

**The Economist as Therapist:  
Methodological Ramifications of “Light” Paternalism**

---

George Loewenstein and  
Emily Haisley

To appear in A. Caplin and A. Schotter (Eds.), "Perspectives on the Future of  
Economics: Positive and Normative Foundations", volume 1 in the Handbook  
of Economic Methodologies, Oxford, England: Oxford University Press.

We thank Dan Akst, Andrew Caplin, Peter Huang, Sam Issacharoff, Ed Glaeser, Ted  
O'Donoghue, Andy Schotter, Nachum Sicherman, Cass Sunstein and Kevin Volpp for helpful  
comments and suggestions.

## Abstract

---

We review methodological issues that arise in designing, implementing and evaluating the efficacy of 'light' paternalistic policies. In contrast to traditional 'heavy-handed' approaches to paternalism, light paternalistic policies aim to enhance individual choice without restricting it. Although light paternalism is a 'growth industry' in economics, a number of methodological issues that it raises have not been adequately addressed. The first issue is how a particular pattern of behavior should be judged as a mistake, and, relatedly, how the success of paternalistic policies designed to rectify such mistakes should be evaluated – i.e., the welfare criterion that should be used to judge light paternalistic policies. Second, paternalism, and especially light paternalism, introduces new motives for attempting to understand the psychological *processes* underlying economic behavior. An enhanced understanding of process can help to explain why people make mistakes in the first place, and, more importantly, provide insights into what types of policies are likely to be effective in correcting the mistakes. Third, there is an acute need for testing different possible policies before implementing them on a large scale, which we argue is best done in the field rather than the lab. Fourth, in addition to methodological issues, there are pragmatic issues concerning who will implement light paternalistic policies, especially when they involve positive expenditures. We discuss how economic interests can be rechanneled to support endeavors consistent with light paternalism.

## Introduction

---

Much economic behavior is, or at least appears to be, rational and self-interested. People balance price and quality when they decide where to shop and what to buy. They decide how much schooling to get and what to study based at least in part on likely returns to different forms of training and in part on their enjoyment of different topics and types of work. They carefully consider investment decisions and hire experts to get good advice. Even if some may view voting itself as irrational, economic interests seem to play at least some role in patterns of voting.

There are areas of life, however, in which people seem to display less than perfect rationality. For example, although the United States is one of the most prosperous nations in the world, with a large fraction of its population closing in on retirement, the net savings rate is close to zero and the average household has \$8,400 worth of credit card debt.<sup>1</sup> Fifty percent of U.S. households do not own any equities,<sup>2</sup> but the average man, woman, and child in the U.S. lost \$284 gambling in 2004— close to \$85 billion in total.<sup>3</sup> Many workers don't "max out" on 401k plans despite company matches (effectively leaving free money "on the table"), and what they do invest often goes undiversified into their own company's stocks or into fixed income investments with low long-term yields. At lower levels of income, many individuals and families sacrifice 10–15% of their paycheck each month to payday loans, acquire goods through rent-to-own establishments that charge effective interests rates in the hundreds of percent, or spend large sums on lottery tickets that return approximately fifty cents on the dollar. Worldwide, obesity rates and associated diseases are high and rising rapidly. Yet people with, or at risk for, life-threatening health conditions often fail to take the most rudimentary steps to protect themselves. One recent estimate is that modifiable behaviors such as tobacco use, overeating, and alcohol abuse account for nearly one-third of

all deaths in the United States [Flegal, Graubard, Williamson, and Gail, 2005; Schroeder, 2007]. Moreover, realization of the potential benefit of proven mediations, some targeted at the same medical problems caused by adverse health behaviors, is stymied by poor adherence rates among patients. For example, by one year after having a heart attack, nearly half of the patients prescribed cholesterol-lowering medications have stopped taking them [Jackevicius, Mamdani, and Tu, 2002].

As economists, how should we respond to the seemingly self-destructive side of human behavior? We can deny it and assume as an axiom of faith that people can be relied upon to do what's best for themselves. We can assume that families paying an average of \$1,000 per year financing credit card debt are making a rational trade-off of present and future utility, that liquidity constraints prevent investing in employer-matched 401k plans, that employees have good reasons for investing in their own company's stock instead of a diversified portfolio, that individuals' coefficients of relative risk aversion are high enough to justify investing in bonds instead of equities, that low-income families have good reasons for spending a large fraction of their paycheck on payday loans, usurious interest rates at rent-to-own establishments, and state lotteries, and that people are obese because they have calculated that the pleasure from the extra food, or the pain from the forgone exercise, is sufficient to compensate for the negative consequences of obesity. Indeed, some economists argue exactly that.<sup>4</sup>

Even among economists, however, this may no longer represent a majority view. Stimulated in part by developments in behavioral economics, increasing numbers of economists are questioning whether people really are such reliable pursuers of self-interest, and are coming to recognize that in some predictable situations people are prone to systematic errors.

In some cases, these errors arise from a lack of information, insight, or limited

computational ability. For example, people may not recognize that company matches on pension funds effectively represent “free money”; they may not understand why it doesn’t make sense to put one’s nest egg in one’s employer’s stocks, and they may not realize that stocks, on average, yield a higher return than bonds. In other cases, people are well aware of the best course of action but due to self-control problems or limited self-insight, are unable to implement it [e.g., Loewenstein, 1996]. Obesity and cigarette smoking may best fit into this latter category; few people have any illusions about the health risks of smoking or obesity, and many smokers and obese individuals do not believe that the benefits exceed the costs (which is why they often spend large amounts of time and money on attempts to quit). But in many cases this knowledge is insufficient to motivate behavior change.

### **“Light” Paternalism**

Part of the historic antagonism of economists toward behavioral economics may have been driven by a fear that documenting flaws in human decision making would inevitably lead to calls for paternalism. If so, it seems that such fears were well founded. Beyond documenting such apparent violations of rationality and their consequences for economic behavior, behavioral economists have indeed begun to take the next logical step: they have begun to devise “paternalistic” policies designed to steer economic behavior in more self-interested directions. Paternalistic policies have the goal of benefiting people on an individual basis, premised on the idea that people cannot be relied upon to invariably pursue self-interest. Whereas the conventional justification for government regulation is to limit *externalities*—costs people impose on other people that they don’t internalize—to promote the public good, the justification for paternalism is to limit *internalities*—costs that people impose on themselves that they don’t internalize [Hernstein, Loewenstein, Prelec, and Vaughan, 1993]. Although some of the behaviors that are targeted by paternalistic policies do

generate externalities (e.g., the failure to wear a motorcycle helmet imposes psychic and monetary costs on people other than the rider), paternalistic policies are generally aimed at helping the person whose behavior is altered. Existing examples of paternalistic regulations include banning narcotics, protection of the economically desperate with usury laws, health and safety regulations (for dangerous occupations), warnings on cigarettes, public health advertising, FDA drug approval, and the social security system.

In contrast to these existing forms of “heavy-handed” paternalism, however, behavioral economists have been advocating a new form of what could be called “light” paternalism. Going by labels such as “libertarian paternalism” [Thaler and Sunstein, 2003] and “asymmetric paternalism” [Camerer, Issacharoff, Loewenstein, O’Donoghue, and Rabin, 2003; Loewenstein, Brennan, and Volpp, 2007], the common goal of these approaches is to steer human behavior in more beneficial directions while minimizing coercion, maintaining individual autonomy, and maximizing choice to the greatest extent possible. Light paternalism aims to enhance decision making without restricting it.

In their treatment of “libertarian paternalism,” for example, Thaler and Sunstein [2003] note that paternalism is often simply not avoidable. In many situations, they point out, organizations or governments must make decisions that will necessarily affect the choices and welfare outcomes of their constituents. It would seem ridiculous not to consider how such decisions will impact the welfare of those affected. They illustrate the point with the hypothetical case of a company cafeteria manager who must either place healthy items before unhealthy items in a cafeteria line or the reverse, but does not have the option of doing neither. Thaler and Sunstein [2003] argue that in such situations it makes perfect sense for managers to adopt the option that they believe is better for employees—namely, placing the healthy food ahead of the unhealthy food. Another example that has received considerable attention is default options for 401(k) retirement plans. If it is beneficial to invest in a 401k

plan, but people tend to stick with the status quo, then it may make sense to change the usual default from not contributing (with the possibility of signing up) to contributing (with the possibility of opting out). The organization must make a choice about whether the default option is enrolled or unenrolled and, if enrolled, at what contribution level. Even if the organization were to have no default option and force employees to select whether they want to be in or out, this still qualifies as a decision of the organization that would lead to a different rate of enrollment and thus affects the welfare of its employees (see Choi, Laibson, Madrian, and Metrick [2005] and see Halpern, Ubel, and Ash [2007] for a discussion in the context of healthcare).

The central insight of Camerer et al.'s [2003] notion of "asymmetric paternalism" is that it is often possible to produce benefits for people who make suboptimal decisions while imposing minimal or no restrictions on those who make rational decisions that optimize their own welfare. In the most pure cases of asymmetric paternalism, people behaving suboptimally are benefited without imposing any costs on those behaving optimally. To continue with the example of defaults on 401k plans, if people, contrary to the dictates of conventional economics, are influenced by the default option, then changing the default could potentially benefit them; if people are not influenced by the default, then changing it will have no effect on behavior and little if any cost.<sup>5</sup> Such policies not only provide benefits to agents who make mistakes without hurting those who are making a deliberate decision, but should also appeal to economists both who do and who do not believe in rationality. Economists who believe that people are less than perfectly rational will perceive such policies as beneficial, while economists who believe in rationality should see them as, at worst, little more than a low-cost nuisance. Policies of this type use relatively subtle psychological factors to influence behavior, making it possible to accomplish policy goals without imposing more draconian mandatory measures such as raising the contribution rate

of social security. Exactly such an approach was adopted in the Pension Protection Act of 2006, which encourages companies to automatically enroll employees into 401(k) plans, and which passed with bipartisan support in an otherwise highly contentious political year. Other examples of policy interventions that fit the criteria for pure asymmetric paternalism include decision framing and expanding choice to offer commitment devices that aid self-control problems (as discussed below).

### **Critiques of Light Paternalism**

Despite the desire to enlist the support of economists who oppose more heavy-handed forms of paternalism, light paternalism is not without its critics. For example, Glaeser [2006] argues that the bureaucrats who guide paternalistic policies cannot be counted on to be any more rational than those affected by the policies and can be counted on to be less interested in the welfare of those affected than in their own welfare. There is certainly some validity to the point, yet there are predictable situations in which the more detached perspectives of policy makers or experts can be more rational than those of individual decision makers. For example, the individual may be faced with tempting choices that are hard to resist but at odds with his or her long-term interests. Policy makers can predict that people will yield to these temptations and may be able to steer such individuals toward making better choices. Similarly, policy makers may have the information processing resources to figure out the best course of action when it comes to complex decisions, such as when it makes sense to receive different types of health care procedures, in situations in which individuals often make mistakes due to the difficulty of interpreting information.

In a different vein, Sugden [2005] and Klick and Mitchell [2006] argue that there is an inherent value to autonomy—to letting people make mistakes (and, one would hope, learn from them). This may be true in many cases but does not apply when there is no opportunity to learn from experience, as would be the case if one discovered that one's retirement savings



were insufficient only upon nearing retirement age. Moreover, this argument seems to reject the very premise of light paternalism— that it is possible to implement paternalistic policies that *do not* restrict individual autonomy or, at worst, do so very minimally. Additionally, paternalistic policies do not preclude learning. Steering individuals toward a welfare-enhancing choice in one situation will be met with positive reinforcement and facilitate learning, which can inform the individual's decisions in other situations.

Finally, in Chapter 1, Gul and Pesendorfer do not provide any kind of principled argument against light paternalism, but one that is based purely on convention. They argue that whether such interventions help or hurt economic agents is irrelevant because economists simply should not be in the business of directing social policy. “The standard approach” to economics, Gul and Pesendorfer argue, “assumes a separation between the economist’s role as social scientist and the role that some economists may play as advisors or advocates.” They dub the economist who crosses that dividing line an “economist/therapist.”

Although Gul and Pesendorfer seem to view “therapist” as a pejorative label, we see no reason to not embrace it. Therapy is, in fact, not a bad metaphor for the new types of policies that behavioral economists have been proposing. Much like a therapist who attempts to steer clients toward more beneficial thoughts and behaviors without forcing them to do anything, all of these variants of light paternalism retain the ultimate autonomy of the individual while at the same time attempting to guide individuals toward courses of action that are seen as advantageous. Just as a psychotherapist endeavors to correct for cognitive and emotional disturbances that detract from the mental health of the patient, the economist/therapist endeavors to counteract cognitive and emotional barriers to the pursuit of genuine self-interest.

## **Methodological Issues Underlying Light Paternalism**

Although light paternalism is a “growth industry” in economics, it is not yet sufficiently “mature” as an enterprise to have developed standard operating procedures or for its practitioners to have fully thought out the range of methodological issues that it raises. The purpose of this chapter is to begin to address this void in the literature by exploring some of the issues that light paternalism raises for economic methods.

The first issue is how a particular pattern of behavior should be judged as a mistake and, relatedly, how the success of paternalistic policies designed to rectify such mistakes should be evaluated. That is, an informed application of paternalism, whether light or not, requires some form of welfare criterion. Clearly, the traditional welfare criterion used by economists, which involves satisfying people’s preferences to the maximum extent possible, cannot be used to evaluate policies that are premised on the view that people do not always choose what is best for themselves. We discuss the question of what type of welfare criterion should be used to evaluate paternalistic interventions.

Second, paternalism, and especially light paternalism, introduces new motives for attempting to understand the psychological processes underlying economic behavior. An enhanced understanding of process can help to explain why people make mistakes in the first place and, more importantly, can provide insights into what types of policies are likely to be effective in correcting the mistakes. We describe how an understanding of psychological process can inform, and already has informed, light paternalistic policies.

Third, in part because light paternalism is such uncharted territory, there is an acute need for testing different possible policies before implementing them on a large scale. There are good reasons why such tests should be carried out in the field rather than in the lab. Hence, the new paternalism points to the need for an expansion of field experiments—a trend that has already begun [DellaVigna, forthcoming].

In addition to methodological issues, there are pragmatic issues concerning who will implement light paternalistic policies, especially when they involve positive expenditures. We discuss how economic interests can be rechanneled to support endeavors consistent with light paternalism. In some cases, it may be in the interests of private sector industries to offer products or create incentives that help individuals to do what is in their own best interests. In other cases, the government can help align the interests of individuals and private industry. We conclude this chapter with a discussion of how recent trends in economic research on light paternalism relate to positive and normative economics.

In the course of discussing these methodological issues underlying light paternalism, we review a wide range of such interventions that have already been tested, as well as some that are still in the development phase. Therefore, a secondary purpose of this review is to give readers unfamiliar with the topic an overview of the wide range of light paternalistic interventions that are already being implemented and tested.

## **What Welfare Criterion?**

---

In their paper introducing the notion of libertarian paternalism, Thaler and Sunstein [2003: 175] state that “a policy counts as ‘paternalistic’ if it is selected with the goal of influencing the choices of affected parties in a way that will make those parties better off,” and then continue, “We intend ‘better off’ to be measured as objectively as possible, and we clearly do not always equate revealed preference with welfare.” But what does it mean to measure “better off” “objectively”? As Thaler and Sunstein hint, preference-based measures of welfare are not up to the job because they equate utility with preference and hence automatically assume that anything a person voluntarily chooses to do must be welfare enhancing. Clearly, it does not make sense to assess whether someone is committing an error using a measure that is premised on the assumption that people do not commit errors.

In their discussion of asymmetrical paternalism, Camerer et al. [2003] propose, as the ideal, purely asymmetric paternalistic policies that help people who behave suboptimally but have little or no negative impact on who behave optimally. Some examples that fit this criterion include establishing defaults and framing alternatives so as to steer individuals toward advantageous alternatives, and possibly offering commit options to people with self-control problems.<sup>6</sup> However, Camerer et al. [2003] acknowledge that purely asymmetric policies are not always possible.<sup>7</sup> To extend the applicability of the approach, they propose a looser criterion which simply requires that the net benefit to irrational consumers must exceed the aggregate costs both to rational consumers and any other affected entities such as businesses or taxpayers. This criterion shifts the debate regarding paternalism from philosophical issues about autonomy and freedom to pragmatic issues of benefits and costs (with loss of autonomy potentially treated as a cost). Evaluating costs and benefits, however, once again requires some concept of welfare, and one that does not encode anything an individual does, or would do, as welfare improving by assumption. Several different types of welfare have been proposed that have this property.

### **Experience Utility**

One possible approach, advocated first by Daniel Kahneman, and subsequently embraced by a number of economists, is to base evaluations of welfare on empirically reported happiness, or what Kahneman labels “experience utility” (as distinguished from “decision utility,” which corresponds to the modern notion of preference inferred from choice). Layard [2005], for example, argues that maximizing happiness rather than income should be the goal of government policy, and others have argued that happiness data should be used to identify appropriate societal tradeoffs between, for example, inflation and unemployment [Di Tella, MacCulloch, and Oswald, 2003] or between money and airport noise [van Praag and Baarsma, 2005]. Others argue for making happiness a goal of policy, on the basis of evidence

that happiness leads to such positive consequences as higher incomes, better work performance, citizenship behaviors, stronger more stable relationships and better health [Diener and Seligman, 2004]. Happiness has a major advantage over revealed preference as a welfare criterion: it is independent of the choices that people make, and hence can be used to evaluate which choices are welfare enhancing and which detract from welfare. However, as discussed in detail by Loewenstein and Ubel [forthcoming], using self-reported happiness as a policy criterion has several problems.

One problem is that people adapt to both unfortunate and fortunate circumstances, such that after sufficient time they return to their original happiness “set point” (see Frederick and Loewenstein [1999] for review). For example, dialysis patients do not experience significantly different levels of happiness than do healthy controls, even when measured “on line” by multiple reports elicited randomly at different points in the day [Riis, Loewenstein, Baron, et al., 2005]. If we were to use experienced utility as a metric for evaluating welfare, we could not conclude that chronically poor health was an undesirable outcome, a result that few would endorse. Moreover, a recent study found that although colostomy patients reported similar levels of happiness to people who did not have colostomies, they also expressed a willingness to give up 15% of their remaining life span if it could be lived with normal bowel function (i.e., no colostomy) [Smith, Sherriff, Damschroder, Loewenstein, and Ubel, 2007]. Despite being about as happy as healthy people, these patients indicated that they placed a high value on having their former health restored. Measures of welfare based on experience utility would fail to pick up such preferences.

Additionally, there are serious problems with all existing measures of happiness. For example, people tend to naturally “norm” happiness scales to their general circumstances or those of the people around them [Kahneman and Miller, 1986; Ubel, Loewenstein, Schwarz, and Smith, 2005]. Happiness scales are also sensitive to a wide range of non-normative

factors, such as current mood, the weather, and earlier questions in the survey [Kahneman and Krueger, 2006]. Finally, existing measures of happiness may miss brief periods of intense grief or regret that might have a substantial negative effect on well-being. Even the best measure of experience utility, using experience sampling techniques, can only measure happiness several times a day. In sum, while happiness measures may provide useful inputs into public policy, it would be a major mistake to base such policies solely on measures of happiness.

### **Limiting Welfare to “Valid” Choices**

An alternative approach, advocated by Bernheim and Rangel (chapter 7), is to adhere to a choice-based measure of welfare (i.e., “decision utility” in Kahneman’s parlance) but to limit the range of choices that “count” as indicative of welfare. Intuitively, their idea is that a person’s choices usually promote their well-being, but in some limited situations, such as when a person is overwhelmed by drives or emotions, they may not. Their proposal, therefore, is to adopt a welfare criterion that, in effect, surgically removes “bad” choices from the set of choices that count.

The crux of the problem is then to specify which choices count and which do not. Bernheim and Rangel consider several alternative means of selecting which choices should count, such as “preponderance” (only selecting choices that are made with some frequency) and “self-officiating” (allowing the individual to decide which subset of choices should be taken as valid indicators of welfare), but find objections to all. Ultimately, they conclude that determining which choices are commensurate with welfare and which are not will require “nonchoice data,” such as evidence from brain scans to determine when decision making is overwhelmed by visceral states or distorted by “circumstances where it is known that attention wanders, memory fails, forecasting is naive, and/or learning is inexplicably slow.” As they express it, “In these instances, we say that the [generalized choice criterion] is

suspect.”

Although such an approach might be useful in theory, we suspect that it will be many years, if ever, before we are able to interpret patterns of brain activation to make inferences about what types of choices should count as welfare enhancing. How, for example, could patterns of brain activation help to differentiate the many legitimate, intense, pleasures that short-circuit rational thinking (and, indeed, are sometimes all the more pleasurable for doing so) from intense impulses that lead us to behave contrary to self-interest? Likewise, it seems questionable that social scientists will come up with a way to distinguish between the excitement of buying something one really wants and the excitement of squandering part of one’s nest egg on a worthless trinket. In practice, we suspect, adjudicating between self-interested and non-self-interested choices will need to be done at least partially on the basis of an evaluation of which behaviors are most likely to confer long-run happiness—that is, on the basis of experience utility. Despite their explicit rejection of experience utility as a welfare criterion, therefore, we suspect that adoption of Bernheim and Rangel’s criterion would inevitably lead to an implicit reliance on judgments of experience utility, albeit in a more subjective and less systematic fashion.

### **Informed Decision Utility**

Another possible approach discussed, but not advocated, by Loewenstein and Ubel [forthcoming] involves honoring people’s choices as a utility-maximizing welfare criterion, but only if attempts are made to ensure that the decision maker is truly informed. Like the approach proposed by Bernheim and Rangel in chapter 7, this is a choice-based approach, but one that seeks to improve the quality of choice by providing decision makers with information rather than by selecting out a subset of choices that are deemed representative of welfare based on nonchoice data. Informed decision utility would include, but goes well beyond, such measures as food and drug labels. Beyond information labels, such an approach

might involve providing warnings about potential decision biases, such as how framing an outcome as a loss or a gain can lead to inconsistent choice.

Further, in situations in which information, however detailed and accurate, fails to provide a real anticipation of consequences, elaborate interventions could be devised to truly inform decision making. For example, one existing program intended to discourage childbearing by those who are not ready for it provides teenagers who are deemed at risk for pregnancy with dolls that require constant attention. The rationale is that, absent such a vivid experience, girls may have an overly romantic view of parenting, even if they are provided with more pallid information about the demands of parenting. Similarly, while smokers may appreciate the health risks of smoking at an abstract level, and may even overestimate such risks, they may not truly understand what it is like to die of lung cancer. In such a situation, again, more innovative interventions might be necessary to truly inform decision making.

The informed decision utility approach, however, suffers from two significant problems. The first is very similar to the fundamental weakness of the approach proposed by Bernheim and Rangel; in practice it is unlikely to avoid the need for recourse to judgments of experience utility. Given the wide range of different informational interventions that are possible, it will be necessary to decide which ones are worthwhile and which are not. The very act of providing information may frame a decision in a particular way that influences decisions in a particular direction, so it will also be necessary to decide how information intended to inform decision utility should be presented. For example, differences in small risks can be made to seem dramatic if they are presented in terms of ratios or percentages (e.g., “regular exercise can reduce your risk of disease X by 100%” ) as opposed to absolute terms (e.g., “regular exercise can reduce your risk of disease X by .0001—from .0002 to .0001”). Deciding which decisions to inform and how to inform them, therefore, will require some independent welfare criterion, the lack of which is the very problem that informed



decision utility was intended to solve. As was true for the choice-subset notion proposed by Bernheim and Rangel, therefore, we suspect that in practice such decisions are going to be informed, at least in part, by recourse to judgments about which types of information will make decision makers happy or well off in some other sense—that is, by experience utility.

The second problem is that informational interventions are effective against only one of the two broad categories of mistakes that people make—those that result from incorrect information—and not against the other: self-control problems. As noted in the introduction to this chapter, there are many situations in which people lose control of their own behavior and knowingly behave in ways that they know are not in their own long-term self-interest. While information might help people to avoid such situations, once one is in the situation, the most accurate information that it is possible to impart is unlikely to have much if any impact on behavior.

## **Capabilities**

Yet another approach, advocated by Amartya Sen [1985, 1992] and elaborated on by Martha Nussbaum [2000] is the capabilities approach. This approach is specifically intended to deal with, among other problems, that of adaptation. It rejects the revealed-preference framework for measuring welfare because people adjust their preferences as they adapt to poor social and physical conditions, characterized by poverty or injustice that, most people would agree, objectively reduce the quality of life. In other words, preference and desire can be diminished by “habit, fear, low expectations, and unjust background conditions that deform people’s choices and even their wishes for their own lives” [Nussbaum, 2000: 114]. Sen [1985] gives the example that a person living in impoverished conditions may learn to have “realistic desires” and derive pleasure from “small mercies” and, as a result, may have more desires met than a person in dramatically better living conditions with overambitious desires. Note that this problem with a revealed-preference framework is similar to the

problem of adaptation that we discuss in relation to using experienced utility as a welfare criterion. Just as adaptation causes problems for hedonic measures of welfare because people adapt hedonically to situations that virtually everyone would agree are adverse, it can cause problems for preference-based measures if people adapt their preferences to their circumstances and, as a result, become satisfied in situations that would be widely deemed to be unsatisfactory.

The solution proposed by Sen and Nussbaum is to construct a normative theory of welfare that is based on human capabilities—that is, what people are capable of achieving based on the opportunities and living conditions afforded them. Nussbaum delineates several “central human functional capabilities,” such as health, freedom from assault, political voice, property rights, equal employment, and access to education, which resemble basic human rights, as well as others that involve self-actualization, such as emotion expression, affiliation with others, and recreation.

The capabilities approach avoids the problem of hedonic adaptation, which is one of the central weaknesses of the experience utility approach. It also avoids the problem of the standard revealed preference approach of treating anything that someone does as welfare enhancing. However, the capabilities approach suffers from crippling problems of its own. Specifically, the approach is impractical to implement because policy makers are unlikely to reach a consensus about which capabilities should be valued and, even if a set of valued capabilities can be agreed upon, the relative values of those capabilities. However, there are similarities between this welfare criterion and the one we propose below. At some point, policy makers should have some discretion to impose “values,” such as the improvement of health or the reduction of poverty, on others, even if these changes are not deemed necessary by a preference-based or experienced utility welfare criterion—particularly if it can be done without limiting individual autonomy.

## **An Imperfect but Pragmatic Approach**

What welfare criterion, then, should be used? We suspect that in most instances the problem will not be as severe as it seems. Although the threshold for light paternalism can be and should be lower than that for more heavy-handed forms of paternalism, we would still advocate that even light paternalistic policies should only be put into play when welfare judgments tend to be relatively straightforward. This is the case when one of the following conditions is met:

1. *Dominance*: In some cases, such as the failure of employees to take advantage of company matches on retirement accounts, a simple dominance criterion will suffice. In the case of company matches, as long as employees have monotonic preferences—that is, prefer more income over less income—they will be better off if they maximize their own contribution, at least up to the level of the maximum company match. The underutilization of 401(k) matching programs most convincingly illustrates that many people do not save optimally, since failing to take advantage of such a match effectively “leaves money on the table.” This is the case even after taking into account tax penalties for early withdrawal. The mistake is particularly egregious, and by no means rare, when an employee past retirement age does not make the maximal allowable contribution, since in this case the contribution could be made, matched, and then both the contributed funds and the matched funds withdrawn the next day without penalty [Choi, Laibson, and Madrian, 2005]. Thus, from the perspective of the employee, a default contribution equal to the level of the maximum company match makes perfect sense.

A somewhat weaker form of dominance is, “stochastic dominance,” which involves minimizing risk at any level of return, or maximizing return at any level of risk. The case of including one’s own company’s stock in a retirement portfolio would seem to come close to violating stochastic dominance.

2. *Clearly negative outcomes*: Given certain circumstances, people make decisions that lead them down a detrimental path. The resulting outcomes are clearly undesirable, unintended, and not in the decision maker's self-interests. In these cases, a precise welfare criterion is not required because it is clear that people would be better off if they could avoid these negative pitfalls. For example, using a regression discontinuity model, Skiba and Tobacman [2006] found that people who use payday loans have a higher chance of filing for Chapter 13 bankruptcy relative to people who were not approved for the loan. Bankruptcy is a clearly negative outcome leading to filing costs, reorganization of debt, and a 10-year stigma on one's credit report. The shocking statistic that there are more payday loan establishments in the United States than there are McDonalds suggests, at minimum, that government policies which encourage or offer alternative forms of credit could be welfare enhancing for many people.<sup>8</sup>
3. *Self-officiating*: Despite Bernheim and Rangel's dismissal of this criterion in chapter 7, which effectively lets people choose their own goals and then helps them to achieve them through restrictions, incentives, or information to aid self-control, we think this criterion is generally a good one, assuming that the choice of goals is not done in the heat of the moment. If people who are overweight consistently believe that they would be better off were they not overweight, and consistently report that some proposed light paternalistic policy would make them better off, this would seem to be another relatively straightforward situation in which light paternalism can be justified. Thus, for example, if employees at a company themselves decided that they would be better off if, to avoid exposure to temptation, no soda machines were on the premises, a self-officiating criterion would dictate that soda machines should be removed. This is, admittedly, a form of heavy-handed paternalism. A lighter version would keep the soda machines on premises but engineer a system that renders them operable only by employees who have

elected ahead of time to give themselves access.

Bernheim and Rangel are very explicit in advocating a welfare criterion based on choice rather than on preference. Our own opinion is that the welfare criterion for evaluating paternalistic policies should be based on preference. Much as a psychotherapist would likely take at face value a client's professed desire to become happier, more sociable, or less anxious, even if she engaged in patterns of thinking and behavior that led to the opposite result, we would argue that the economist-as-therapist should treat verbal statements of preference as useful information, even if choice is not in line with professed preference. If people express a desire to lose weight but make choices that cause them to gain weight; if they express a desire to be financially solvent but make choices that lead to burdensome debt; if they want to stop smoking but continue to smoke; if they want to take prescription medications but fail to do so, these are all situations in which paternalistic interventions could be helpful. Indeed, the very hallmark of a situation in which paternalism may be justified is a divergence between stated preference and choice. Only in cases where such divergence exists should light paternalistic policies be devised, and they should endeavor to bring choice more in line with stated preference.

As further developments in the measurement of welfare occur, it may ultimately be possible to come up with less conservative measures of welfare that allow for a useful balancing of costs and benefits. Perhaps more fine-grained, domain-specific measures of experienced utility will help get around current problems with the measurement of happiness, allowing for the identification of a broader range of beneficial light paternalistic interventions. Until that happens, however, we would advocate that even light paternalistic policies only be enacted in the clear-cut situations just enumerated.

## **The Importance of Process**

---

Light paternalism provides new motivation for looking inside the “black box” of human behavior. A better understanding of the processes underlying economic behavior can help to identify when light paternalistic interventions would be helpful and, perhaps more importantly, can help to inform the policies themselves. As we show below, many light paternalistic interventions exploit nonstandard behavioral regularities (e.g., loss aversion, hyperbolic time discounting, and the status quo bias), which ordinarily undermine the optimality of decision making, to instead enhance it.

To illustrate the point, consider the Save More Tomorrow (SMarT) program designed and implemented by Thaler and Benartzi [2004]. The program was designed to deal with the problem that many employees fail to take advantage of the tax breaks and company matches on 401(k) plans and, as a result, fail to save adequately for retirement. The failure to save adequately for retirement stems in part from hyperbolic time discounting (which overweighs the pleasures of current consumption over the pleasures of deferred consumption), loss aversion (because putting money into 401(k) plans is seen as a cut in take-home pay), and the status quo bias (which, when the default contribution rate was zero, encouraged noncontribution).

Employees at companies that participate in the SMarT plan are asked if they would increase their 401(k) contribution rates beginning at the time of their next pay raise. Since the contribution rate does not increase until after a raise, employees do not perceive the increased savings as a cut in take-home pay. Once employees sign up for the plan, they remain enrolled, and the process repeats each year until they reach the maximum contribution rate, unless they opt out. The SMarT plan is designed to make biases that typically discourage saving, such as hyperbolic time preferences, loss aversion, and the status quo bias, work instead to promote saving.

Hyperbolic time preference, a concept first identified by Strotz [1955], refers to the tendency for people to be more impatient in the present (when trading off present against future gratifications) than they are with respect to the future (when trading off future against even more future gratifications). As Strotz [1955:177] expressed it, hyperbolic time discounting implies that individuals who

naively resolve now what they “will do” in the future, commonly do not schedule the beginning of austerity until a later date. How familiar the sentence that begins, “I resolve, starting next[...]” ! It seems very human for a person who decides that he ought to increase his savings to plan to start next month, after first satisfying some current desires; or for one to decide to quit smoking or drinking after the week-end, or to say that “the next one is the last one.”

The SMarT program plays directly on these inclinations, presenting people with the option of doing what comes naturally—spending in the present but saving in the future—a plan that is especially attractive to people with hyperbolic time preferences.

The program also takes account of loss aversion, which describes the tendency for people to put greater weight on the psychological cost of a loss than on the psychological benefit of an equivalent gain. Due to loss aversion, people are more likely to tolerate a forgone gain than a loss of equal value. The program removes saving from future wage increments (perceived as a forgone gain) rather than having people simply contribute out of income (perceived as a loss).

If that were the whole story, of course, the SMarT plan would not work, because when tomorrow became today people would once again prefer spending over saving. However, at this point another factor comes into play that weighs against such an outcome: The program exploits the status quo bias to maximize continuing adherence by putting into place a decision rule (save a certain fraction out of future wage increases) that remains in effect unless it is explicitly rescinded.

This combination of ingredients seems to work. Initial evaluations of the program found that enrollment was very high (78%), that very few who joined dropped out, and that there

were dramatic increases in contribution rates (from 3.5% to 11.6% over 28 months).

### **Harnessing Decision Biases to Improve Decision Making**

Redirecting patterns of behavior that usually hurt people to help them instead is a common pattern among light paternalistic interventions. In this subsection, we discuss a variety of behavioral regularities that can be exploited by the economist/therapist.

#### *The Importance of Immediate Feedback and Reinforcement*

In the discussion of hyperbolic time discounting in connection with the SMarT plan, the emphasis was on not imposing immediate out-of-pocket costs on program participants. An even more important implication of hyperbolic time discounting is the need to design interventions that provide participants with very immediate costs and benefits—that is, reinforcement—as well as feedback about their behavior.

Thus, for example, hyperbolic time discounting probably plays a role in drug addiction (because the benefits of taking a drug are immediate and the consequences delayed), and one of the most effective treatments of addiction exploits hyperbolic time discounting to provide addicts with short-term financial incentives to quit [Higgins, Wong, Badger, Ogden, and Dantona, 2000]. Addicts are given coupons for consumer goods each day when they come in for treatment if their urine sample is negative for drug use. Most of the addicts treated in this program have experienced devastating losses as a result of their addiction, and would seem to have every incentive for quitting. But these small payments often succeed where much larger benefits fail, probably because they are delivered with a frequency that resembles that of drug-taking itself. A general principle is that many suboptimal patterns of behavior are caused by the overweighting of immediate costs and benefits, and hence any attempt to deliver incentives to overcome such patterns needs to provide incentives that can be small but must be frequent.



A line of research in which this insight is already well understood has involved using financial incentives to combat behaviors resulting from self-control problems. Financial incentives have been used to get people to stop smoking [Volpp, Gurmankin, Aschet al., 2006], lose weight [Jeffrey, Thompson, and Wing, 1978; Jeffrey, Gerber, Rosenthal, and Lindquist, 1983], stop taking addictive drugs such as heroin, cocaine, and cigarettes [e.g., Higgins et al., 2000; Heil, Tidey, Holmes, and Higgins, 2003], and get better grades [Angrist, Lang, and Oreopoulos, 2006]. Such interventions can be seen as an even more extreme version of “light” paternalism in that, not only is participation voluntary, but also the introduction of financial incentives (assuming they are rewards and not punishments) actually puts individuals into financial positions that are better than their positions before the intervention. Although people may know that in the long run it is in their best interests to diet, take their medications, or stop using illicit drugs, they often have difficulty implementing such decisions. Financial incentives seem to help mainly by offering short-term payoffs that bring the short-term incentives in line with long-term self interests.<sup>9</sup>

This insight can and should be, but to the best of our knowledge has yet been, applied to savings behavior.<sup>10</sup> Thus, many interventions to increase saving involve attempts to make the prospect of a destitute (or prosperous) retirement more salient to individuals, for example, by presenting vivid images of people suffering poverty in retirement. Such interventions are unlikely to have much of an impact because the prospect of retirement is so remote when people need to begin saving, and because any one day or even month of saving constitutes an inconsequential “drop in the bucket.” Savings interventions that provide people with more immediate and frequent reinforcement are more likely to succeed. Short-term success in implementing saving plans could be reinforced by providing people with daily or weekly feedback of the form: “If you continue to save at this rate, this is where you will be at retirement.” And achieving short-term saving goals—even at a daily or weekly level—could

be reinforced through small rewards, including lottery prizes. Much as addicts respond to small, immediate gift vouchers, even after failing to respond to the seemingly much larger life benefits of being drug free, it is very likely that small short-term rewards for saving could have an impact that the objectively much larger prospect of a prosperous retirement does not.

### *Overweighting of Small Probabilities*

It is well established that people tend to overweight small probabilities, which contributes to, among other things, the attractiveness of playing the lottery. Although playing the lottery is often viewed as self-destructive, the overweighting of small probabilities can be exploited to individuals' benefits by using it to magnify the value of rewards. Thus, in an ongoing collaboration with Kevin Volpp, Stephen Kimmel, and Jalpa Doshi at the University of Pennsylvania, the first author has been providing people with a lottery-based incentive to take warfarin—a medication that dramatically lowers the likelihood of a second stroke at minimal cost and with few side effects if taken regularly. Patients get an electronic drug dispenser that electronically signals a central office if the correct drawer has been opened on a particular day, indicating that the patient, in all probability, took the pill. Every evening, a number is drawn and, if the number matches the patient's personal lottery number and the drawer was opened during the day, the patient receives a substantial cash prize. The incentive mechanism plays not only on the overweighting of small probabilities, but also on regret aversion—the distaste for being in a situation in which one would have experienced a better outcome had one taken a different action. It does so by informing participants who fail to take their drug during the day and who win the lottery that they would have won had they taken the drug. The research on drug adherence is funded by an insurance company that is interested in the possibility that the program could be cost-effective if the cost of promoting adherence is lower than the cost of caring for the people who would have stokes as a result of failing to adhere to their drug regimen. Playing on the overweighting of small probabilities

and regret aversion increases the “bang for the buck” and hence the likelihood that the program will be cost-effective. Initial results are promising; two pilot-tests of the intervention, each involving 10 patients followed for one month, resulted in an increase in adherence rates from a baseline of 66% to adherence rates of 96% in one study and 97% in the other.

### *Loss Aversion*

A second program, currently being pilot tested with obese U.S. veterans who want to lose weight, and developed by Volpp, Loewenstein, and Carnegie Mellon University graduate student, Leslie John, is an incentive scheme for promoting weight loss that involves “deposit contracts.” In an innovative study, Mann [1972] found that participants who deposited money and other valuables with a therapist and signed contracts in which return of their valuables was contingent on progress toward pre-specified weight loss lost tremendous amounts of weight: an average of 32 pounds. A subsequent study that also involved deposit contracts produced similarly stunning results, with 47% losing more than 20 pounds and 70% losing more than 15 pounds. In contrast, interventions in which people have simply been paid for weight loss have produced more modest results.

In our in-progress intervention, people who are already motivated to lose weight (a precondition for this being treated as an instance of light paternalism) are invited to deposit an amount up to \$90 per month (\$3 per day), which the experimenters match one for one. The individual then receives a payment of two times the daily amount deposited for every day that his weight falls below a line that entails losing one pound per week. Deposit contracts play on loss aversion, but instead of playing on the underweighting of forgone gains (as does the SMarT program), it plays on the relatively greater weighting of out-of-pocket costs, which renders especially distasteful the prospect of losing one’s own deposited money, as well as the experimenter’s match. Deposit contracts also play on optimism, which encourages obese

people who want to lose weight to put their own money at risk in the first place. The hope is that, when combined with the subsequent motivational force of loss aversion, optimism about future weight loss will become self-fulfilling.

### *Framing Effects*

Diverse lines of research show that changing superficial features in the presentation of a decision can produce predictable shifts in preference. Such “framing effects” can be exploited to help people make beneficial decisions and, at the very least, should be taken into consideration when presenting people with important information they need to make decisions about government programs, investment decisions, medical decisions, and so forth. Making use of framing effects is consistent with asymmetric paternalism in that it does not limit choice in any way, but can be used to help people make beneficial decisions. Similarly, it is consistent with the guiding principle of libertarian paternalism that information must be presented in some way to the public, so why not present it in a fashion that is advantageous to its recipient? Recent research by Schwartz, Bertrand, Mullainathan, and Shafir [2006] takes advantage of framing effects and loss aversion to increase take-up into employer-sponsored health care flexible spending accounts, which are economically beneficial for the vast majority of employees. Contribution rates were higher when the decision was framed as a loss (“Stop losing money now”) compared to when the decision was framed as a gain (“Start saving money now”).

### *Goal Gradients*

In another program at an even more preliminary stage of development, the two authors have been developing innovations to increase the efficacy of Individual Development Accounts (IDAs). IDAs are matched savings accounts that allow low-income families to accumulated assets to purchase a home, pay for education, or start a small business. One of

these innovations involves changing the schedule of deposit goals from a constant goal each month to a schedule based on the goal gradient hypothesis, first proposed by Hull [1932], which states that effort and motivation increase as one gets closer to completing a goal. This principle has been shown to apply to consumer behavior in reward programs, including the finding that even the illusion of progress toward a goal or, in this case, a reward can increase purchases [Kivetz, Urminsky, and Zheng, 2006].<sup>11</sup> Consistent with the goal gradient hypothesis, the schedule of savings deposits starts very small, increases slowly, and is highest right before the savings goal is met. This feature also makes the plan attractive to people with inconsistent time preferences who weigh immediate consumption much more heavily than future consumption. Initial payments will reduce current consumption only marginally, while the larger payments at the end of the plan reduce consumption more significantly but are heavily discounted.

### *Summary*

The foregoing examples illustrate how, consistent with chapters in this volume that argue against a strict revealed preference approach, an understanding of human psychology can help us both to understand the causes of self-destructive behavior and to devise policies intended to counteract it. New developments will inevitably lead to creative new policies. For example, new research on the neural underpinnings of intertemporal choice [e.g., McClure, Laibson, Loewenstein, and Cohen, 2004] are drawing attention to the important role played by affect in many self-control problems. By drawing on insights about affect—namely, the tendency for “hot” emotions to “cool off” over time—this research may help inform and further the reach of cooling off regulations which already exist in a wide range of domains (e.g., when it comes to door-to-door sales). A challenge for future research will be to kindle the motivational force of hot emotions for beneficial rather than self-destructive ends.

## **The Need for Expanded Field Research**

---

Conventional economists sometimes accuse behavioral economics of being rife with different effects (e.g., as discussed above, loss aversion, hyperbolic time discounting, and regret aversion), with competing effects sometimes coming into play simultaneously, making it difficult to predict the net impact of a particular exogenous change. There is some validity to this charge, although this state of affairs may reflect the real complexity of human psychology rather than any limitation of behavioral economics. People have different identities and behave differently depending on which identity is activated in a particular situation [LeBoeuf and Shafir, 2005]. They come to decisions “armed” with an array of different “choice heuristics,” and which they employ depends on what type of situation they view themselves as facing [Frederick and Loewenstein, 2006]. At a more physiological level, behavior is the product of multiple neural systems that often act in concert but in some cases come into conflict [e.g., Sanfey, Loewenstein, McClure and Cohen, 2006]. The consequence is that small changes in circumstances or institutions can sometimes have large unforeseeable effects on behavior.

The multiplicity of psychological effects decreases the predictability of individual responses to policy interventions, and, as economists understand particularly well, interactions between individuals create further opportunities for unpredictable effects. To avoid unintended consequences, therefore, there is a pressing need for careful testing of specific interventions before they are implemented on a broad scale. Careful small-scale pilot testing is essential to ensure that the benefits of a large-scale implementation will outweigh the societal costs. Although we do not endorse what seems to be an emerging hostility toward laboratory studies [e.g., Levitt and List, 2008], there is probably no substitute for field studies when it comes to testing light paternalistic interventions.

An example of a paternalistic intervention with unexpected and unintended consequences

was the “Move to Opportunity” experiment that was conducted in several major U.S. cities in the 1990s [Katz, Kling, and Liebman, 2001]. Although not an example of light paternalism, the study is useful for illustrating the utility of field experiments as a tool for evaluating any kind of paternalistic intervention. Families receiving subsidized housing were randomly assigned to one of three conditions: a group given a restricted housing voucher that could only be used in low-poverty neighborhoods (less than 10% below the poverty line), a group given an unrestricted housing voucher, and a control group. The purpose of the study was to provide the first unconfounded test of the impact of neighborhood characteristics on economic and noneconomic outcomes. Although not framed by its developers as a test of paternalism, providing restricted vouchers can be interpreted as a form of paternalism, since they limit the choices of those who receive them, presumably with their best interest in mind.

The results of the Move to Opportunity experiment were surprising [Kling, Liebman, and Katz, 2007]. Although moving to a more economically advantaged neighborhood did have some beneficial effects, especially for girls, it also had some surprising negative effects that were concentrated mainly among boys. Girls had beneficial outcomes in the areas of mental health, educational outcomes (staying in school, reading and math achievement), risky behaviors (alcohol use, cigarette use, and pregnancy), and physical health. However, for boys there were substantial negative effects on physical health and risky behaviors. Results for adults were also disappointing. Contrary to expectations, there was no evidence of economic improvement in earnings, employment, or welfare usage for adults. Follow-up interviews indicate that these effects may be due to disrupted social networks and transportation difficulties. However, there were significant beneficial effects for adult obesity and mental health.

The Move to Opportunity study underlines the importance of testing paternalistic interventions on a small scale, but in the field. Although moving poor families into affluent

neighborhoods may have clear benefits, such as increasing the safety of children, there may be a host of unintended consequences that could not have been anticipated at the outset. Moreover, the disappointing results from the Move to Opportunity study underscore the importance of collecting information about process, which was the theme of the preceding section. Beyond the disappointing results of the program itself, an unfortunate aspect of the research component of the program was the failure to collect sufficient qualitative data to shed light on why the program produced some of the perverse results that it did. Such process data could be used as an input into developing an improved follow-up program.

Whatever its limitations when it came to monitoring process variables, the Move to Opportunity program did provide extremely good outcome measures, which enabled a very clear delineation of its effects. This is an essential practice that should be applied more diligently in other field evaluations of light paternalism, and that applies most significantly to what is unquestionably the most important application of light paternalistic policies to date: interventions to increase saving.

As already touched upon, a number of researchers have tested interventions designed to encourage people to save more of their income. Note that these interventions are paternalistic in the sense that they assume that people do not naturally save as much as they want to or should. They are “light” in the sense that all are voluntary; none force people to save money. Although some do impose restrictions on withdrawals from savings, these are purely voluntary. These studies have employed a wide range of methods.

Several “natural experiments” (or “quasi experiments” as the psychologists who developed the techniques refer to them; e.g., Campbell [1969]), have examined the effects of increasing default contributions on increasing participation and contribution rates to 401(k) plans (see Choi, Laibson, and Madrian [2004] for review). These studies track changes in the savings and investment behavior of employees at companies that abruptly change some



aspect of their policy. Presumably, such a change in policy does not coincide with an equally sudden and simultaneous change in the preferences of employees. Such studies show that simply by changing the default from unenrolled to enrolled dramatically increases enrollment, even though in either case the employee retains total decision-making autonomy, making this a near-perfect example of asymmetric paternalism [Choi et al., 2004; Madrian and Shea, 2001]. Employees are also highly influenced by the default level of contribution and the default for the asset allocation among available investment funds, underscoring the need to set optimal default contribution rates and diversification strategies.

Other research examining interventions to promote saving has involved field experiments in which a variable of interest was manipulated exogenously. For example, Duflo and Saez [2002] examined the impact of an educational intervention to increase enrollment [Duflo and Saez, 2003]. A random sample of employees in a subset of departments were offered a \$20 payment for attending an informational fair, and their 401(k) contributions were tracked as well as those of their coworkers. The most interesting finding from the study was that social information plays an important role in participation in 401(k) plans. Enrollment was significantly higher in departments where some individuals received the monetary inducement to attend the fair than in departments where no one received the inducement. However, increased enrollment within these treated departments was almost as high for individuals who did not receive any monetary inducement as it was for individuals who did, demonstrating the influence of social information.

Another field experiment focusing on saving examined the interest in, and response to, the introduction of a voluntary commitment savings product that restricted access to deposits [Ashraf, Karlan, and Yin, 2005]. Existing customers of a bank in the Philippines were randomly assigned to one of three conditions: a commitment group who were given the option of opening the restricted account, a marketing group who received a special visit to

encourage savings, and a control group who were not contacted. Twenty-eight percent of the commitment group enrolled in the restricted account. After 12 months, individuals in the commitment group were significantly more likely to have increased their savings by 20% than were participants in the marketing group or the control group. Average savings balances of the commitment group increased by 81% relative to the control group. Further, this study sheds light on which individuals are most likely to enroll in restricted savings accounts. Results of a pre-experiment survey show that impatience over immediate trade-offs, but patience over future trade-offs (consistent with hyperbolic discounting), predicts program enrollment, particularly for women.

A major, although seemingly unavoidable, limitation of all of these studies is the paucity of outcome measures that were collected. All of the studies of saving behavior examined the impact of, for example, changing retirement savings defaults on the affected account (the account for which the default rule is changed) but did not look at the impact on the overall financial position of the individuals and families involved. The problem with such a limited focus is that the change in retirement saving may have had other undesirable effects that were not measured by existing studies. If the increase in retirement saving comes out of frivolous consumption, that might be a good thing, but what if it leads to an increase in credit card debt, or a cutback of spending on nutrition or children's education? Without knowing the answer to these questions, it is difficult to come to any confident conclusion about the benefits of the seemingly "successful" programs to increase retirement saving. Indeed, even if it were shown that increasing retirement saving did not come at the expense of increased debt or decreased investments in human capital, it still would be difficult to evaluate the effects of such programs in a comprehensive fashion. For example, if the increase in retirement saving came out of vacation trips, is this necessarily a good thing? Might it be better for a family to take nice vacations while the children are young and then to live on a

shoestring during retirement?

Another limitation of most of the field experiments that have been conducted is their failure to manipulate program parameters in a fashion that, if an intervention were successful, would provide insight into what specific features of the intervention matter. For example, the Save More Tomorrow plan, which combines several features, has been proven successful in increasing saving. However, the relative importance of each specific feature is unclear. Thus, perhaps a program that committed people to save in the future but did not deduct that saving from future pay increases would work just as well as the current SMarT plan. Without studies that randomly assign participants to different configurations of plan features, we will never know the answer to questions of this type.

Beyond field research examining the impact of light paternalistic interventions, there is a need for basic research on topics that will inform the design and evaluation of effective policy. First, and consistent with the discussion above, the question of the optimal welfare criterion is in some sense an empirical question. Research could potentially address questions such as which criteria most closely mirror people's lay theories and values (e.g., whether people are more comfortable with choice-based or happiness-based policy decisions) and could also examine the types of trade-offs between autonomy and guidance that people endorse.<sup>12</sup> Additionally, to understand the trade-offs between different welfare criteria, it is important to have basic research on reliable and valid welfare measures. Progress has been made on the development of methodology to measure experience utility, such as with the use of ordinal scales to minimize the problem of scale recalibration and the use of experience sampling techniques [see Kahneman and Krueger, 2006; Riis et al., 2005]. Future research could focus on measures that correspond to different welfare criteria. For example, the self-officiating welfare criterion entails an attempt to ascertain what an individual desires most of the time, but preferences often fluctuate. Just as experience sampling has been used to

capture fluctuations in happiness over time, it could also be used to measure fluctuations in preferences over time.

Second, consistent with the need for expanded research on process discussed above, there is a need for basic research on topics that will inform the design of policy. For example, we still have an extremely imperfect understanding of the psychological factors leading to undersaving, overeating, and a variety of other problems. To what extent is undersaving due to the overweighting of immediate gratifications, to procrastination (the intention to start saving tomorrow and the belief that one will do so), the “drop-in-the-bucket” effect (the view that one small indulgence or act of self-denial will have a negligible impact on one’s overall level of saving), to overoptimism about future revenue sources, or a host of other possible contributing factors. A better understanding of why people fail to save could aid in the design of light paternalistic interventions. Similarly, many light paternalistic interventions involve giving people feedback and/or rewards for behaving in a self-interested fashion. However, we still have little understanding of what types of rewards are most motivating (e.g., lotteries vs. cash payments vs. in-kind rewards) or about what types of rewards pose the greatest threat of crowding out people’s intrinsic motivation to do what’s best for themselves.

Third, there is a need for new technologies to aid in the implementation and assessment of paternalistic interventions. For example, devices that measure weight, blood sugar levels, and blood pressure and that, like the electronic pill dispenser we have been using to improve warfarin adherence, permit two-way communications with a central administrator, could introduce a range of new possibilities for light paternalistic interventions.

## **Implementing Light Paternalism: Rechanneling Economic Interests**

---

Currently, there are a wide range of economic interests aligned, in effect, against consumers—entities that profit when, for example, consumers consume large amounts of food or alcohol, smoke cigarettes, play the lottery, incur credit card debt, or overdraw their bank accounts (incurring overdraft charges that provide a substantial flow of revenues to banks). These efforts are not necessarily driven by malicious motives; a company that failed to play on consumer weaknesses but faced competitors that did would be likely to lose business (see Loewenstein and O’Donoghue [2006] and Issacharoff and Delaney [2006] for a discussion of this issue).<sup>13</sup>

Admittedly, there are economic forces arrayed on the other side, for example, the diet industry, sellers of nicotine patches, and financial companies that benefit when people amass financial assets. But the forces that play on consumers’ weaknesses tend to be much stronger than those that bolster consumer defenses, and the motives of those arrayed on the other side are often ambivalent.<sup>14</sup> For example, nicotine patches are sold to people who are addicted to cigarettes, so their makers have, at one level, an interest in promoting addiction. Likewise, although the sellers of commercial diets would probably attract more customers if they were effective in promoting weight loss, they make the most money by selling hope rather than actual results. Hospitals similarly have the goal of curing sickness, but they have little motivation in promoting preventive medicine, which would just hurt their bottom line. An important goal for economists interested in light paternalistic solutions to such problems, therefore, is not only to devise clever solutions to suboptimalities in consumer behavior but to figure out creative ways to implement and fund such solutions.

In some situations, incentives for light paternalistic policies could be put into place via legislation or other forms of government regulation. For example, companies could be given

tax breaks that are dependent on employee contribution rates to 401(k) plans, in which case they could potentially be motivated to change defaults or, perhaps, introduce the SMarT plan. Through tax incentives or granting agencies, governments can promote business models that make it easier for individuals to act in their own best interests, such as nutritious and affordable fast food. The so-called “fat tax” is an example of a much more heavy-handed intervention that could work against the ever-declining prices of high-calorie foods, a situation that many economists hold responsible for growing levels of obesity.

In other situations, however, it is going to require the creativity of economists to play matchmaker and to identify areas of mutual interest that might not have spontaneously emerged without their intervention. Take obesity, for example. Although, as described, there are a number of economic entities (including, possibly, the medical industry) that stand to gain from obesity or the behaviors that cause it, there are also some economic interests that lose when people gain weight. Prominent among those who stand to lose are insurance companies. Although, as an industry, insurance companies may be indifferent to whether people are thin or fat, individual life insurance companies would benefit if their customers lost weight. If creative, low-cost interventions could be designed, therefore, it is quite possible that insurance companies would be motivated to underwrite the costs. Insurance companies would also be in a position to lobby for legislation that would allow them to adjust their rates based on the weight of a prospective customer, which would pass the economic benefits of weight loss on to consumers or their employers.

As another example, take drug adherence. Here, health insurers could potentially be motivated to provide funding for interventions that had the potential to reduce health costs. In fact, as already alluded to, the first author, along with researchers at the University of Pennsylvania, have secured funding from an insurance company to pilot test an intervention intended to increase adherence to warfarin—an antistroke medication. Pharmaceutical

companies also have a direct stake in drug adherence although their interests are somewhat more conflicted than those of insurance companies.

Saving is an example where there is a confluence of interests between customers and the bank. Further, people's difficulty in saving and desire to save more create a circumstance in which banks can even extract rents by aiding customers in saving more. A recent study conducted in the Philippines examined the impact of hiring deposit collectors, bank employees who come to customers' house to pick up savings deposits, a practice that is prevalent in some developing countries [Ashraf, Karlin, and Yin, 2006]. The use of deposit collectors increased savings by 25% relative to control groups, and people were willing to pay for this service. The study suggests that people are willing to pay because the service reduces the transaction costs of having to go to the bank, facilitates adherence to financial planning, and restricts the spending of spouses. Banks in the United States are just starting to take advantage of people's difficulty in saving to develop marketable products, such as American Express's "Savings Accelerator Plan" for their One Card that contributes 1% of eligible purchases into a savings account.

As a final example, consider lotteries. Despite the fact that state lotteries return only 50 cents on the dollar—the lowest payout rate of any form of legal gambling [Clotfelter and Cook, 1989]—in fiscal year 2003 Americans spent almost \$45 billion on lotteries, or \$155 for every man, woman, and child in the United States. Lotteries are played disproportionately by low-income individuals, with many studies finding that poor people put a larger fraction of their income into lotteries and others finding that they actually spend a larger absolute amount per capita. The purchase of lottery tickets by the poor could be considered a type of "poverty trap"—a cycle of behavior that prevents poor people from improving their situations.

The most obvious solution to this problem might seem to be to regulate the lottery, but

that is very unlikely to happen since the lottery generates a sizable amount of revenue for states, and because any restriction of availability is likely to lead to the reemergence of illegal, unregulated alternatives. A “rechanneling of economic interests” would entail that the financial services industry market investment alternatives that have lottery-like properties—i.e., that have a small cost and a small probability of yielding a large payout—but that, unlike lotteries and other forms of gambling, yield a positive expected return. Trying to “pull” people away from gambling and toward investing could potentially be much more effective than trying to “push” people away from gambling. The potential money amounts to be reaped are staggering, and allocating this money to capital formation instead of operating lotteries would be socially productive.

We believe that the key to selling these low-cost, high-risk investments is to make it possible to invest small amounts at a time and make the investments convenient to purchase on a daily basis. We have conducted experiments on state lottery ticket purchases in a low-income population and found that rates of ticket purchases are high when people make purchase decisions one at a time, that is, myopically. This finding can be explained in part by what is termed the “peanuts effect” [Prelec and Loewenstein, 1991; Weber and Chapman, 2005]. For each decision, the dollar they spend on a ticket is underweighted—that is, merely considered a “peanut”—and so they go for the gamble. However, rates of purchases are significantly lower when the decision to purchase several tickets is aggregated into a single decision. Then people are less likely to write off the amount necessary to purchase several tickets as insignificant.

This insight into decision making under uncertainty can be used to help low-income individuals to invest and to save. Though people may not be willing to take a substantial sum of money to invest (or may not have the self-control to save the minimum balances necessary to open an investment account), they may be willing to devote small amounts of money,



spread out over time, to investments options. The startup costs are quite high for the convenient sale of low-cost investments. However, there is a lot of potential to market other types of investments in addition to those designed to dissuade gambling, such as investments in equity index funds and savings in money market accounts.

The convenient sale of low-cost investments in a system that minimizes transaction costs by providing only a few investment options has great potential to increase the money that the average individual devotes to investing and saving, especially for low-income individuals and for those who typically play the lottery. At a minimum, investment companies should market investments as an alternative to gambling. An ad could feature two people, one who spends a dollar a day on the lottery, and show the money being put on a pile and then shrinking or burning, and one who invests it, and show the money accumulating gradually into a huge pile.

## **Conclusion: A Methodology of Normative Economics**

.....

Milton Friedman, in his famous 1953 paper “The Methodology of Positive Economics,” distinguished between two approaches to economic methods, which he termed positive and normative economics.<sup>15</sup> Friedman defined positive economics as a “body of systematized knowledge concerning what is,” which, he continued, could “provide a system of generalizations that can be used to make predictions about the consequences of any change in circumstances.” Normative economics, in contrast, encompassed a “body of systematized knowledge discussing criteria of what ought to be,” and a “system of rules for the attainment of a given end.” [Friedman, 1953: 3].

Although Friedman devoted most of his essay to a discussion of the methodology of positive economics, he did not dismiss the value of normative economics. Rather, he lamented that normative economics would be unavoidably contentious, because, he believed,

issues of values were much more difficult to resolve than issues of fact.<sup>16</sup> Friedman himself, of course, never shied from the normative [Krugman, 2007]. In fact, as typified by his famous *Free to Choose*, much of his professional life was devoted to arguing about what ought to be and what system of rules would be most successful in achieving his vision of the good society. Believing as he did in rational choice and the benefits of free markets, his conclusions were generally fairly predictable: eliminate regulations and eliminate any barriers to unrestricted competition.

In the last several decades, however, a new view of human behavior has taken root among many economists, one that recognizes through methods of positive economics limitations in people's pursuit of self-interest. Research on the psychology of decision making, the role of affect in decision making, and neuroeconomics have led to the recognition that human behavior can in some cases be suboptimal or even self-destructive, and have contributed to our understanding of when, why, and how deviations from self-interest occur. The new research has, in turn, spawned a whole new area of normative economics focused on the two elements of normative economics identified by Friedman: the measurement of welfare and the design of economic and social systems that maximize welfare.

Although embracing an interventionism that conservative thinkers such as Milton Friedman generally disdain, the new light paternalism can be viewed as in fact quite sympathetic to their arguments and philosophy. Eschewing traditional forms of heavy-handed command and control, light paternalism endorses diversity in policy experimentation, the use of market incentives rather than mandates, and the use of improved informational and feedback mechanisms to verify effects, push objectives, and guard against unintended consequences. Although light paternalism is still in its infancy, it has already produced insights into regulation and incentive design that are likely to have far-reaching consequences. Economists, we believe, should be and, as we have documented, to a very

great extent already are in the business of “discussing criteria of what ought to be” and attempting to devise economic institutions that maximize the likelihood that what ought to be in fact occurs. If this brands us economist/therapists, then we embrace this label with pride.

## NOTES

---

We thank Dan Akst, Erik Agner, Andrew Caplin, Peter Huang, Sam Issacharoff, Ed Glaeser, Ted O'Donoghue, Andy Schotter, Nachum Sicherman, Cass Sunstein, and Kevin Volpp for helpful comments and suggestions.

1. See [www.spendonlife.com/content/CreditCardDebtEliminationAndFactsAboutDebtInAmerica-1-223-3.ashx](http://www.spendonlife.com/content/CreditCardDebtEliminationAndFactsAboutDebtInAmerica-1-223-3.ashx).
2. According the Investment Company Institute, this includes equity and mutual fund holdings in employee-sponsored retirement plans ([www.ici.org/statements/res/rpt\\_05\\_equity\\_owners.pdf](http://www.ici.org/statements/res/rpt_05_equity_owners.pdf)).
3. American Gaming Association ([www.americangaming.org/Industry/factsheets/statistics\\_detail.cfv?id=7](http://www.americangaming.org/Industry/factsheets/statistics_detail.cfv?id=7)).
4. Murphy [2006].
5. There is a third class of people who could potentially be made worse off by a default. For example, a high savings default would not be optimal for people carrying high credit card debt, and these people may fail to “rationally” opt out of default. This point is further discussed in a later section that highlights the need for pilot testing and good outcome measures to ensure against net negative consequences of paternalistic policies.
6. However, one potential problem with precommitment options is that people who are in “cold states”—e.g., not hungry or not craving drugs—may be unable to fully appreciate the force of their own future motivation and hence may be excessively prone (i.e., more prone than would be optimal) to commit their own future behavior [see, e.g., Badger et al., 2007; Nordren, van der Plight, and Harreveld, in press].
7. Legislation that regulates information disclosure, e.g., the Federal Truth in Lending Act, is close to purely asymmetric and would probably satisfy this criterion. Other forms of information disclosure might be more questionable. For example, food labels can make one miserable if one fails to diet [Loewenstein and O'Donoghue, 2006]. Cooling off periods that require a mandatory waiting period for certain purchases or activities, such marriage, are even less strictly asymmetric. Cooling off periods are designed to prevent people from making mistakes when they are in a state of arousal that they will later regret, but they do impose real costs on those who must delay their purchase. In such situations, asymmetric paternalism can be justified only if the benefits (in this case, the utility that otherwise would have been lost from making purchasing errors) exceeds the costs for people who engage in the behavior regardless of visceral state (in this case, the cost of having to delay the purchase by those who do, in fact, want to make it).
8. There are approximately 30,000 payday loan outlets in the United States, which is about double the number of McDonald's restaurants.
9. A review of 47 studies on the effect of financial incentives to encourage preventative health care reveals that overall these interventions are successful, at least in the short run [Kane, Johnson, Tawn, and Butler, 2004]. The incentives were effective 74% of the time for simple preventative measures, such as vaccinations, and 72% of the time for complex preventative measures that required sustained behavioral change, such as weight loss. A variety of different types of incentives (cash, coupons, free medical care, lotteries, gifts, and punishment) were effective.
10. Individual development accounts offer financial incentives for saving through monthly matching and have been an extremely promising tool for helping low-income families build assets. However, like employee matching of 401(k) contributions, matching is not immediate and frequent enough to be an optimal reinforcer.

11. The motivational effect of the illusion of progress toward a goal was demonstrated by greater purchase acceleration when people were given a “buy 12 coffees, get one free” card with two preexisting bonus stamps than when they were given a “buy 10 coffees, get one free” card.
12. For an example of using empirical research to elicit values about policy trade-offs, see Ubel and Loewenstein [1996] and Ubel, Loewenstein, Scanlon, and Kamlet [1996].
13. Although not necessarily malevolent, in some cases these forces can take on sinister forms. For example, the “Center for Consumer Freedom (Promoting Personal Responsibility and Protecting Consumer Choice)” ([www.consumerfreedom.com/index.cfm](http://www.consumerfreedom.com/index.cfm)) describes itself (see the “About Us” link) as a “nonprofit coalition of restaurants, food companies, and consumers working together to promote personal responsibility and protect consumer choices” and as being in opposition to “the growing cabal of ‘food cops,’ health care enforcers, militant activists, meddling bureaucrats, and violent radicals who think they know ‘what’s best for you’ [that] are pushing against our basic freedoms.”
14. This is true even when it comes to the biggest success story to date for light paternalism: savings behavior. The first author had the experience of pitching an idea for increasing employee retirement saving to a company that offered an employer match, only to be discreetly informed that it wasn’t in the company’s interest to encourage its employees to save more since an increase in company matches would only detract from the bottom line.
15. In doing so, he drew on the earlier work of John Neville Keynes [1891].
16. In contrast to his respectful views of normative economics, Friedman was less favorable toward economists who ignore data altogether. Writing in 1953, Friedman failed to anticipate the remarkable methodological advances that were to occur in the next half-century, some of them enabled by the development of the computer. “One effect of the difficulty of testing substantive economic hypotheses has been to foster a retreat into purely formal or tautological analyses...economic theory must be more than a structure of tautologies if it is able to predict and not merely describe the consequences of action” [Friedman, 1953, 11–12].

Moreover, Friedman believed that many apparent disputes over values actually revolve around issues of fact and hence could be resolved empirically—that is, through the methods of positive economics. As an example, he cited disputes over the desirability of minimum wage legislation that seemingly revolved around values but, which he posited, hinged on, and hence could be resolved by knowledge of, the impact of an increase in the minimum wage on employment. While not denying the significance and utility of normative economics (which he hardly could have done without risking the label of hypocrite), Friedman believed it would be possible to diminish the scope of normative economics by expanding that of positive economics. Casual empiricism, as well as empirical research, however, suggests that issues of value are rarely resolved by recourse to data [see, e.g., Mitroff, 1974]. Empirical testing usually has a sufficient subjective element such that clever investigators can, by framing the question in the right way, or by using the right methods, come up with the answer they seek (see Glaeser, chapter 13). Thus, for example, Plott and Zeiler [2005] show that with a magical mixture of experimental manipulations, they are able to reduce the magnitude of the endowment effect. Indeed, even on the issue that Friedman used to illustrate the capacity of positive economics to supplant normative economics—the impact of an increase in the minimum wage on employment—there has been a remarkable tendency for empirical research conducted by proponents of raising the minimum wage to conclude that doing so has minimal or even positive impact on employment, with the opposite pattern observed in the research of opponents. Fuchs, Krueger, and Poterba [1998] conducted a survey of

labor and public economists at leading research universities that elicited, among other things, respondents' beliefs about the impact of an increase in the minimum wage on youth employment, their degree of support for an increase in the minimum wage, and various questions about values and political orientation. Despite many decades of research on the topic, they found a remarkable lack of convergence among researchers regarding the impact of a minimum wage hike on employment. Moreover, there was also little evidence that settling the positive issue would, in fact, help to resolve the normative one. Support for an increase in the minimum wage was strongly correlated with a researcher's social and political values but barely related to economists' beliefs about the impact of an increase in the minimum wage on employment.

## REFERENCES

---

- Angrist, Joshua, Daniel Lang, and Philip Oreopoulos. 2006. Incentives and Services for College Achievement: Evidence from a Randomized Trial. Online. Available: [www.stanford.edu/group/SITE/Web%20Session%207/Oreopoulos\\_Abstract.pdf#search=%22](http://www.stanford.edu/group/SITE/Web%20Session%207/Oreopoulos_Abstract.pdf#search=%22).
- Ashraf, Nava, Dean Karlan, and Wesley Yin. 2005. Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines. *Quarterly Journal of Economics* 121: 635–672.
- . 2006. Deposit Collectors. *Advances in Economic Analysis and Policy* 6(2): 1–22.
- Badger, Gary J., Warren K. Bickel, Louis A. Giordano, Eric A. Jacobs, and George Loewenstein, 2007. Altered States: The Impact of Immediate Craving on the Valuation of Current and Future Options. *Journal of Health Economics* 26: 865–876.
- Camerer, Colin, Samuel Issacharoff, George Loewenstein, Ted O’Donoghue, and Matthew Rabin. 2003. Regulation for Conservatives: Behavioral Economics and the Case for “Asymmetric Paternalism.” *University of Pennsylvania Law Review* 151(3): 1211–1254.
- Campbell, Donald. 1969. Reforms as Experiments. *American Psychologist* 24: 409–429.
- Choi, James J., David Laibson, and Brigitte C. Madrian. 2004. Plan Design and 401(k) Savings Outcomes. *National Tax Journal* 57: 275–298.
- . 2005. \$100 Bills on the Sidewalk: Suboptimal Savings in 401(k) Plans. National Bureau of Economic Research NBER Working Paper 11554.
- Choi, James J., David Laibson, Brigitte C. Madrian, and Andrew Metrick. 2005. Optimal Defaults and Active Decisions. National Bureau of Economic Research NBER Working Paper: 11074.
- Clotfelter, Charles T., and Philip J. Cook. 1989. *Selling Hope: State Lotteries in America*. Cambridge, MA: Harvard University Press.
- DellaVigna, Stefano. Forthcoming. Psychology and Economics: Evidence from the Field. *Journal of Economic Literature*.
- Diener, E., and M. Seligman 2004. Beyond Money: Toward an Economy of Well-Being. *Psychological Science in the Public Interest* 5: 1–31.
- Di Tella, Raphael, Robert J. MacCulloch, and Andrew J. Oswald. 2003. The Macroeconomics of Happiness. *Review of Economics and Statistics* 85: 809–827.
- Duflo, Esther, and Emmanuel Saez. 2002. Participation and Investment Decisions in a Retirement Plan: The Influence of Colleagues’ Choices. *Journal of Public Economics* 85: 121–48.
- . 2003. The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment. *Quarterly Journal of Economics* 118: 815–842.
- Flegal, Katherine M., Barry I. Graubard, David F. Williamson, and Mitchell H. Gail. 2007. Cause-specific Excess Deaths Associated with Underweight, Overweight, and Obesity. *JAMA: Journal of the American Medical Association* 298: 2028–2037.
- Frederick, Shane, and George Loewenstein. 1999. Hedonic Adaptation. In *Well-Being: The Foundations of Hedonic Psychology*, ed. Daniel Kahneman, Ed Diener, and Norbert Schwarz, 302–329. New York: Russell Sage Foundation.
- . 2006. Preference Authority in the Evaluation of Sequences. Center for Behavioral Decision Research Working Paper, Carnegie Mellon University.
- Friedman, Milton. 1953. “The Methodology of Positive Economics,” In *Essays in Positive Economics*. Chicago: University of Chicago Press.
- Fuchs, Victor R., Alan B. Krueger, and James M. Poterba. 1998. Economists’ Views about

- Parameters, Values, and Policies: Survey Results in Labor and Public Economics. *Journal of Economic Literature* 36: 1387–1425.
- Glaeser, Edward L. 2006. Paternalism and Psychology. *University of Chicago Law Review* 73: 133–156.
- Halpern, Scott D., Peter A. Ubel, and David A. Asch. 2007. Harnessing the Power of Default Options to Improve HealthCare. *New England Journal of Medicine* 357: 1340–1344.
- Heil, Sarah H., Jennifer W. Tidey, Heather W. Holmes, and Stephen T. Higgins. 2003. A Contingent Payment Model of Smoking Cessation: Effects of Abstinence and Withdrawal. *Nicotine and Tobacco Research* 5: 205–213.
- Herrnstein, Richard, George Loewenstein, Drazen Prelec, and William Vaughan. 1993. Utility Maximization and Melioration: Internalities in Individual Choice. *Journal of Behavioral Decision Making* 6: 149–185.
- Higgins, Stephen T., Conrad J. Wong, Gary J. Badger, Doris E. Ogden, and Robert L. Dantona. 2000. Contingent Reinforcement Increases Cocaine Abstinence During Outpatient Treatment and One Year of Follow-up. *Journal of Consulting and Clinical Psychology* 68: 64–72.
- Hull, Clark L. 1932. The Goal-Gradient Hypothesis and Maze Learning. *Psychological Review* 39: 24–43.
- Issacharoff, Samuel, and Erin F. Delaney. 2006. Credit Card Accountability. *University of Chicago Law Review* 73: 157–182.
- Jackevicius, Cynthia A., Muhammad Mamdani, and Jack V. Tu. 2002. Adherence with statin therapy in elderly patients with and without acute coronary syndromes. *JAMA : the Journal of the American Medical Association*. 288: 462-467.
- Jeffrey, Robert W., Wendy M. Gerber, Barbara S. Rosenthal, and Ruth A. Lindquist. 1983. Monetary Contracts in Weight Control: Effectiveness of Group and Individual Contracts of Varying Size. *Journal of Consulting and Clinical Psychology* 51: 242–248.
- Jeffrey, Robert W., Paul D. Thompson, and Rena R. Wing. 1978. Effects on Weight Reduction of Strong Monetary Contracts for Calorie Restriction or Weight Loss. *Behaviour Research and Therapy* 16: 363–369.
- Kahneman, Daniel, and Alan B. Krueger. 2006. Developments in the Measurement of Subjective Well-Being. *Journal of Economic Perspectives* 20: 3–24.
- Kahneman, Daniel, and Dale Miller. 1986. Norm Theory: Comparing Reality to Its Alternatives. *Psychological Review* 93: 136–153.
- Kane, Robert L., Paul E. Johnson, Robert J. Town, and Mary Butler. 2004. A Structured Review of the Effect of Economic Incentives on Consumers' Preventive Behavior. *American Journal of Preventive Medicine* 27: 327–352.
- Katz, Lawrence F., Jeffrey R. Kling, and Jeffrey B. Liebman. 2001. Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment. *Quarterly Journal of Economics* 116: 607–654.
- Keynes, John N. 1891. *The Scope and Method of Political Economy*. London: Macmillan and Co.
- Kivetz, Ran, Oleg Urminsky, and Yuhuang Zheng. 2006. The Goal-Gradient Hypothesis Resurrected: Purchase Acceleration, Illusionary Goal Progress, and Customer Retention. *Journal of Marketing Research* 43: 39–58.
- Klick, Jonathan, and Gregory Mitchell. 2006. Government Regulation of Irrationality: Moral and Cognitive Hazards. *Minnesota Law Review* 90: 1620–1663.
- Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. Experimental Analysis of Neighborhood Effects. *Econometrica* 75: 83–119.
- Krugman, Paul. 2007. Who Was Milton Friedman? *New York Review of Books* 54: 27–30.



- Layard, R. 2005. *Happiness. Lessons from a New Science*. London: Allen Lane.
- LeBoeuf, Robyn A., and Eldar B. Shafir. 2005. Decision Making. In *The Cambridge Handbook of Thinking and Reasoning*, ed. Keith J. Holyoak and Robert G. Morrison, 243–265. New York: Cambridge University Press.
- Levitt, Steven D., and John A. List. 2008. What Do Laboratory Experiments Measuring Social Preferences Reveal about the Real World? *Journal of Economic Perspectives*.
- Loewenstein, George. 1996. Out of Control: Visceral Influences on Behavior. *Organizational Behavior and Human Decision Processes* 65: 272–292.
- Loewenstein, George, Troyen Brennan, and Kevin G. Volpp. 2007. Asymmetric Paternalism to Improve Health Behaviors. *JAMA: Journal of the American Medical Association* 298: 2415–2417.
- Loewenstein, George, and Ted O’Donoghue. 2006. We Can Do This the Easy Way or the Hard Way: Negative Emotions, Self-Regulation, and the Law. *University of Chicago Law Review* 73: 183–206.
- Loewenstein, George, and Peter A. Ubel. Forthcoming. Hedonic Adaptation and the Role of Decision and Experience Utility in Public Policy. *Journal of Public Economics*, Special Issue on Happiness and Public Economics.
- Madrian, Brigitte C., and Dennis F. Shea. 2001. The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior. *Quarterly Journal of Economics* 116: 1149–1525.
- Mann, Ronald A. 1972. The Behavior-Therapeutic Use of Contingency Contracting to Control an Adult Behavior Problem: Weight Control. *Journal of Applied Behavioral Analysis* 5: 99–102.
- McClure, Samuel M., David I. Laibson, George Loewenstein, and Jonathan D. Cohen. 2004. Separate Neural Systems Value Immediate and Delayed Monetary Rewards. *Science* 306: 503–507.
- Mitroff, Ian I. 1974. *The Subjective Side of Science: A Philosophical Inquiry into the Psychology of the Apollo Moon Scientists*. New York: Elsevier.
- Murphy, Kevin. 2006. Keynote Address. 2006. Healthchallenge Think Tank, McGill University.
- Nordgren, Loran F., Joop van der Pligt, and Frenk van Harreveld. In press. The Instability of Health Cognitions: Visceral States Influence Self-Efficacy and Related Health Beliefs. *Health Psychology*.
- Nussbaum, Martha. 2000. *Women and Human Development: The Capabilities Approach*. Cambridge: Cambridge University Press.
- Plott, Charles R., and Kathryn Zeiler. 2005. The Willingness to Pay–Willingness to Accept Gap, the “Endowment Effect,” Subject Misconceptions, and Experimental Procedures for Eliciting Valuations. *American Economic Review* 95: 530–45.
- Prelec, Drazen, and George Loewenstein. 1991. Decision Making over Time and under Uncertainty: A Common Approach. *Management Science* 37: 770–786.
- Riis Jason, George Loewenstein, Jonathan Baron, Christopher Jepson, Angela Fagerlin, and Peter A. Ubel. 2005. Ignorance of Hedonic Adaptation to Hemo-dialysis: A Study Using Ecological Momentary Assessment. *Journal Experimental Psychology: General* 134: 3–9.
- Sanfey, Alan, George Loewenstein, Sam M. McClure, and Jonathan D. Cohen. 2006. Neuroeconomics: Cross-Currents in Research on Decision Making. *Trends in Cognitive Science* 10: 108–116.
- Schroeder, Steven A. 2007. Shattuck Lecture. We Can Do Better–Improving the Health of the American People. *The New England Journal of Medicine* 357 (12): 1221–1228.
- Schwartz, Janet, Marianne Bertrand, Sendhil Mullainathan, and Eldar Shafir. 2006. Boosting Program Take-up: An Experiment with Flexible Spending Accounts. Paper

presented at the Behavioral Decision Research in Management Conference, Santa Monica, June 15–17.

- Sen, Amartya. 1999. *Development as Freedom*. Oxford: Oxford University Press.
- . 1985. *Commodities and Capabilities*. Amsterdam: North-Holland.
- Skiba, Paige, and Jeremy Tobacman. 2006. Measuring the Effect of Access to Credit: Evidence from Payday Loans. Paper presented at the Society for Judgment and Decision Making Conference, Houston, November 17–20.
- Smith, Dylan M., Ryan L. Sherriff, Laura J. Damschroder, George Loewenstein, and Peter A. Ubel. 2007. Misremembering Colostomies? Former Patients Give Lower Utility Ratings Than Do Current Patients. *Health Psychology* 25: 688–695.
- Strotz, Robert H. 1955. Myopia and Inconsistency in Dynamic Utility Maximization. *Review of Economic Studies* 23: 165–180.
- Sugden, Robert. 2005. Capability, Happiness and Opportunity. Paper presented at Capabilities and Happiness: An International Conference, Department of Economics, University of Milano-Bicocca, June 16–18.
- Thaler, Richard H., and Shlomo Benartzi. 2004. Save More Tomorrow: Using Behavioral Economics to Increase Employee Savings. *Journal of Political Economy* 112: 164–187.
- Thaler, Richard H., and Cass R. Sunstein. 2003. Libertarian Paternalism. *American Economic Review* 93: 175–179.
- Ubel, Peter, and George Loewenstein. 1996. Distributing Scarce Livers: The Moral Reasoning of the General Public. *Social Science and Medicine* 42: 1049–1055.
- Ubel, Peter, George Loewenstein, Dennis Scanlon, and Mark Kamlet. 1996. Individual Utilities Are Inconsistent with Rationing Choices: A Partial Explanation of Why Oregon's Cost-Effectiveness List Failed. *Medical Decision Making* 16:108–116.
- Ubel, Peter A., George Loewenstein, Norbert Schwarz, and Dylan Smith. 2005. Misimagining the Unimaginable: The Disability Paradox and Healthcare Decision Making. *Health Psychology* 24: 57–62.
- Van Praag, Bernard M. S., and Barbara E. Baarsma. 2005. Using Happiness Surveys to Value Intangibles: The Case of Airport Noise. *Economic Journal* 115: 224–246.
- Volpp, Kevin G., Andrea Gurmankin., David A. Asch, Jesse A. Berlin, John J. Murphy, Angela Gomez, Harold Sox, Zhu Jungsan, and Caryn Lerman. 2006. A Randomized Control Trial of Financial Incentives for Smoking Cessation. *Cancer Epidemiology Biomarkers and Prevention* 15: 12–18.
- Weber, Bethany J., and Gretchen B. Chapman. 2005. Playing for Peanuts: Why Is Risk Seeking More Common for Low-Stakes Gambles? *Organizational Behavior and Human Decision Processes* 97: 31–46.