similar process of give-and-take between businesses and libertarian paternalist regulators.
6. See the *circular flow diagram* in any introductory economics textbook while you hum the tune "The Wheels on the Bus Go Round and Round."
7. Edward Glaeser, "Paternalism and Psychology," *University of Chicago Law Review* 73 (2006): 133–156, at p. 135. Note that he uses the word "persuasion" in a less-than-ideal sense, not what we usually think of reasoned and respectful debate.
8. These criticisms are hardly new, appearing the some form in the work of economists Thorstein Veblen and John Kenneth Galbraith, to name just two. Modern scholars who have emphasized the role market forces play in determining consumer preferences include economist David George (*Preference Pollution: How Markets Create the Desires We Dislike*, Ann Arbor: University of Michigan Press, 2001, esp. chap. 3), legal scholars Jon D. Hanson and Douglas A. Kysar ("Taking Behavioralism Seriously: The Problem of Market Manipulation," *New York University Law Review* 74 (1999): 101–217), and physician Peter A. Ubel (*Free Market Madness: Why Human Nature Is at Odds with Economics—and Why It Matters*, Boston, MA: Harvard Business Press, 2009). The theme also appears in much of Cass Sunstein's pre-*Nudge* work; see, for instance, his book *Free Markets and Social Justice* (Oxford: Oxford University Press, 1997).
9. Or "wantons" in philosopher Harry Frankfurt's sense; see note 5 in chapter 3.
10. I say "nearly" because, technically, the employee can leave the company out of protest, an option often unavailable when facing government paternalism.
11. Immanuel Kant, *The Metaphysics of Morals*, ed. Mary Gregor (Cambridge: Cambridge University Press, 1996), p. 451 (emphasis mine).
12. For Smith's view, see *The Theory of Moral Sentiments* (New York: Penguin, 2010), part VI, section II, chap. 1; for more on Kant's version, see *Metaphysics of Morals*, pp. 450–452.
13. This is true even if, in some specific aspects, we know them better than they know themselves; see note 2 in chapter 4.
14. There oughta be a song…
15. Kant, *Metaphysics of Morals*, p. 454.

16. To be sure, not everyone agrees with this view of care. There is a school of moral philosophy called *care ethics* that maintains that care should govern all interactions between individuals, business, and government, and that concepts of justice and respect should be understood in terms of care. For a good introduction by a central figure in the field, see Virginia Held's *The Ethics of Care: Personal, Political, and Global* (Oxford: Oxford University Press, 2007).
17. Dan Brock, "Paternalism and Autonomy," *Ethics* 98 (1988): 550–565, at p. 551.
18. Thaler and Sunstein, *Nudge*, p. 90; more generally, see pp. 90–91, 188–194.
19. Jonathan Klick and Gregory Mitchell, "Government Regulation of Irrationality: Moral and Cognitive Hazards," *Minnesota Law Review* 90 (2006): 1620–1663, at p. 1633. The research they draw from is James P. Byrnes, *The Nature and Development of Decision Making: A Self-Regulation Model* (Mahwah, NJ: Lawrence Erlbaum Associates, 1998).
20. Klick and Mitchell, "Government Regulation of Irrationality," p. 1635, quoting Jennifer S. Lerner and Philip E. Tetlock, "Accounting for the Effects of Accountability," *Psychological Bulletin* 125 (1999): 255–275, at p. 265.
21. Luc Bovens, "The Ethics of *Nudge*," in Till Grüne-Yanoff and Sven Ove Hansson (eds.), *Preference Change: Approaches from Philosophy, Economics and Psychology* (Dordrecht: Springer, 2009), pp. 207–219, at p. 218.
22. Klick and Mitchell, "Government Regulation of Irrationality," pp. 1626–1627.
23. Ibid., p. 1639. They suggest a third source of self-fulfilling prophecy: as businesses adapt around nudged consumer behavior, new nudges must be designed to counter the new business practices, in a never-ending cycle of call and response between business and the regulators (pp. 1640–41).
24. Cass Sunstein and Richard Thaler, "Libertarian Paternalism Is Not an Oxymoron," *University of Chicago Law Review* 70 (2001): 1159–1202, at p. 1200.
25. On the slippery slope argument, see Mario J. Rizzo and Douglas Glen Whitman, "Little Brother Is Watching You: New Paternalism on the Slippery Slopes," *Arizona Law Review* 51 (2009): 685–739.
26. Sunstein and Thaler, "Libertarian Paternalism," pp. 1199–1200.
27. Psychologist Jonathan Baron writes that "biases may work together to restore a kind of artificial equilibrium that works

for normal situations, so that correcting one bias without correcting another one can make things worse" ("Cognitive Biases, Cognitive Limits, and Risk Communication," *Journal of Public Policy and Marketing* 23 (2004), 7–13, at p. 11). Even Ted O'Donoghue and Matthew Rabin, early proponents of some version of libertarian paternalism, realize the danger of nudges casting too wide a net: "A policy that helps those agents making one common error may hurt those making another common error, even if the policy is virtually harmless to fully rational people" ("Studying Optimal Paternalism, Illustrated by a Model of Sin Taxes," *American Economic Review Paper and Proceedings* 93 (2003): 186–191, at p. 191).

28. Richard Posner, "Rational Choice, Behavioral Economics, and the Law," *Stanford Law Review* 50 (1998): 1551–1575, at p. 1575.
29. Jeffrey J. Rachlinski, "The Uncertain Psychological Case for Paternalism," 97 *Northwestern University Law Review* 97 (2003): 1165–1225, at p. 1214.
30. "Of course," as Cinderella's stepmother said, "I said *if*."

### 7 Why Choice Matters So Much—and What Can Be Done to Preserve It

1. For more on Kant's conception of autonomy, see Roger J. Sullivan, *Immanuel Kant's Moral Theory* (Cambridge: Cambridge University Press, 1989), chap. 5 and Thomas E. Hill, Jr., *Autonomy and Self-Respect* (Cambridge: Cambridge University Press, 1991), chaps. 3 and 4.
2. Mark D. White, *Kantian Ethics and Economics: Autonomy, Dignity, and Character* (Stanford, CA: Stanford University Press, 2011), especially chap. 1.
3. For a fascinating look into "what we care about," see Harry Frankfurt's title essay in his collection *The Importance of What We Are About* (Cambridge: Cambridge University Press, 1988), pp. 80–94.
4. John Stuart Mill, *On Liberty*, ed. D. Spitz (New York: W. W. Norton, 1975), p. 54.
5. Actually, if I ever acted "normally," that would more likely be seen as out of character!
6. Christine M. Korsgaard, *Creating the Kingdom of Ends* (Cambridge: Cambridge University Press, 1996), p. 378.
7. Hill, *Autonomy and Self-Respect*, p. 99. The chapter containing this quote discusses suicide, and Hill suggests his "Author Perspective" as one way to understand (if not justify) the

decision to kill oneself, by ensuring that one can end the story as one wants (with dignity, for instance). "If you value being an author and have just one story to write, you should not hurry to conclude it. But sometimes, to give it the meaning you intend, you must end it before you spoil it" (p. 101).

8. For a concise introduction to Sartre, see Thomas Flynn's entry on him at the *Stanford Encyclopedia of Philosophy*: http://plato.stanford.edu/entries/sartre/.
9. Christine M. Korsgaard, *Self-Constitution: Agency, Identity, and Integrity* (Oxford: Oxford University Press, 2009), p. 25.
10. As Korsgaard begins her book *Self-Constitution*, "human beings are *condemned* to choice and action" (p. 1).
11. This discussion of autonomy was all too brief. The interested reader would be suggested to look to books such as the following: Gerald Dworkin, *The Theory and Practice of Autonomy* (Cambridge: Cambridge University Press, 1988); Ellen Frankel Paul, Fred D. Miller, Jr., and Jeffrey Paul (eds.), *Autonomy* (Cambridge: Cambridge University Press, 2003); and James Stacey Taylor (ed.), *Personal Autonomy* (Cambridge: Cambridge University Press, 2005). If that's not enough, see the references at the end of John Christman's "Autonomy in Moral and Legal Philosophy" at the *Stanford Encyclopedia of Philosophy*, at http://plato.stanford.edu/entries/autonomy-moral/.
12. Luc Bovens, "The Ethics of *Nudge*," in Till Grüne-Yanoff and Sven Ove Hansson (eds.), *Preference Change: Approaches from Philosophy, Economics and Psychology* (Dordrecht: Springer, 2009), pp. 207–219, at p. 217.
13. Cass Sunstein and Richard Thaler, "Libertarian Paternalism Is Not an Oxymoron," *University of Chicago Law Review* 70 (2001): 1159–1202, at p. 1162.
14. Daniel M. Hausman and Brynn Welch, "Debate: To Nudge or Not to Nudge," *Journal of Political Philosophy* 18 (2010): 123–136, at p. 127. Hausman and Welch also assert that such informational campaigns are not paternalistic, which is questionable: although they're certainly not manipulative or coercive, they are motivated by concern for people's own well-being. A libertarian may object to government funds being used for this purpose, arguing that in many cases (such as nutritional labeling) the market will provide the information that consumers demand. Nonetheless, even though these measures can be regarded as paternalistic, it is certainly one of the more innocuous forms of paternalism possible.

15. The US labeling may have changed by the time you read this book; the Food and Drug Administration has mandated larger and more graphic warnings, but while I weave this wondrous web of wordage, the law is being argued in court. (Let's hope they never outlaw alliteration.)
16. Ezekiel J. Emanuel, "Healthy Labels, Not Stealthy Labels," *The New York Times*, March 5, 2012, available at http://opinionator.blogs.nytimes.com/2012/03/05/healthy-labels-not-stealthy-labels/. (I will restrain myself from commenting on the delicious irony in the title.)
17. For more on this system, part of the National Health Service's "Live Well" initiative, see their website at http://www.nhs.uk/Livewell/Goodfood/Pages/food-labelling.aspx. (Again, I will not comment on the irony of the phrase "Your health, your choices" that is splashed across of the top of their website. Nope.)
18. Not to worry, however; according to *New York Times* food writer Mark Bittman, in discussing taxes on unhealthy foods, "we have experts who can figure out how 'bad' a food should be to qualify" ("Bad Food? Tax It, and Subsidize Vegetables," *The New York Times*, July 23, 2011, available at http://www.nytimes.com/2011/07/24/opinion/sunday/24bittman .html).
19. For a technical discussion of combatting cognitive biases and heuristics directly, see Richard P. Larrick, "Debiasing," in Derek J. Koehler and Nigel Harvey (eds.), *Blackwell Handbook of Judgment and Decision Making* (Malden, MA: Blackwell, 2004), pp. 316–337.
20. Colin Camerer, Samuel Issacharoff, George Loewenstein, Ted O'Donoghue, and Matthew Rabin, "Regulation for Conservatives: Behavioral Economics and the Case for 'Asymmetric Paternalism,'" *University of Pennsylvania Law Review* 151 (2003): 1211–1254, at pp. 1238–1247.
21. Interestingly, Sunstein and Thaler are careful to identify cooling-off periods with Camerer et al.'s asymmetric paternalism rather than nudges, though they do agree with the logic behind them; see *Nudge: Improving Decisions about Health, Wealth, and Happiness* (New Haven, CT: Yale University Press, 2008), pp. 250–251.
22. Gregory Mitchell, "Libertarian Paternalism Is an Oxymoron," *Northwestern University Law Review* 99 (2005): 1245–1277, at pp. 1255–1256.
23. Oren Bar-Gill and Elizabeth Warren, "Making Credit Safer," *University of Pennsylvania Law Review* 157 (2008): 1–101, at

p. 26–27. Warren was involved in the creation and administration of the Consumer Financial Protection Bureau, created to implement many of the policies she recommended in her work, and later became a US Senator from Massachusetts.

24. Ibid., pp. 27–32.
25. Ibid., p. 50.
26. Michael S. Barr, Sendhil Mullainathan, and Eldar Shafir, "Behaviorally Informed Financial Services Regulation." New American Foundation (2008), available at http://www.newamerica.net/files/naf_behavioral_v5.pdf, p. 9.
27. Michael M. Grynbaum, "New York Plans to Ban Sale of Big Sizes of Sugary Drinks," *The New York Times*, May 30, 2012, available at http://www.nytimes.com/2012/05/31/nyregion/bloomberg-plans-a-ban-on-large-sugared-drinks.html.

# Further Reading

Since I chose to include complete references in the endnotes to each chapter, I thought it would be helpful to include a consolidated and expanded list of books and articles on behavioral law and economics, libertarian paternalism, and nudges. Many of the works listed here were not cited in the book, and those that were cited are worth reading for more than what little I drew from them. Finally, this list is hardly complete or up-to-date—especially by the time you read this! While I'm sure I missed some important references, the ones below have been very useful in contributing to my thoughts on the issue.

### Classic and Contemporary Work on Paternalism (in General)


Arneson, Richard J. "Mill versus Paternalism." *Ethics* 90 (1980): 470–489.

Brock, Dan W. "Paternalism and Autonomy." *Ethics* 98 (1988): 550–565.

Coons, Christian, and Michael Weber, eds. *Paternalism: Theory and Practice*. Oxford: Oxford University Press, 2013.

Dworkin, Gerald. "Paternalism." In *Stanford Encyclopedia of Philosophy*, 2010, available at http://plato.stanford.edu/entries/paternalism/.

———. "Paternalism." *The Monist* 56 (1972): 64–84.

———. *The Theory and Practice of Autonomy*. Cambridge: Cambridge University Press, 1988, chap. 8.

Dworkin, Ronald. *Sovereign Virtue: The Theory and Practice of Equality*. Cambridge, MA: Harvard University Press, 2000, chap. 5.

Feinberg, Joel. *Harm to Self*. Oxford: Oxford University Press, 1986.

———. "Legal Paternalism." Reprinted in *Rights, Justice, and the Bounds of Liberty*. Princeton, NJ: Princeton University Press, 1980, pp. 110–129.

Kant, Immanuel. *The Metaphysics of Morals.* Edited by Mary Gregor. Cambridge: Cambridge University Press, 1797/1996.

Kleinig, John. *Paternalism.* Manchester: Manchester University Press, 1983.

Mill, John Stuart. *On Liberty.* Edited by D. Spitz. New York: W. W. Norton, 1859/1975.

Sartorius, Rolf, ed. *Paternalism.* Minneapolis: University of Minnesota Press, 1983.

Scoccia, Danny. "In Defense of Hard Paternalism." *Law and Philosophy* 27 (2008): 351–381.

———. "Paternalism and Respect for Autonomy." *Ethics* 100 (1990): 318–334.

Shiffrin, Seana Valentine. "Paternalism, Unconscionability Doctrine, and Accommodation." *Philosophy and Public Affairs* 29 (2000): 205–250.

Zamir, Eyal. "The Efficiency of Paternalism." *Virginia Law Review* 84 (1998): 229–286.

### Original Works on Behavioral Law and Economics and Libertarian Paternalism

Bar-Gill, Oren, and Elizabeth Warren. "Making Credit Safer." *University of Pennsylvania Law Review* 157 (2008): 1–101.

Barr, Michael S., Sendhil Mullainathan, and Eldar Shafir. "Behaviorally Informed Financial Services Regulation." New American Foundation, 2008, available at http://www.newamerica.net/files/naf_behavioral_v5.pdf.

Camerer, Colin, Samuel Issacharoff, George Loewenstein, Ted O'Donoghue, and Matthew Rabin. "Regulation for Conservatives: Behavioral Economics and the Case for 'Asymmetric Paternalism.'" *University of Pennsylvania Law Review* 151 (2003): 1211–1254.

Jolls, Christine, Cass Sunstein, and Richard Thaler. "A Behavioral Approach to Law and Economics." *Stanford Law Review* 50 (1998): 1471–1550. Also reprinted in Sunstein, *Behavioral Law & Economics*, 13–58.

———. "Theories and Tropes: A Reply to Posner and Kelman." *Stanford Law Review* 50 (1998): 1593–1608.

Korobkin, Russell B., and Thomas S. Ulen. "Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics." *California Law Review* 88 (2000): 1051–1144.

O'Donoghue, Ted, and Matthew Rabin. "Studying Optimal Paternalism, Illustrated by a Model of Sin Taxes." *American Economic Review Papers and Proceedings* 93 (2003): 186–191.

Parisi, Francesco, and Vernon L. Smith, eds. *The Law and Economics of Irrational Behavior.* Stanford, CA: Stanford University Press, 2005.

Sunstein, Cass, ed. *Behavioral Law & Economics.* Cambridge: Cambridge University Press, 2000.

Sunstein, Cass, and Richard Thaler. "Libertarian Paternalism Is Not an Oxymoron," *University of Chicago Law Review* 70 (2001): 1159–1202.

———. "Preferences, Paternalism, and Liberty." In *Preferences and Well-Being*, edited by Serena Olsaretti. Cambridge: Cambridge University Press, 2006, pp. 233–264.

Thaler, Richard, and Cass Sunstein. "Libertarian Paternalism." *American Economic Review Papers and Proceedings* 93 (2003): 175–179.

———. *Nudge: Improving Decisions about Health, Wealth, and Happiness.* New Haven, CT: Yale University Press, 2008.

#### Later Commentary and Criticism of Behavioral Law and Economics and Libertarian Paternalism

Amir, On, and Orly Lobel. "Stumble, Predict, Nudge: How Behavioral Economics Informs Law and Policy." *Columbia Law Review* 108 (2008): 2098–2139.

Blumenthal-Barby, J. S., and Hadley Burroughs. "Seeking Better Health Care Outcomes: The Ethics of Using the 'Nudge.'" *American Journal of Bioethics* 12, no. 2 (2012): 1–10.

Bovens, Luc. "The Ethics of Nudge." In *Preference Change: Approaches from Philosophy, Economics and Psychology*, edited by Till Grüne-Yanoof and Sven Ove Hansson. Dordrecht: Springer, 2009, pp. 207–219.

Burgess, Adam. "'Nudging' Healthy Lifestyles: The UK Experiments with the Behavioural Alternative to Regulation and the Market." *European Journal of Risk Regulation* 1 (2012): 3–16.

Epstein, Richard A. "Behavioral Economics: Human Errors and Market Corrections." *University of Chicago Law Review* 73 (2006): 111–132.

"Free to Choose? A Symposium on Behavioral Law and Economics," at the Truth on the Market blog, December 6–7, 2010, available at http://truthonthemarket.com/free-to-choose-symposium/. (Includes contributions from David Friedman, Larry Ribstein, David Levine, Henry G. Manne, Geoffrey A. Manne, Thom Lambert, Christopher Sprigman and Christopher Buccafusco, Judd E. Stone, Ronald Manne, Richard A. Epstein, Claire Hill, Kevin

McCabe, Tom Brown, Stephen Bainbridge, Erin O'Hara, Todd Henderson, and Douglas A. Ginsburg and Joshua Wright.)

Glaeser, Edward. "Paternalism and Psychology." *University of Chicago Law Review* 73 (2006): 133–156.

Hausman, Daniel M., and Brynn Welch. "Debate: To Nudge or Not to Nudge." *Journal of Political Philosophy* 18 (2010): 123–136.

Hayden, Grant M., and Stephen E. Ellis. "Law and Economics after Behavioral Economics." *Kansas Law Review* 55 (2007): 629–675.

Hill, Claire A. "Anti-Anti-Anti-Paternalism." *NYU Journal of Law & Liberty* 2 (2007): 444–454.

Huigens, Kyron. "Law, Economics, and the Skeleton of Value Fallacy." *California Law Review* 89 (2001): 537–568.

John, Peter, Sarah Cotterill, Alice Moseley, Liz Richardson, Graham Smith, Gerry Stoker, and Carinne Wales. *Nudge, Nudge, Think, Think: Experimenting with Ways to Change Civic Behaviour.* London: Bloomsbury Academic, 2011.

Kelman, Mark. "Behavioral Economics as Part of a Rhetorical Duet: A Response to Jolls, Sunstein, and Thaler." *Stanford Law Review* 50 (1998): 1577–1591.

Klein, Daniel B. "Statist Quo Bias." *Econ Journal Watch* 1 (2004): 260–271.

Klick, Jonathan, and Gregory Mitchell. "Government Regulation of Irrationality: Moral and Cognitive Hazards." *Minnesota Law Review* 90 (2006): 1620–1663.

Korobkin, Russell. "What Comes after Victory for Behavioral Law and Economics?" *University of Illinois Law Review* 2011 (2011): 1653–1674.

Loewenstein, George, and Emily Haisley. "The Economist as Therapist: Methodological Ramifications of 'Light' Paternalism." In *The Foundations of Positive and Normative Economics: A Handbook*, edited by Andrew Caplin and Andrew Schotter. Oxford: Oxford University Press, 2008, pp. 210–245.

McKenzie, Richard B. *Predictably Rational? In Search of Defenses for Rational Behavior in Economics.* Dordrecht: Springer, 2010, especially chap. 10.

Mitchell, Gregory. "Libertarian Paternalism Is an Oxymoron." *Northwestern University Law Review* 99 (2005): 1245–1277.

———. "Taking Behavioralism Too Seriously? Unwarranted Pessimism of the New Behavioral Analysis of Law." *William and Mary Law Review* 43 (2002): 1907–2021.

———. "Why Law and Economics' Perfect Rationality Should Not Be Traded for Behavioral Law and Economics' Equal Incompetence." *Georgetown Law Journal* 91 (2002): 67–167.

Posner, Richard A. "Rational Choice, Behavioral Economics, and the Law." *Stanford Law Review* 50 (1998): 1551–1575.

Rachlinski, Jeffrey J. "Cognitive Errors, Individual Differences, and Paternalism." *University of Chicago Law Review* 73 (2006): 207–229.

———. "The Uncertain Psychological Case for Paternalism." *Northwestern University Law Review* 97 (2003): 1165–1225.

Rebonato, Riccardo. *Taking Liberties: A Critical Examination of Libertarian Paternalism*. New York: Palgrave Macmillan, 2012.

Rizzo, Mario J., and Douglas Glen Whitman. "The Knowledge Problem of New Paternalism." *Brigham Young University Law Review* 2009 (2009): 905–968.

———. "Little Brother Is Watching You: New Paternalism on the Slippery Slopes." *Arizona Law Review* 51 (2009): 685–739.

———. "Paternalist Slopes." *NYU Journal of Law & Liberty* 2 (2007): 411–443.

Saint-Paul, Gilles. *The Tyranny of Utility: Behavioral Social Science and the Rise of Paternalism*. Princeton, NJ: Princeton University Press, 2011, chap. 7.

Selinger, Evan, and Kyle Powys Whyte. "Nudging Cannot Solve Complex Policy Problems." *European Journal of Risk Regulation* 1 (2012): 26–31.

Sugden, Robert. "Why Incoherent Preferences Do Not Justify Paternalism." *Constitutional Political Economy* 19 (2008): 226–248.

"Symposium: *Homo Economicus*, *Homo Myopicus*, and the Law and Economics of Consumer Choice." *University of Chicago Law Review* 73(1) (2006). (Includes papers by Ian Ayres, Douglas G. Baird, Oren Bar-Gill, Tom Brown and Lacey Plache, Colin F. Camerer, Richard A. Epstein, Edward L. Glaeser, Samuel Issacharoff and Erin F. Delaney, George Loewenstein and Ted O'Donoghue, Jeffrey J. Rachlinski, Luis Rayo and Gary S. Becker, and Cass R. Sunstein.)

Trout, J. D. "Paternalism and Cognitive Bias." *Law and Philosophy* 24 (2005): 393–434.

———. "A Restriction Maybe, But Is It Paternalism? Cognitive Bias and Choosing Governmental Decision Aids." *NYU Journal of Law & Liberty* 2 (2007): 455–469.

Whitman, Douglas Glen. "Against the New Paternalism: Internalities and the Economics of Self-Control." *Policy Analysis* no. 563 (February 22, 2006): 1–16.

Whyte, Kyle Powys, Evan Selinger, Arthur L. Caplan, and Jathan Sadowski. "Nudge, Nudge or Shove, Shove—The Right Way for Nudges to Increase the Supply of Donated Cadaver Organs." *American Journal of Bioethics* 12 (2012): 32–39.

# Index