Commitment Devices

Gharad Bryan, Dean Karlan, and Scott Nelson

Department of Economics, Yale University, New Haven, Connecticut 06520; email: dean.karlan@yale.edu

Annu. Rev. Econ. 2010. 2:671–98

First published online as a Review in Advance on June 4, 2010

The *Annual Review of Economics* is online at econ.annualreviews.org

This article's doi: 10.1146/annurev.economics.102308.124324

1941-1383/10/0904-0671$20.00

Key Words

commitment contracts, present-biased preferences, self-control, temptation models, hyperbolic preferences

Abstract

We review the recent evidence on commitment devices and discuss how this evidence relates to theoretical questions about the demand for, and effectiveness of, commitment. Several important distinctions emerge. First, we distinguish between what we call hard and soft commitments and identify how soft commitments, in particular, can help with various dilemmas, both in explaining empirical behavior and in designing effective commitment devices. Second, we highlight the importance of certain modeling assumptions in predicting when commitment devices will be demanded and examine the laboratory and field evidence on the demand for commitment devices. Third, we present the evidence on both informal and formal commitment devices, and we conclude with a discussion of policy implications, including sin taxes, consumer protection, and commitment device design.

1. INTRODUCTION

People set goals. They promise to smoke less, to save more, to drink less, to work more, to eat less, and to exercise more. They set goals with varying levels of formality, making promises to themselves, vows to spouses, or bets with friends. These same people regularly fail to meet their goals. They keep smoking, spend money on things they later regret, and throw out perished vegetables while eating chips. Why do people set goals, and why do they fail to reach them? These are important and difficult questions for the economics profession; they draw into doubt our model of human behavior and problematize our definition of welfare.

This review is concerned with self-commitment devices—the arrangements people make to formalize and facilitate their goals. Broadly, a commitment device is an arrangement entered into by an individual with the aim of helping fulfill a plan for future behavior that would otherwise be difficult owing to intrapersonal conflict stemming from, for example, a lack of self-control. We exclude actions that accrue significant current benefits or that are taken with a strategic motive.[1] Therefore, whereas painting one's nails with Control-It Nail Biting Treatment® is a commitment device, buying nail polish that looks beautiful but that also happens to taste foul, with the goal of having attractive nails, is not. And whereas signing up for automatic savings transfers from a checking account may be a commitment device, doing so to limit a spouse's spending is not.

We refer to commitment devices that call for real economic penalties for failure, or rewards for success, hard commitments. And we refer to any device that has primarily psychological consequences a soft commitment. This determination is not perfectly binary, as some hard commitments also have psychological costs, and most soft commitments will also accrue some nonzero economic cost. An example of a hard commitment would be a commitment savings account on which interest is forfeited if a monthly deposit is not made; note that the depositor could also incur a psychological cost, such as shame or loss of self-esteem, for missing a deposit. A soft commitment would be a separate savings account labeled "send kids to college"; if someone withdraws money from that account to pay for a holiday party, he or she incurs costs that are primarily psychological, such as disappointment, but perhaps also faces a small economic cost, such as the opportunity cost of time spent at the bank.

The paper is structured as follows. In Section 2 we provide a more detailed definition of commitment, as well as background discussion of the variety of domains and informal mechanisms observed. In Section 3 we briefly review economic theories that can explain a demand for commitment, highlighting the multiplicity of explanations and some issues related to sophistication and naïveté, and soft versus hard commitments. We also briefly discuss how a lack of commitment for the poor can lead to poverty traps. In Section 4 we examine evidence of the use of commitment in both laboratory and field settings, and in Section 5 we discuss evidence showing a demand for commitment. In Section 6 we discuss issues for the market for commitment devices. Finally Section 7 discusses welfare and concludes with a discussion of areas for further research.

[1]Naturally, the phrase commitment device is also appropriately used to describe strategic arrangement with respect to others. However, we limit the scope of this review to self-commitment devices, as defined above.

2. DEFINITION AND MOTIVATION

In planning this review, we intended to include a concise one-line definition of a commitment device in the introduction. Writing such a definition proved to be a difficult but enlightening task. Instead of a concise definition, we therefore include a long definition and several examples, which we hope clearly illustrate what we mean by a commitment device, and show that defining a commitment device ultimately requires some appeal to an agent's motivation.

2.1. Definition

We define a commitment device as an arrangement entered into by an agent who restricts his or her future choice set by making certain choices more expensive, perhaps infinitely expensive, while also satisfying two conditions: (*a*) The agent would, on the margin, pay something in the present to make those choices more expensive, even if he or she received no other benefit for the payment, and (*b*) the arrangement does not have a strategic purpose with respect to others.

Condition (*a*) excludes many common economic transactions that are not germane to our discussion, such as simple pay-in-advance purchases. Buying concert tickets in advance and booking plane tickets, for example, are behaviors that could be seen as commitments because they shrink the future budget set, lock in certain consumption goods (as long as there is not free disposal), and alter the price of future consumption. But, in most cases, it is inappropriate to see these as commitment devices, unless condition (*a*) is satisfied—that is, unless the agent values the commitment aspect of the transaction enough to pay something simply to gain the commitment without any other benefit.

Condition (*a*) also allows us to distinguish between arrangements that facilitate future consumption plans merely for transactional or logistical reasons and arrangements that are desirable because they lock in (commit) a particular consumption path. For example, buying in bulk could be done to commit oneself to consume more of a particular good, but it also could be done merely to lower transaction costs. We are not interested in the latter, just the former.

Condition (*b*) rules out a well-studied set of commitment devices that are used to influence the actions of others. Consider, for example, the doomsday device in the movie *Dr. Strangelove*, which committed Russia to set off an all-destructive nuclear device in response to any nuclear attack. The committing party did not want to commit themselves to set off the all-destructive device, but rather wanted to deter attacks from others. Of particular interest, this condition means that we do not study arrangements that are entered into with the aim of controlling one's spouse or neighbors. We note, however, that our exclusion does not mean that we believe these motivations to be unimportant. In developing countries there is some evidence that commitment devices arise because of spousal, familial, and neighbor bargaining (e.g., see Anderson & Baland 2002, Plateau 2001).

At times we discuss commitment contracts as compared to devices. A commitment contract refers to a commitment device that is an actual contract between two parties, rather than a unilateral arrangement employed by an individual to restrict his or her own choices.[2]

[2]Commitment contracts have generated legal discussions because, if they are designed to benefit one party, the enforcing party must also receive consideration (i.e., some benefit) for the contracts to be legally binding. This requirement is typically easy to satisfy through the so-called peppercorn principle of consideration, i.e., that even if a party receives value that is no more than a peppercorn, this is sufficient consideration to be a legally binding contract.

2.2. A Formal Example: The Runner's Dilemma

Our definition requires an understanding of the motivation behind people's choices. In this section we provide a specific example of a commitment contract and the thought process that motivates it. Suppose we observe a runner (Rachel) about to embark on a 10-mile training session on which she can either run or walk. Before departing, Rachel signs a contract with a friend that says that if she walks, she will pay $1,000 to the friend.

What motivates this decision, and why is it a commitment contract? Assume that the true benefit of running is B and the true cost is C, and think of Rachel as two selves, the time 0, or prerunning, Rachel and the time 1, or running, Rachel. We think of these Rachels as having beliefs regarding the costs and benefits of running, where these beliefs may differ over time. B_t and C_t represent Rachel's belief about B and C at time t.

Furthermore, time 0 Rachel must have a belief about what time 1 Rachel will believe. We define B_1^0 as Rachel's belief at time 0 as to what her belief will be in time period 1, and C_1^0 is similarly defined. Throughout we assume that $B_1 < C_1$ so that Rachel will actually walk if no commitment contract is signed.

Within this context, a decision to commit to running can be rationalized by the belief of time 0 Rachel that $B_0 > C_0$ but that $B_1^0 < C_1^0$. That is, time 0 Rachel believes that it is valuable to run but believes that time 1 Rachel does not. Any choice to restrict future consumption that is motivated by these beliefs fits our definition of a commitment contract.

2.3. Some Examples of Commitment

The motivation for studying commitment devices is simple: Evidence persists in many domains of life that individuals create ad hoc arrangements that can be construed as commitment devices. Understanding these arrangements better, and understanding whether they are suboptimal, helps to understand the potential for more formal commitment contracts, as well as highlights some of the challenges in the development of a market for formal commitment contracts. Later in the review we present harder and more careful evidence on commitment devices, but we begin with a set of anecdotes.

2.3.1. Smoking. The majority of smokers say they want to quit, often try to quit, but continuously fail in their quit attempts. Gallup (2008) polling suggests 74% of smokers would like to give up smoking, and the CDC reports that 70% want to quit completely. Quit attempts are also extensive. In 2006, an estimated 44.2% of adult smokers stopped smoking for at least one day as part of an attempt to quit (CDC 2006). However, success rates are low. Data variously suggest that fewer than 5% (Am. Lung Assoc. 2008, p. 8), or 4%–7% (Hughes et al. 2004), of quit attempts annually are successful.

2.3.2. New Year's Resolutions. New Year's resolutions are (infamously) unsuccessful. Although the evidence is light, one study followed 3,000 resolution makers from 2007. Whereas 52% of resolution makers were confident of success, only 12% actually managed to follow through with their plans (Quirkology 2007). Evidence suggests some of these failures are cyclical: Roughly 60% of resolutions made in 2009 were repeats of failed 2008 resolutions (Resolut. Res. 2009).

2.3.3. Weight Loss. In Gallup polling from 2006, 56% of Americans said they wanted to lose weight (Gallup 2006). More tellingly, 59% of those interviewed in 2001 said they

wanted to lose weight, implying that at least 15% of Americans were still trying five years later (Gallup 2006).

2.3.4. Various ad hoc Commitment Devices. Other limited anecdotal evidence suggests ad hoc behavior that can be construed as commitment devices, including the following (many of these are cited in Schelling 1984): cutting up one's credit cards, only taking a fixed amount of cash when heading out to party for a night, buying junk food in small packages rather than buying in bulk, not keeping alcohol in the house, brushing one's teeth earlier in the evening to avoid late night snacking (Hamermesh 2010), drinking a lot of water before going to bed to make oneself get out of bed upon waking in the morning,[3] going to a hotel room for the day or taking a longer than necessary train ride to get writing done, leaving one's laptop or papers at the office so that time at home is not spent working, drinking through a straw to avoid gulping, and buying long-term gym memberships rather than paying by the day.[4]

3. THEORY

In this section we discuss three models that are most often used to rationalize a demand for commitment. We discuss the quasi-hyperbolic discounting model of Laibson (1997), highlighting how a lack of sophistication can explain failures of commitment; the temptation and self-control theory of Gul & Pesendorfer (2001); and the set of dual-self theories studied by Thaler & Shefrin (1981) and Fudenberg & Levine (2006). We note, however, that there are many models that imply a demand for commitment and that the welfare and market implications of commitment may differ depending on the model used.[5] We also discuss extensions of these theories that deal specifically with the commitment problems of the poor, and with soft commitments.

In all the models that follow below, we discuss a single decision maker as if he or she were a collection of decision makers over many time periods $t = \{0,1,\ldots\}$ and refer to the time t decision maker as the time t self.

3.1. Hyperbolic Discounting

The demand for commitment was first modeled by Strotz (1956). Strotz noted that a slight generalization of the discounted utility model implied that an individual's intertemporal marginal rate of substitution would differ depending on the time period in which a decision was made and that this would lead to a preference for commitment. Specifically, an

[3]We owe this example to Bart Simpson in *The Simpsons*' third Christmas episode "Miracle on Evergreen Terrace." Lisa Simpson claims that Native Americans also used this commitment device—we cannot confirm this conjecture.

[4]This phenomenon has been studied by DellaVigna & Malmendier (2006), who find that, for many people, paying daily would be cheaper than long-term memberships. This is interpreted as evidence of unsuccessful attempts at committing oneself to exercise.

[5]Preferences that explain a demand for commitment include the following: (*a*) preferences that change over time such as the quasi-hyperbolic model discussed below or the model of Banerjee & Mullainathan (2009); (*b*) preferences that depend on the set of choices, such as Gul & Pesendorfer (2001), Dekel et al. (2009), Sarver (2008), and Noor (2007); (*c*) preferences that breach the assumption of independence of irrelevant alternatives [for example, the optimal expectations model of Brunnermeier & Parker (2005) breaches the irrelevance of independent alternatives, as shown by Spiegler (2008), and can generate a demand for commitment]; and (*d*) preferences that are determined in part by past decisions, such as Koszegi & Rabin (2009).

individual with preferences that are time separable and additive assesses the utility from consumption over time $\{c_t, c_{t+1}, \ldots, c_T\}$ according to

$$U(\{c_t,c_{t+1},\ldots,c_T\}) = \sum_{k=0}^{T-t} D(k)u(c_{t+k},t+k),$$

where u is a one-period utility function and $D(k)$ any discount function.

In this framework, preferences are time consistent if the decision maker deciding at time t would agree with a decision maker at time $t + \tau$ on the relative importance of consumption across any two periods, for all τ.[6] Strotz showed that this is only the case if $D(k) = \delta^k$ (i.e., the time-consistent exponential discounting model); for any other function, the preferences are time inconsistent. Time inconsistency implies that different selves differ in their assessment of the best course of action and consequently that each time period's decision maker would like to restrict the set of choices available to his or her future selves.

Strotz (1956) argued that a function that accurately represented human behavior would overvalue current consumption. A simple formulation that allows for such preferences was provided by Phelps & Pollak (1968) and Laibson (1997):

$$U(\{c_t,c_{t+1},\ldots,c_T\}) = u(c_t) + \beta\sum_{k=1}^{T-t} \delta^k u(c_{t+k}),$$

where $\beta \leq 1$. Laibson argued that this model captures the essence of hyperbolic discounting favored by psychological evidence (for a review, see Ainslie 1992) but retains the tractability of the discounted utility model. This model has come to be called the quasi-hyperbolic discounting model.

3.2. Sophistication and Naïveté

The quasi-hyperbolic model leaves open the question of whether the decision maker knows his or her future preferences. O'Donoghue & Rabin formalize the distinction and show that it can have an important impact on behavior (e.g., see O'Donoghue & Rabin 1999).[7] They distinguish between two extreme types: sophisticates, who know their own β, and naifs, who do not. In between these two types are agents who are partially naïve: They know that they are prone to overconsumption but underestimate the extent of their weakness.

For our purposes there are two major implications of naïveté. First, the completely naïve need commitments but do not recognize the fact. They are, therefore, difficult to help through market-provided commitment contracts. Second, partially naïve agents may undercommit their future selves, making commitments on which they do not follow through. It is therefore possible to exploit partially naïve agents by charging them for the chance to undercommit (for more in-depth discussion of this and related issues, see DellaVigna & Malmendier 2004, Eliaz & Spiegler 2006, Gottlieb 2008, Heidhues & Koszegi 2008). We return to this general theme when we discuss market provision of

[6]Formally we require that the marginal rate of substitution $\frac{\partial U}{\partial c_t} / \frac{\partial U}{\partial c_{t+\tau}}$ is the same regardless of the date at which it is assessed.

[7]A second formulation of naïveté is due to Eliaz & Spiegler (2006). The two formulations often agree but may lead to different welfare assessments (e.g., see Heidhues & Koszegi 2008).

commitment, as this self-awareness issue is critical to the question of the take-up of commitment contracts.

3.3. Choice-Set-Dependent Utility

The main alternative to the quasi-hyperbolic model is the temptation preference model of Gul & Pesendorfer (2001, 2004) (GP), which considers preferences over sets of choices. In the GP model there is a cost of avoiding the most tempting item in a choice set. Agents, therefore, benefit from removing tempting items from their choice sets, implying a demand commitment. For example, the GP model applies to an agent's choice over restaurant menus. Consider the choice between two menus, a menu with just fish $\{f\}$ and a menu with fish and steak $\{f,s\}$. When facing the menu $\{f\}$, the agent receives the ordinary utility associated with eating fish. However, when facing the menu $\{f,s\}$, the agent may end up eating fish, but will pay a psychological cost of avoiding the more tempting steak option. The agent seeks commitment to avoid this temptation cost.

More formally, when choosing from a menu consisting of only one item, the decision maker's preferences are given by an ordinary utility function u. So the utility associated with the menu $\{f\}$ is $u\{f\}$. Because a singleton menu is essentially a commitment, the preferences represented by u are often termed commitment preferences, and in the spirit of our example we assume that $u\{f\} > u(s)$. The decision maker also finds some items tempting, with the amount of temptation measured by a second utility function v. For most people it is safe to assume that steak is more tempting than fish, so let us assume that $v(s) > v(f)$. When choosing from the menu $\{f,s\}$, the decision maker receives utility

$$U(\{f,s\}) = \max_{b\in\{f,s\}} (u(b) + v(b)) - \max_{c\in\{f,s\}} v(c). \quad (1)$$

Given our assumptions about v, we know that the last term in Equation 1 is equal to $v(s)$; that is, the decision maker faces a cost equal to the temptation value of the most tempting item on the menu, the steak. The first term in Equation 1 then says that the decision maker weighs up the commitment and temptation utility in making his or her decision. If $u(f) + v(f) > u(s) + v(s)$, then he or she will choose the fish and utility will be given by $u(f) - (v(s) - v(f))$. That is, his or her utility is the commitment value of the fish minus the difference in temptation values between the chosen item and the most tempting item. Because the last term must be negative, it is always the case that $u(f) \geq u(\{f,s\})$, and so the decision maker would prefer to choose from the smaller menu $\{f\}$. If, alternatively, the agent gives in to temptation and orders the steak, his or her utility is given by $u(s)$, and again we see that $u(f) > u(s)$, so he or she would prefer to commit to the smaller menu.

Two key characteristics distinguish the GP model from the quasi-hyperbolic model. First, the GP model is consistent with revealed preference. So, a time 1 self that chooses the steak is taking the preference-maximizing action given the temptation cost associated with the menu $\{f,s\}$. Second, disutility in the GP model is a function of the choice set, and thus there is a possibility for costly self-control. In contrast, utility in the hyperbolic model is strictly a function of the actions taken. For example, in the hyperbolic model, if fish is chosen in period 1, then there is no cost associated with steak being on the menu. In the GP model, however, the existence of the steak may be a cost. Formally, the GP model can rationalize $\{fish\} \succ_0 \{steak, fish\}$ and $fish \succ_1 steak$ (where $\succ_t$ are period t preferences), whereas the hyperbolic model cannot.

3.4. Dual-Self Models

The GP representation is also open to the following interpretation: The decision maker consists of a long-run, planning, self and a series of short-run, doing, selves. The long-run self has preferences given by u, whereas the short-run selves have preferences given by v. The long-run self can exert influence on the short-run self to change choices but must pay a mental cost that is equal to $v(a) - \max_{c \in B} v(c)$ to implement choice a. The GP model can therefore capture the possibility that the brain is made up of two decision-making entities or selves.[8] This observation provides a link between the GP model and a long tradition in psychology and a smaller selection of formal dual-self models in economics.

Dual-self models differ in their structure, but all posit a long-run self and a short-run self that are in some kind of conflict. Formal models currently in the literature differ in two key ways. First, the source of conflict differs: For example, the short-run selves may be myopic (Thaler & Shefrin 1981, Fudenberg & Levine 2006), face a distorted utility function (Benhabib & Bisin 2005), or be addicted (Bernheim & Rangel 2004). Second, the means of control may differ. For example, the long-run self may face a cost of manipulating the short-run preferences (Thaler & Shefrin 1981, Fudenberg & Levine 2006), may have a limited ability to keep decisions consistent with a preset plan (Benhabib & Bisin 2005), or may face a stochastic process that determines whether the long-run or short-run self has complete control (Bernheim & Rangel 2004).

Regardless of the specifics, these models all predict that the long-run decision maker would benefit from commitment contracts that curtail the set of actions available to its future short-run selves. These models have the advantages of directly incorporating psychological processes and providing clear welfare rules based on the preferences of the long-run self.

3.5. Income, Poverty, and Commitment

Much of the empirical work discussed here centers on the poor. This reflects a simple reality: Behavioral anomalies may be costly to individuals, but the poor have less slack, i.e., disposable income, with which to absorb errors (Mullainathan & Shafir 2009). Banerjee & Mullainathan (2009) formalize this idea in the context of commitment, putting forward a temptation model that helps explain the existence of poverty traps. Their model allows for two types of goods, temptation goods (z) and nontemptation goods (x). They make one critical assumption: The portion of income spent on temptation goods decreases as income grows. As an example, inexpensive items such as candy are consumed by rich and poor alike and are probably tempting to most. More expensive items, however, such as cars and foreign vacations, are consumed exclusively by the wealthy and seem less likely to be tempting. It seems intuitively appealing, therefore, that the portion of income spent on temptation is lower for the wealthy.

Their model is a variation of the hyperbolic discounting model presented above. In a two-period version, the time 1 self makes decisions using the utility function

$$u(x_1) + v(z_1) + \delta u(x_2),$$

where δ is the discount rate. The time 2 agent, however, decides according to the utility function $u(x_2) + v(z_2)$. Thus, the time 2 agent values z_2, whereas the time 1 agent does not.

[8]Fudenberg & Levine (2006) and Benabou & Pycia (2002) formalize this claim. See also Loewenstein & O'Donoghue (2004) for a discussion of the relationship between the quasi-hyperbolic model and dual-self models.

In this model, z_2 is an increasing function of the resources available in period 2, and the fact that the time 1 agent does not value z_2 implies that there is, in effect, a temptation tax on period 1 savings. Banerjee & Mullainathan (2009) show that if preferences are such that the portion of income spent on temptation goods is decreasing in income (z_2 is concave in period 2 resources), then the model can generate a poverty trap. Intuitively, the poor do not save for the future, as they know a relatively large amount of that saving will be spent on z goods—a high marginal temptation tax. At higher levels of income, however, only a small portion of marginal saving will be spent on z goods. Therefore, those with low income do not save, and those with high income do save, leading to polarization of income. If this model is correct and temptation does in fact represent a smaller cost to the rich, then commitment devices may provide a means of pulling the poor out of poverty.

3.6. Soft Commitments

Although there has been much discussion of concepts such as mental accounting, which would allow for soft commitment (e.g., Thaler 1985), there is a dearth of papers that formalize these intuitions and their welfare and behavioral implications. Because soft commitments work by imposing psychological costs, it seems natural that models of soft commitment will need to incorporate a better understanding of the actual decision process itself. We discuss here two possible approaches, but note that each has its difficulties.

A promising approach can be found in Benabou & Tirole (2004). These authors develop a model of an agent who sets unenforced rules for him or herself, such as jogging regimens or reduced cigarette use. The central feature of their model is that the agent faces cravings c and exercises willpower β in the midst of a willpower activity (such as jogging or quitting smoking), but both c and β are random variables that cannot be accurately remembered ex post and cannot be predicted ex ante. Thus the agent creates personal rules and invests in his or her own track record of following these rules. The agent wishes to protect this track record during times of great temptation and hence follows through on goals when in the absence of such a rule he or she might not. In turn, this track record gives the agent the willingness to begin new willpower activities in the future, which the agent desires to do. Although this model explains personal rules, it does little to explain why formalization (e.g., labeling an account "the children's college fund") can be an effective commitment device.

Alternatively, the cue-based theory of Bernheim & Rangel (2004) might explain the success of some kinds of soft commitment. In their model, certain environmental factors are complementary to specific temptation goods: smoking with friends, overspending at a mall, or overeating at a ballpark. It may, therefore, be more effective to avoid those particular friends, or the mall, or the ballpark, rather than trying to avoid cigarettes, cash, or food altogether. A similar model could help explain the effectiveness of soft commitments. By making plans that avoid the complementary good, an agent can avoid consuming the temptation good, but this is clearly a soft commitment, in that it is likely to succumb easily to some temptation, albeit in a different environment (e.g., shopping online rather than at the mall).

4. EVIDENCE OF ACTUAL COMMITMENT DEVICES

This section presents evidence of people actually making use of commitment devices. First, we review laboratory evidence in which individuals impose commitments on themselves.

Second, we discuss a series of studies explaining how informal institutions often serve as commitment devices. Third, we discuss research that explores more formal, market-supplied, commitment contracts and consider their effectiveness in changing behavior and outcomes.

4.1. Laboratory Evidence

Read et al. (1999) study simultaneous and sequential choice of virtue (long-term benefits) and vice (short-term benefits).[9] An individual is said to make a simultaneous choice if he or she decides now what to consume in the future and is said to make a sequential choice if he or she must decide at the time of the consumption. Commitment, therefore, is transforming a decision from a sequential into a simultaneous decision. Subjects in the Read et al. (1999) experiment chose a series of three movies, either highbrow or lowbrow movies, to watch on three different days. In the simultaneous treatment, the subjects chose all three movies ahead of time, which resulted in 44% of subjects choosing virtue for their first movie, 64% for the second, and 71% for the third. In the sequential treatment, subjects chose virtue approximately 45% of the time for all three movies. This experiment gives evidence for the success of a commitment device in changing behavior—promoting the choice of more virtuous movies—but does not address the demand for commitment because subjects were not given the option to choose whether to select movies simultaneously or sequentially.

Trope & Fishbach (2000) study self-imposed penalties in the context of a medical test. The subjects were told that the test required them to avoid sugary foods for either 3 days or 6 hours if the test is to be effective. This situation can, consequently, be seen as one with a long-term gain (successfully completing the test) and a short-term cost (the pain of avoiding sugary foods). Subjects were then asked to choose a penalty to be imposed on themselves if they failed to avoid sugary foods. On average, those required to fast for 6 hours set a penalty of $1.49, and those required to fast for 3 days set a penalty of $3.86. This is direct evidence of self-imposed commitment—both sets of agents chose to make the choice to eat the sugary foods more expensive. As subjects facing greater temptation imposed higher costs on their own potential failure, the experiment also suggests that people require larger penalties to motivate more difficult tasks.

In the same paper, Trope & Fishbach (2000) studied whether subjects were willing to receive a reward contingent on completing a task. Subjects were told that they would undertake testing to assess their risk of heart disease. Subjects were assigned to two treatments. In one treatment, the test was described as "strenuous and unpleasant," and in the other it was described as "easy and comfortable." The subjects were told that they would receive a payment, which they could either receive before the test (unconditional on completing the test) or after the test (conditional on taking the test), and were asked to give a preference on a six-point scale for receiving the reward conditional on the test. Of those students to whom health was important, the average interest in the conditional payment was 4 in the unpleasant treatment and 2.5 in the easy and comfortable treatment. Among those to whom health was not important, the impact of the treatment was reversed, and overall interest was lower. Agents seemed to demand commitment to help them with a difficult task but only sought that commitment if they saw a long-term benefit.

[9]Their definition of a virtue is similar to DellaVigna & Malmendier's (2004) definition of an investment good.

Ariely & Wertenbroch (2002) studied the choices and performance of students who had to undertake three tasks in 21 days. The students were divided into three groups. First, the even-spaced group was told that their papers would be due at evenly spaced intervals throughout the 21 days. Second, the maximum-flexibility group was told that they could hand the papers in at any time before the end of the course. Third, the free-choice group was given the choice to set deadlines at the beginning of the course. The students were told that any deadlines they set would be enforced. Ariely & Wertenbroch find that students in the free-choice group chose to set deadlines on themselves that were before the last day possible. Furthermore, students in the free-choice group performed better than students in the maximum-flexibility group and completed fewer of their tasks late. Interestingly, those in the even-spaced group performed best of all. The results indicate a demand for commitment and show that providing the ability to commit can improve performance; however, the results also show that agents may not be able to demand the optimal commitments, perhaps because of some form of naïveté, temptation over menus, or, as Ariely & Wertenbroch argue, because the decision makers can see the normative appeal of removing all restrictions on themselves.

4.2. Informal Commitments in the Field

A preference for commitment also provides an explanation for several financial behaviors observed in developing countries, including informal deposit collection, rotating savings and credit associations (ROSCAs), and participation in microcredit programs. Although similar issues arise regardless of income, evaluations tend to focus on the poor for reasons discussed above—a lack of commitment is more consequential for the poor (Mullainathan & Shafir 2009), and temptation spending may represent a larger share of total expenditures (Banerjee & Mullainathan 2009).

Besley (1995) finds that local savings opportunities in West Africa offer negative interest on savings; in other words, some people are willing to pay to have money taken out of their hands. Whereas one interpretation of this finding is simply that individuals need to protect their savings from theft, from spouses, or from neighbors, another interpretation is that people are willing to pay to have their savings protected from themselves and their own impulses. However, the most direct evidence on this proposition—that some people may value a savings product just for its commitment value—is unclear. Ashraf et al. (2006a) asked time-discounting questions of a group of individuals who were offered a deposit-collection service. They see a lower uptake of deposit-collection services among those who exhibited preference reversals consistent with quasi-hyperbolic discounting.[10] This offer was made subsequent to the offer of a commitment savings account SEED, which we discuss in detail below, and thus the results could be a confound because of their commitment needs already being satisfied. Or it could be the simpler explanation that in this context the deposit-collection service was not perceived as a commitment device.

For other low-income savers, ROSCAs may serve the role of a commitment device. ROSCAs function as group-savings mechanisms, in which group members periodically meet and contribute their savings to a communal pot, which is then awarded at the end of the meeting to one of the group members. Many theories have been put forward to explain the use of ROSCAs and in particular why individuals participate rather than save on

[10]See the discussion in Section 5 for more details of the exact tests used and other studies that use similar tests.

their own. It is important to understand that once one joins a ROSCA, failure to save incurs the cost of a loss of social collateral (i.e., ire from one's fellow ROSCA members). Thus ROSCAs have been put forward as a commitment device. Gugerty (2007) presents econometric as well as ethnographic evidence supporting this hypothesis and quotes from Kenyan women that "saving money at home can make you extravagant in using it," "sitting with other members helps you to save," and "you can't save alone." However, as mentioned above, an alternative but not mutually exclusive explanation is strategic, to keep money away from others (Anderson & Baland 2002).

Ambec & Treich (2007) also show how ROSCAs can serve as a self-control mechanism. Hyperbolic consumers desire to make extra cash unavailable to themselves so that the cash is not spent on temptation goods of no long-term value. If accumulated into larger sums, then the temptation for small frivolous goods is overwhelmed by the ability to buy a large indivisible good of long-term value to the consumer (e.g., a roof for their home). By waiting until they win the ROSCA pot, individuals can then spend this accumulated extra cash on such a long-term value good. This model yields the empirical prediction that the poorest people, who have little to no extra cash, will not be ROSCA members, and the amount contributed to the ROSCA will increase with income. Ambec & Treich then show evidence that these predictions are borne out in practice, supporting their model of ROSCAs as commitment devices for hyperbolic consumers.

A third model of ROSCAs as a commitment savings device is found in Basu (2008). Using a novel approach that does not rely on either a contracting assumption [as Ambec & Treich (2007) do] or a social-sanction enforcement mechanism [as Gugerty (2007) does], Basu shows that sophisticated, quasi-hyperbolic savers can use a ROSCA to impose discipline on themselves, independent of their peers. The key insight is that ROSCAs have two commitment features that a time-inconsistent saver could not access on his or her own: the illiquidity of accumulated savings and the fixed schedule of periodic payments into the ROSCA pot. For quasi-hyperbolic savers with a strong present bias, Basu shows that the illiquidity feature can be important for making ROSCA participation worthwhile; for other savers who are more weakly hyperbolic (less present biased), the most important feature is the regularity of the deposit schedule, which commits agents to save more regularly than they would on their own. Because the agents are sophisticated, they know that their participation in future ROSCAs will be welfare enhancing, so they avoid reneging on their commitment to save in the current ROSCA round (which would bar them from future ROSCA participation). And, because this same logic applies equally well in all future time periods, agents know they have a meaningful commitment to continue saving through the ROSCA even when they are tempted not to.

Basu then enriches his model by including a discrete measure of the information problems that may be faced by ROSCA members in screening out individuals who have reneged on previous ROSCA commitments. His analysis of anonymity, "partial reputation," and "full reputation" yields the testable prediction that ROSCAs in urban areas (where anonymity is higher) are more likely to be fixed ROSCAs, whereas ROSCAs in rural areas are more likely to be random. Citing data from Gugerty (2007) and from Anderson & Baland (2002), Basu argues that there is at least limited evidence consistent with that prediction.

Kast & Pomeranz (2009) present further evidence that group savings can function as a commitment device even in the absence of contracting and social sanctions. In a randomized trial with microcredit clients in Chile, they test a regular savings account and a high-interest savings account against a group-savings program, in which individuals

can publicly announce both their savings goals and their weekly savings deposits into their (private) accounts. Evidence suggests this commitment device is successful in getting the clients to save more.[11] The peer-savings treatment resulted in 65% higher balances than the regular savings accounts, but the mechanism for its success is unclear. The authors discuss possible interpretations, such as the information-based model of Battaglini et al. (2005) or, alternately, a reputation and signaling model. Future work could test whether the periodic deposit feature of the model in Basu (2008) could also explain these results and in particular whether this particular commitment feature is more effective for less present biased individuals, as Basu would predict.

The use of microcredit has also been posited as a commitment device to save, ironically (and expensively). If the rate of return to households is as high as indicated by microcredit interest rates, then why are households unable to save enough to take advantage of these rates? Why do many microborrowers borrow repetitively, a behavior that seems to be inconsistent with any theory of credit constraints or demand for short-term liquidity? An alternative explanation for the use of microcredit (or other high-interest-rate debt) is offered by Rutherford (2000) and Bauer et al. (2008). They argue that households use microcredit as a means to save. By borrowing, they are raising the price (social shame or bank sanctions) from failing to save. And likewise, this commitment to future payments forces them to be tighter with their investments and consumption decisions and be more frugal on frivolous consumption. Bauer et al. (2008) provide empirical evidence for this claim by showing that people with preference reversals in time-preference questions (using real money) are more likely to be involved in a microfinance organization.

4.3. Formal Commitments in the Field

Although many firms offer products that provide commitment devices, finding commitment products that clearly satisfy the definition we outline in Section 1 is more difficult. Here we focus on four implementations that come closest to our definition of a pure commitment device. The examples here include a retirement savings device in the United States, a charitable-giving device in Sweden, a smoking-cessation contract in the Philippines, and a fertilizer coupon program in Kenya.

Benartzi & Thaler (2004) designed an employer-based saving plan, Save More Tomorrow™ (SMarT), with the aim of increasing retirement savings. For our purposes, the SMarT plan has three interesting features. First, the plan provides two commitments: Contributions to retirement savings are automatic, and the rate of contribution increases as the employee's salary increases. Second, enrollment is voluntary, implying that those who enroll demonstrate a demand for commitment. Third, the commitment is soft: Employees can withdraw from the plan at any time without incurring a cost beyond filling out a form. Benartzi & Thaler (2004) report on the implementation of the program at three employers. The results at the first employer—with the longest period of operation—strongly support the joint hypothesis that people demand commitment and that soft commitments are effective. Of those offered the program, 78% enrolled, indicating strong demand for the commitment, and after four pay raises, 80% of those individuals were still enrolled,

[11]Note that these individuals are all borrowing at interest rates much higher than they earn on these additional savings, so further research will examine whether the success of saving more made them better off through a behavioral mechanism or worse off because of the negative arbitrage in which they were encouraged to engage by their peers.

indicating that a soft commitment was sufficient to ensure their continued participation. Benartzi & Thaler (2004) also argue that SMarT increased savings rates; the most compelling evidence for this comes from the third employer for which SMarT was only offered to one division, leading to a 1.5% increase in savings rates, whereas savings at the other divisions stayed constant.

The success of SMarT leads us to ask which of its features are most important. Although there are many dimensions along which the plan could be altered, two areas seem to us to be key. First, although the program shows that soft commitments are effective, we wonder how the results would be affected if the commitment were hard, involving a penalty for withdrawal from the program. We suggest that a hard commitment would decrease enrollment, but also increase retention. If that is true, it is interesting to ask about the optimal penalty level. Second, we wonder what the optimal rate of savings increase is and whether the rate of increase should be a choice. Again we imagine that there is a trade-off between take-up and effectiveness that would be worth exploring.

Similar to SMarT, Breman (2009) tests the same approach for charitable fundraising. In two field experiments, charities offered prior donors the opportunity to "give more tomorrow," i.e., to increase their future, but not current, donations. In both experiments, the average gifts increased when this soft commitment was offered relative to a request to increase current donations (i.e., a "give more now" option), and the increase persisted after the initial increase.

Another formal commitment savings device was provided by the SEED (Save, Earn, Enjoy Deposits) accounts discussed in Ashraf et al. (2006b). These authors worked with a rural bank in the Philippines to offer 700 individuals an option of opening a commitment savings account. The SEED program offered the choice of two commitment features: either (*a*) a time-based maturity, in which the account balance would become available only at a future date, or (*b*) an amount-based maturity, in which the account balance would become available only once the account holder reached a specified savings goal. SEED accounts, therefore, offered reduced liquidity without any compensating increase in interest rate but nevertheless had a high (28%) take-up. Importantly, the SEED accounts were effective in increasing savings. Altogether, those who were offered accounts increased their savings held at the bank by approximately 80% compared with a control group after one year, and the treatment-on-the-treated estimate showed that the subgroup of individuals who actually opened the account saved an estimated 300% more than they would have without the account. Although this evidence seems to support the demand for and effectiveness of a strong commitment product, it is still open to a soft-commitment interpretation. Specifically, agents may have been willing to pay the price of having a hard commitment to gain access to an account that was merely labeled as being about achieving their goals.

Managing and overcoming addictions make up another area for which commitment devices may be important. Bernheim & Rangel (2004) document addicts' ability to manage the cues that trigger their addictive behavior. For example, recent ex-smokers avoid bars, restaurants, or other circumstances that might provide complementary cue goods and increase their likelihood of smoking. These are all examples of soft commitment, insofar as the ex-smoker is avoiding the psychological cues of being at a bar, but not imposing any direct penalty on him or herself for smoking. Some addicts also make use of hard commitments. Alcoholics are known to use certain metabolism-inhibiting drugs, such as Antabuse (also known as disulfiram), which temporarily modify the body processes that metabolize alcohol and produce a disagreeable physical reaction upon alcohol

consumption. These behaviors represent ad hoc means for addicts to make a behavioral change now that reduces their desire to engage in an impulsive behavior in a later state; because the addicts do not derive any utility from this current commitment, this behavior seems to be unambiguously a commitment device. Although these hard commitments for alcoholics have found some success (Goldstein 2001), field studies show that retention rates for Antabuse are poor—often less than 20% (Galanter & Kleber 2008).

A commitment device to stop smoking (CARES) was tested by Giné et al. (2010). In this study a Philippine bank offered smokers an opportunity to open a CARES (Committed Action to Reduce and End Smoking) savings account for the express purpose of giving themselves an incentive to quit. Six months after opening the account, smokers were required to take a urine test for smoking cessation, putting their balance on the line if the test showed they had been unable to quit. The contract was taken up by 11% of smokers offered the account, and on average participants had a balance of 585 pesos ($US11) after six months, some 535 pesos more than the minimum balance. Most importantly, smokers randomly offered CARES were 3 percentage points more likely to pass the six-month test than the control group, and these results held up in a 12-month follow-up study (six months after the smokers were allowed to withdraw their money).

Duflo et al. (2009) consider a novel commitment device in the context of fertilizer use in Kenya. They argue that, similar to the Eliaz/Spiegler version of naïveté, farmers believe that there is some chance that they will be too impatient at the time of planting to purchase fertilizer. Farmers, however, are incorrect in their predictions, being overly optimistic about their chances of investing, leading to constant underinvestment. This formulation implies that a small time-limited discount available immediately after harvest will have a larger impact on take-up than a similar discount at the time of planting because the farmer wishes to reduce the probability that he will not invest in the future. Finally, if the farmer is allowed to choose the time at which a discount is given, he will choose the earlier time period because he realizes that the earlier discount will be more effective. Duflo et al. (2009) test this prediction by providing such a choice between timed discounts in the form of free delivery of fertilizer. Consistent with the model, they find that the early discount leads to greater use of fertilizer than a later discount—39% of farmers versus 20%. Furthermore, they find that 47% of farmers choose to receive the discount earlier, a result that cannot be explained if the farmers are time consistent and have alternative uses for their money. This study suggests that farmers were able to use the discount program to commit themselves to invest in fertilizer. Interestingly, by taking advantage of the stochastic nature of the self-control problem, the commitment device finds a way to help the farmers despite their naïveté. An open question remains: Given the demand for this, and the simplicity with which agricultural supply stores could provide this, why do more stores not offer prepurchased fertilizer coupons?

5. EVIDENCE FOR PARTICULAR THEORIES THAT LEAD TO DEMAND FOR COMMITMENT DEVICES

5.1. Direct Evidence on Discount Rates

Some of the earliest evidence in support of time-inconsistent or temptation preferences—both of which predict a demand for commitment—comes from laboratory experiments in psychology and economics. Two types of evidence are often cited. First, discount rates,

elicited using a subject's willingness to pay for rewards at different dates relative to today, tend to be decreasing over time. Thaler (1991), Loewenstein & Thaler (1989), Loewenstein & Prelec (1992), Kirby & Herrnstein (1995), Ainslie (1992), and Benzion et al. (1989) cite early evidence to this effect.

To take one example, the subjects in Thaler (1991), students from the University of Oregon, were told that they had won a monetary prize that they could take now or at successively later dates. They were asked how much they would require to make them indifferent between receiving the prize now or at the later dates. Thaler's results show that the amount of discounting is initially high but then levels out. For example, the median subject was indifferent between $15 today, $30 in three months, $60 in one year, and $100 in three years. The first comparison implies a three-month discount rate of 0.5, whereas the third implies a three-month discount rate of 0.85. These results reveal time inconsistency, but the data are clearly not rich enough to imply quasi-hyperbolic discounting per se.

Second, and more directly, researchers have asked individuals to choose across same-length time periods that vary in their proximity to now and show that preferences reverse as the first time period moves further into the future. The intuition from Thaler (1991) is that an agent may prefer one apple today to two apples tomorrow, but will almost certainly prefer two apples in a year and a day to one apple in a year; thus many people display preference reversals. Recent evidence comes from two studies in developing countries. In a study in the Philippines, Ashraf et al. (2006b) ask "would you prefer 200 pesos now or 250 pesos in a month" and "would you prefer 200 pesos in 6 months or 250 pesos in 7 months?" They repeat these questions for various peso values and find that 28% of their sample of Filipino microfinance clients exhibit preference reversals in which the client is more patient in the future time period than in the current time period.[12] Bauer et al. (2008) ask similar questions with real money (whereas the Ashraf et al. study used hypothetical questions) of Indian microfinance clients, and find that 20% of their sample exhibit reversals. In a developed-country context, Meier & Sprenger (2010) ask similar questions, also with a six-month lag, and find that 36% of respondents in a low-income urban area exhibit present-biased preference reversals.[13] These findings support some kind of break with the time-consistent exponential discounting model but again can be reconciled with a number of models of time-inconsistent preferences, including both the quasi-hyperbolic and GP models presented above (for more recent evidence consistent with preference reversals, see Tanaka et al. 2010).

It will be important to understand when evidence of time inconsistencies does, and does not, imply a demand for commitment. Although ultimately there is good evidence that preference reversals are correlated with demand for commitment devices, there are also four theories suggesting when, and why, they might not. Before reviewing these theories, however, we note that, because they do not predict demand for commitment devices (in particular, demand from those who demonstrate inconsistent preferences through surveys), none of these arguments can be a complete explanation.

[12]Nearly 15% had reversals in the other direction; that is, they were more patient in the current period than in the future. The impact of such preferences is discussed, for example, by Ameriks et al. (2007).

[13]Intriguingly, Meier & Sprenger also find that present-biased individuals are significantly more likely to carry a balance on their credit cards and carry significantly larger balances than other, time-consistent respondents. This suggests that present-biased reversals may indeed be evidence of self-control problems; however, this evidence is correlative and not necessarily causal.

The first argument, from Rubinstein (2003), claims that a similarity-based decision-making process, rather than utility or preference relations, can explain preference reversals. The idea is that any choice over time is a choice between pairs (x,t), where x is a reward and t is a time. Rubinstein argues that if there is similarity in one dimension, either between the x's or t's, then an agent's choice is made on the basis of the other dimension. Hence preference reversals could be explained by arguing that a year from now and a year and a day are similar, whereas today and tomorrow are not similar. Such an explanation would not imply demand for a commitment device.

The second argument points out a risk confound. Keren & Roelofsma (1995), Andersen et al. (2008), and Fernandez-Villaverde & Mukherji (2002) argue that the time today is privileged in its relationship to uncertainty (or lack thereof). An agent knows everything about today, in particular, if he or she is hungry, for example. Therefore, when deciding between one apple today and two tomorrow, the agent is choosing between a certain outcome (eating, given his or her hunger at this moment) and an uncertain outcome (eating tomorrow, in an uncertain state of hunger), whereas his or her choices for one year from now are made under equally uncertain conditions. Fernandez-Villaverde & Mukherji (2002) show that the differing risks are sufficient to drive preference reversals, without implying a demand for commitment.

The third argument puts forward an alternative functional form that can accommodate the typical answers to time-discounting questions. Benhabib et al. (2006) elicited from a group of NYU students their discount rates for money. They find that a small fixed cost, on the order of $4, is associated with delaying a reward to any time in the future, when compared with today. Such a specification can create preference reversals but does not imply a large demand for commitment.

The fourth argument points out that preference-reversal questions, if viewed as pure finance questions, should be answered as such, and not as questions about consumption that relate to one's utility function (if the individual has access to complete financial markets for debt and savings). Thus, there is a correct answer in that one answer will be net-present-value maximizing for respondents, and even agents that are hyperbolic, quasi-hyperbolic, GP, and so on, will always maximize the net present value of their future cash flows. This argument has been made in several settings (e.g., see Mulligan 1996) and implies that those who appear time inconsistent must be suffering from other biases in decision making, which may or may not predict commitment.

Reviewing these four arguments, we see that our question—whether preference reversals imply a demand for commitment—may hinge on several factors: (*a*) Is an agent facing a choice for which, as in Rubinstein (2003), similarity-based decision making is appealing? (*b*) Is the risk confound significant enough, as in Fernandez-Villaverde & Mukherji (2002), to sway an agent's choice? (*c*) Is the commitment good an item that, like the sums of money studied in Benhabib et al. (2006), may not be tempting enough to imply more than a fixed cost for delay? (*d*) Finally, is this agent's decision, as in Mulligan (1996), best modeled as a financial decision rather than a consumption decision? Understanding these arguments is key to identifying why, and when, preference reversals might not imply a demand for commitment.

However, as noted above, there is also strong evidence that, under certain circumstances, preference reversals do imply a demand for commitment. Ashraf et al. (2006b) test this in relation to the SEED accounts discussed above. Recall that SEED accounts reduced liquidity without any compensating increase in interest rate and that 202 clients

(28% of those offered) opened a SEED account. Ashraf et al. (2006b) show that clients who exhibited preference reversals were significantly (12.5%) more likely to open a SEED account than those who did not exhibit preference reversals. The question of why preference reversals were correlated with a demand for commitment in the SEED context, but perhaps not in other contexts described by the four arguments above, remains open for further research. We anticipate that the answer to this question will require a better understanding of soft commitments and hence also the cognitive processes behind commitment.

5.2. Long-Term Evidence on Discount Rates

Walter Mischel has a series of studies that correlate the short-term discount rates of four year olds with their long-run life outcomes. The children were given a small amount of candy and told that if they did not consume the candy, they would receive a larger amount later.[14] Although this setup does not document preference reversals, it does document very high discount rates for the substantial portion of children who preferred the smaller amount of candy, and this is, we believe, enough to infer a self-control problem. Specifically, if we observe an agent choose one marshmallow now over two marshmallows in an hour, then standard exponential discounting and linearity in utility from marshmallows would tell us that the same agent will also choose one marshmallow now over 16,777,216 marshmallows in a day.[15] The key point is stated by O'Donoghue & Rabin (2006): "Without [present bias], economists have no coherent model of short-term impatience."[16] Hence Mischel's study of these children's long-run life outcomes can be seen as a laboratory for the long-run impacts of self-control problems.

Generally speaking, Mischel found that the children's behavior in these experiments predicts future life success 10 years later. Those who were better able to delay gratification are more attentive, better able to concentrate, exhibit greater frustration tolerance (Shoda et al. 1990), have higher SAT scores, are perceived as more competent by their parents and peers (Mischel et al. 1989), and are less likely to take drugs (Ayduk et al. 2000). This series of studies suggests that children who are more patient or are able to manage their impatience go on to live more successful lives and suggests that commitment devices may indeed someday have far-reaching impacts.

Ameriks et al. (2007) also present evidence on the life impact of self-control problems in a sample of TIAA-CREF clients. They use a novel survey-based approach to determine whether an individual has a self-control problem. They present subjects with a hypothetical situation in which the subject has won 10 dinners to be used over the next two years. The subjects are first asked how they would ideally spread the dinners over the next two years. Next they are asked how they would actually spread the dinners. Those who expect to use

[14]The reactions of the children to these experiments are worth watching (**http://www.youtube.com/**, search "Mischel marshmallow").

[15]That the agent chooses one marshmallow over two in an hour implies $\delta \leq \frac{1}{2}$. Therefore over 24 hours the comparison is 1 versus $\delta^{24}\, x$. This implies $x \geq \frac{1}{(1/2)^{24}} = 16{,}777{,}216$.

[16]Interestingly, the children often waited for a significant amount of time before giving in to the temptation and consuming the candy. This sort of behavior suggests that children are not sophisticates with regard to their temptation, because an agent that knows that they will eventually give in should give in immediately. This sort of behavior is also reported by Skiba & Tobacman (2008) in relation to pay-day loans. They note that many defaulting borrowers have already paid out many pay-day loans before their default and argue that the borrowers should therefore be modeled as partially naïve agents.

the dinners earlier than the ideal are seen as being tempted by current consumption, whereas those who expect to consume them late are seen as tempted by delay. The data collected suggest that 12.1% of the sample are tempted by overconsumption and that 18.6% of the sample are tempted by delay (taking the dinners too late). The researchers also find that the gap between expected and ideal consumption is strongly predictive of lifetime wealth accumulation, with the average over consumer accumulating 20% less wealth than someone with no self-control problem and the average overdelayer accumulating 25% more. Again the study shows the potential for commitment to have real and far-reaching impacts in people's lives.

5.3. Neuroscientific Evidence

Neuroscience, psychology, and psychiatry have long traditions of treating the brain or person as modular, consisting of at least two parts that are in some kind of conflict. Research in these fields can be mapped onto, and hence can provide evidence for, the economic dual-self theories discussed above. And importantly, this evidence generally complicates economic models of self-control and commitment, and points to the need to better understand soft commitment.

For our purposes we consider two main features of the neuroscientific dual-self theory of the brain. First, the brain has two sources of motivation with conflicting desires. Second, the ability of the controlling center to alter behavior is limited.[17] Both of these propositions have found support in neuroscientific studies.

Evidence consistent with the first hypothesis can be found in the study of McClure et al. (2004). They use functional magnetic resonance imaging to measure brain activity in subjects making intertemporal decisions. They show that the limbic system (often associated with affective decision making) is activated preferentially by decisions involving immediate reward, whereas decisions involving future rewards activate only the lateral prefrontal cortex (often associated with control and long-term planning). McClure et al. (2004, p. 506) summarize their findings as supporting the conclusions that "human behavior is often governed by a competition between lower level, automatic processes that may reflect evolutionary adaptations to particular environments, and the more recently evolved, uniquely human capacity for abstract, domain-general reasoning and future planning." This evidence is at least suggestive, if not conclusive, of the existence of two separate sectors of the brain with different aims (for different evidence that is more consistent with a single motivation, see Glimcher et al. 2007).

Evidence for the second hypothesis is less direct but compelling. Intuitively, if control is limited, then giving a person a task that uses the resources of the controlling center should change behavior, making lower-level behavior more dominant. This is exactly what is found by Shiv & Fedorikhin (1999). They randomly assigned subjects to remember a two- or seven-digit number. The subjects were then asked to walk to another room, being offered on the way a choice between a slice of cake or a bowl of fruit salad. Fifty-nine percent of those in the seven-digit treatment chose the cake, whereas only 37% of those in the two-digit treatment did. One appealing interpretation of this finding is that the controlling center has a limited capacity and that it cannot divert resources to self-control

[17]It is not clear how we decide that the controlling part of the brain should be the part that we use for welfare comparisons.

when it is attempting to do other things.[18] A more direct approach can be found in the work pioneered by Baumeister et al. (1994). This literature first primes treatment subjects, requiring them to engage in an activity that requires self-control (e.g., not looking at words on the bottom of a television screen). They are then asked to make a decision in which self-control is important. In general, the findings show that treatment subjects act more impulsively in the second activity, suggesting that some of the controlling center's capacity was exhausted by the first activity's self-control demands. In an interesting recent study in this vein, Vohs & Faber (2007) ask subjects to write an essay about anything that came into their head and told half the participants not to think about a white bear while doing so. Everyone was then offered the opportunity to spend 10 dollars on impulse consumption items (e.g., gum and playing cards) or keep the money and go home. On average, those who were asked to not think about a white bear spent $4.05 on impulsive goods, whereas the control group spent only $1.21 ($p < 0.001$). It would be interesting to see an experiment that combines the two approaches from McClure et al. (2004) and Shiv & Fedorikhin (1999), simultaneously looking at limited self-control and the area of the brain that is working.[19]

While supporting a dual-self model, this evidence complicates economists' understanding of commitment devices and agents' demand for them. Specifically, many commitment devices do not take options off the table, but rather make one of the options more costly. However, it seems that the level of temptation presented to the so-called primitive brain (e.g., by a chocolate cake) should be independent of its cost, so long as that cost is pushed into the future. The key question is how we imagine the brain to work. On the one hand, self-control may simply be limited; an agent can either remember large numbers, or can exercise self-control, but cannot do both. In this case, changing the price of the chocolate cake would have no impact on the agent's choice. On the other hand, we may consider self-control to be represented by an increasing convex cost function. In this case, increasing the price of the chocolate cake might make it worthwhile for the agent to exert the effort necessary to resist the temptation.

A third interpretation of the experiments, however, is that they confirm the importance of soft commitment. Rather than making a hard commitment to not eat chocolate cake, a more useful commitment in this case would be to avoid encounters with chocolate cake whenever one is, so to speak, trying to remember long numbers, or in a more real-world example, relevant to (one of) the authors, to remove tempting foods from view while working hard on a review paper.

Finally, we note that research on this neuroscientific evidence could fruitfully proceed in the same way as research on preference reversals. Shiv & Fedorikhin (1999) suggest an alternate measure of self-control, different from the preference-reversal questions used by Ashraf et al. (2006b) and the survey questions used by Ameriks et al. (2007): How many digits can an individual remember and still avoid the chocolate cake? It would be interesting to see whether such a measure correlates with the use of any kind of commitment device.

[18]Strictly speaking, this explanation requires that cognition and self-control centers are the same.

[19]Work in this direction has already begun. For example, Knoch & Fehr (2007) use rTMS (repetitive transcranial magnetic stimulation) to disable particular parts of the brain and then consider the ability of the subject to resist temptations. They find evidence that they believe implies that the right prefrontal cortex is responsible for our capacity to resist temptation.

6. MARKET PROVISION OF COMMITMENT DEVICES

Will the market provide efficient commitment contracts? We highlight three challenges that face markets, firms, or governments in the development of commitment products: the trade-off between commitment and flexibility, the difficulty in writing contracts to consume less, and the public-policy concern of potential exploitation of partially naïve agents.

6.1. Trade-Off Between Commitment (Completion) and Flexibility (Take-Up)

A commitment contract must weigh the need for commitment against an agent's wish for flexibility. A time-inconsistent individual may want to sign a contract to exercise four times a week but also want the flexibility to break the obligation if he or she has a cold, or a particularly hard work week. If the conditions requiring flexibility are observable, then contracts can be efficient. For example, the contracts in Ashraf et al. (2006b) allow for money to be withdrawn in the event of a medical emergency or relocation away from a bank branch. However, in many cases the circumstances requiring flexibility cannot be observed or verified (a contract provider, for example, cannot easily determine whether an individual has a hard work week). These issues are addressed theoretically in two recent papers. Amador et al. (2006) consider a two-period consumption saving model with an agent who has GP or hyperbolic preferences and faces a stochastic demand shock. They show that a minimum savings requirement (i.e., a contract that requires the agent to save at least x number of dollars per period) is always part of the optimal contract. The resulting contract, however, is not efficient as the agent underconsumes (overcommits) at some shocks. This result could easily be extended, suggesting that some people will be harmed ex post by an optimal commitment contract. Bond & Sigurdsson (2009) reconsider this result in a three-period version of the same model and provide conditions under which an efficient contract can be provided. The result differs from Amador et al. (2006) as a three-period model allows the agent in time 1 to be punished by a period 2 agent. For example, a period 1 agent can be punished by allowing a period 2 agent to overconsume. In practice we do not see complicated contracts of this second form, but we do often see minimum savings contracts.

6.2. Contracts to Limit Consumption

Contracts directly on consumption can be problematic for two reasons: market response and an imprecise relationship between consumption and one's true goal.

Intuitively, a contract that raises the price of consumption—for example, a contract to raise the price of cigarettes or fatty foods—will be undercut in a competitive market. DellaVigna & Malmendier (2004) study leisure and investment goods in the absence of these competitive pressures, and they find that if sophisticated consumers can sign exclusive contracts with firms, then both leisure and investment goods will be efficiently provided. Leisure goods are goods with immediate benefit and long-term cost (e.g., fatty foods, cigarettes for some), whereas investment goods have long-term benefit and immediate cost (e.g., membership to a health club).

An efficient contract for an investment good involves a high upfront fee and below marginal cost usage fee, encouraging greater use. Gym contracts often take this form (i.e., a monthly or annual fee, and no marginal cost for each trip to the gym). For a leisure good, the contract involves the reverse: an above marginal cost usage fee and subsidized sign-up.

Gottlieb (2008), however, shows that if firms are not able to write exclusive contracts, leisure goods will be provided by others at marginal cost, thus leading time-inconsistent agents to overconsume. The intuition, as above, is that other firms can always undercut the high prices charged by the counterparty to the commitment contract. In a simple example, if one tries to write a contract with their convenience mini-market to not sell them cigarettes (i.e., raise the price to infinity), then the consumer can still go a mile down the road to another store, who will eagerly sell them cigarettes at marginal costs. A similar example is found in consumer financial markets: Easy and tempting credit can potentially unravel a commitment saving contract. Naturally, a consumer who realizes this will not undertake the commitment in the first place.[20] This market response is consistent with credit-card contracts that subsidize take-up (e.g., offer teaser rates) but charge high later interest rates.

Regulation or market innovation might compensate for the inability of the market to efficiently provide leisure goods. Sin taxes are one frequently discussed response. Sin taxes raise the price of common leisure goods, such as cigarettes, alcohol, and soda, and if enough people are time inconsistent, optimal taxation requires imposing sin taxes, even if there is a cost to those who are not time inconsistent (O'Donoghue & Rabin 2003; see also Krusell et al. 2002 and Krusell & Smith 2003 for a discussion of optimal taxation in which not only consumers but also the government lack commitment). Gruber & Mullainathan (2005) provide empirical evidence that people are happier in states with higher cigarette taxes, and that this is true even among smokers, suggesting that the argument for sin taxes may be correct. Alternatively, market innovation could solve the problem by requiring an agent to meet certain goals that are not brand or firm specific. Such a contract effectively raises the price for all consumption of a particular type of good, rather than for the consumption provided by particular firms. So, rather than raising the price of Marlboros, in particular, a third-party contract aims to raise the price of all smoking, and rather than committing to putting money in a bank account, an agent commits to have a certain level of credit available at a certain time. A website that provides such contracts to the general public is **http://stickK.com**. Since its launch in January 2008, it has generated almost 50,000 contracts, mostly on health-related goals such as weight loss, exercise, and smoking. It remains to be seen whether this sort of market-provided contract can remove the need for taxation and regulation.

An alternative response to the difficulty of contracting with respect to consumption (i.e., such as eating candy) is to instead contract on outcomes (i.e., states of the world, such as being fat). However, processes may also be hard to contract. For example, in the case of obesity, the outcome of interest is obesity (or health, more generally), not food consumed. However, physiological differences across bodies can make it less than optimal to write strict contracts on weight. Furthermore, if the goal is more about being healthy, then the outcome is some combination of factors, including muscle mass, cholesterol level, heart condition, and body fat. Naturally one could construct contracts on each, or on some linear or nonlinear combination of each, but this quickly becomes overcomplicated and leads to a contract that is difficult to market.

However, we note that contracts on outcomes by no means dominate those on actions. In many situations, the outcome of interest is determined only partly by the agent's action. Then, it may be suitable to contract on the action of the agent rather than the outcome, as it removes the probabilistic element from the process. For example, one could write a

[20]There is still scope for partially naïve consumers to take-up commitment savings contracts.

contract to send out 100 resumes to look for jobs, but should not necessarily write a contract to get hired. Thus the answer is far more nuanced and requires an understanding of the likely competitive response, and the deterministic relationship between the action and the desired outcome.

6.3. Overenrollment

Naive and partially naïve individuals may also be inefficiently served. Naive individuals are unlikely to demand commitment contracts, implying that some sort of intervention is necessary to allow for commitment. Furthermore, partially naïve agents can be exploited by the market and may need to be protected through regulation. In the CARES study discussed above, some agents signed up, deposited money, and still failed the nicotine test, which may be indicative of partially naïve agents underestimating the size of commitment required to discourage smoking. [Alternately, it may be that customers failed to quit because they were hit with a shock that make quitting more expensive. If this is true, failure to comply is an example of the inefficiency of contracts highlighted by Amador et al. (2006).] Three papers explore these issues in formal models. DellaVigna & Malmendier (2004) consider the issue in a model in which firms know if an agent is naïve, whereas Eliaz & Spiegler (2006) study the general case in which firms do not know an agent's type. Both studies show that although sophisticates are served efficiently, partially naïve agents are exploited by the market. Heidhues & Koszegi (2008) apply a similar model to the credit market and show that the presence of partially naïve consumers implies contracts with high penalty fees for late repayment, as we see in reality. Heidhues & Koszegi (2008) then show that regulation requiring linear credit contracts can be welfare improving. Other options might include educating people as to the extent to which they are likely to default, as in Gabaix & Laibson (2006) and Bertrand & Morse (2009).

7. WELFARE AND CONCLUSIONS

7.1. Welfare

Much of this review has implicitly assumed that if people demand commitment, then it is beneficial to provide that commitment. There are, however, reasons to believe that commitment is not always welfare improving. To illustrate, we return to the runner's dilemma presented above. Recall that Rachel signed a contract requiring her to pay $1,000 if she walked during her training session, and we argued that this would be a commitment contract if time 0 Rachel believes that the benefits of running are greater than the costs, and time 1 Rachel disagrees with this assessment.

Whether Rachel's decision is welfare improving depends on which Rachel has the better information. Rachel's decision to commit will be welfare enhancing if time 1 Rachel will choose to run if the commitment is in place ($B_1 + 1{,}000 > C_1$), and time 0 Rachel is correct in the belief that running is beneficial ($B_0 > C_0 \Rightarrow B > C$). Alternatively, the decision to commit will be welfare reducing if either of these requirements is not met. First, time 0 Rachel may be incorrect in her assessment of the impact of the commitment. That is, it may be that $B_1 + 1{,}000 < C_1$. In this case, Rachel will not run and will forfeit the money—probably a welfare-reducing situation. We return to this issue below under the heading of naïveté. Second, time 0 Rachel may be incorrect in her assessment of the true cost or

benefit of running. That is, we may have $B_0 > C_0$, but $B < C$. In this case, Rachel will still choose to commit so long as $B_1^0 < C_1^0$, but the decision to commit is welfare reducing. This problem is rarely discussed in the literature, but it does not seem unreasonable. On the contrary, it seems natural to assume that time 1 Rachel has good information about the current costs of running, whereas time 0 Rachel may be incorrect in her assessment of the costs. In particular, it seems intuitive that pain becomes less memorable as time goes by, and therefore a nonrunning Rachel may be basing the assessment C_0 on a biased measure. Kahneman et al. (1997) present evidence in favor of this view. Subjects in their experiments reported a more positive memory of a longer, more painful event, so long as the event ended in a less painful manner. This suggests that pain is remembered differently from how it is experienced.

The difficulty with welfare analysis arises because we observe two choices that are in conflict, and the raw choice data are insufficient to tell us which Rachel is better informed. Two main approaches have been proposed to tackle this difficulty. The first approach is to look for preferences that rationalize the choice data, preserving revealed preference. Such an approach is championed most notably by Gul & Pesendorfer (2008). This approach implies that both Rachel's time 0 and time 1 decisions are correct given their circumstances, and that, because the commitment choice occurs first, it is welfare enhancing.

The second approach is to deem one of the decisions as a mistake (i.e., not utility maximizing) and to attempt a welfare evaluation by using the correct decision to establish a welfare criterion. This approach is taken up, for example, by Koszegi & Rabin (2008) and requires that we specify a model that is capable of differentiating correct from mistaken choices, perhaps making use of richer data than are usually used by economists.[21] For example, in the quasi-hyperbolic model discussed above, it is often argued that the preferences of the time 0 agent should be used to measure welfare. This choice in effect assumes that Rachel's choice is always welfare improving. We believe that settling on a particular approach and providing empirical support or clear philosophical arguments for one of these approaches are the hard questions for commitment, which deserves more thought and research. One can imagine, as an objective outsider, giving someone advice to "live a little" or also to "buckle down and get your paper written." So the answer is not as simple as "choose the planner/do-gooder self" when making welfare statements.

7.2. Conclusions

This review attempts to demonstrate two theoretical challenges and three empirical observations. On the theory side, we argue that the welfare implications of commitment devices hinge critically on modeling assumptions and that there is insufficient work to understand the demand for soft commitments. On the empirical side we put forward evidence from the laboratory and field on the demand for commitment devices and the creation of informal commitment devices, as well as the use and impact from formal market-offered commitment devices. The market for commitment devices is young, and several policy questions, directly motivated by the theory, remain unanswered. First, how should commitment devices be targeted? Do commitment devices only work for the sophisticated or partially

[21] An interesting recent contribution by Bernheim & Rangel (2009) takes what might be considered an intermediate approach. They suggest using choice data to infer all the possible welfare-maximizing actions. In some situations, this may be sufficient to make comparisons. In situations in which it is not, they argue that some decisions can clearly be considered to be better.

naïve, and if naïve individuals take up commitment devices, why do they do so and what is the welfare implication? Second, which are more effective, soft or hard commitments? If severe consequences increase effectiveness but lower demand for a commitment device, what is the right balance to maximize impact, and how can devices and contracts be offered so that optimal sorting of individuals to contract strength occurs? Third, what is the role of habit formation? Can commitment devices be employed to generate long-term behavioral change, or merely short-term changes that then require ongoing commitment devices to maintain behavior? For example, Giné et al. (2010) find that long-term behavioral changes occurred, as do Charness & Gneezy (2009) for incentives to exercise, but will such long-term behavioral changes occur in other domains, such as weight loss or savings behavior?

DISCLOSURE STATEMENT

Dean Karlan is President and owns equity in stickK.com.

ACKNOWLEDGMENTS

Dean Karlan would like to thank David Laibson, Matthew Rabin, and Richard Thaler for conversations over the years and Nava Ashraf, Xavier Giné, Wesley Yin, and Jonathan Zinman for collaboration on commitment-related projects. The authors thank Julian Jamison, Richard Mansfield, and Jonathan Morduch for helpful advice and Editorial Committee member Esther Duflo for comments. Karlan is also grateful to the National Science Foundation and Sloan Foundation for support. All errors remain our own.

LITERATURE CITED

Ainslie G. 1992. *Picoeconomics: The Interaction of Successive Motivational States within the Person*. Cambridge, UK: Cambridge Univ. Press

Amador M, Werning I, Angeletos GM. 2006. Commitment vs. flexibility. *Econometrica* 74:365–96

Ambec S, Treich N. 2007. ROSCAs as financial agreements to cope with self-control problems. *J. Dev. Econ.* 82(1):120–37

Am. Lung Assoc. 2008. Trends in tobacco use. *Tob. Trend Rep.*, Am. Lung Assoc., Washington, DC

Ameriks J, Caplin A, Leahy J, Tyler T. 2007. Measuring self-control problems. *Am. Econ. Rev.* 97(3):966–72

Andersen S, Harrison GW, Lau MI, Rutstroem EE. 2008. Eliciting risk and time preferences. *Econometrica* 76(3):583–618

Anderson S, Baland J-M. 2002. The economics of ROSCAs and intrahousehold allocation. *Q. J. Econ.* 117(3):963–95

Ariely D, Wertenbroch K. 2002. Procrastination, deadlines, and performance: self-control by precommitment. *Psychol. Sci.* 13(3):219–24

Ashraf N, Karlan D, Yin W. 2006a. Deposit collectors. *BE J. Econ. Anal. Policy* 6(2):5

Ashraf N, Karlan D, Yin W. 2006b. Tying Odysseus to the mast: evidence from a commitment savings product in the Philippines. *Q. J. Econ.* 121(2):635–72

Ayduk O, Mendoza-Denton R, Mischel W, Downey G, Peake PK, Rodriguez M. 2000. Regulating the interpersonal self: strategic self-regulation for coping with rejection sensitivity. *J. Pers. Soc. Psychol.* 79(5):776–92

Banerjee AV, Mullainathan S. 2009. The shape of temptation: implications for the economic lives of the poor. *Tech. Rep.*, MIT

Basu K. 2008. *Hyperbolic discounting and the sustainability of rotational savings arrangements*. Work. Pap., Univ. Chicago

Battaglini M, Bénabou R, Tirole J. 2005. Self-control in peer groups. *J. Econ. Theory* 123(2):105–34

Bauer M, Chytilová J, Morduch J. 2008. Behavioral foundations of microcredit: Experimental and survey evidence from rural India. *Res. Pap.*, Financ. Access Initiat., New York

Baumeister RF, Heatherton TF, Tice DM. 1994. *Losing Control: How and Why People Fail at Self-Regulation*. New York: Academic

Benabou R, Tirole J. 2004. Willpower and personal rules. *J. Polit. Econ.* 112(4):848–86

Benabou R, Pycia M. 2002. Dynamic inconsistency and self-control: a planner-doer interpretation. *Econ. Lett.* 77(3):419–24

Benartzi S, Thaler R. 2004. Save more tomorrow: using behavioral economics to increase employee saving. *J. Polit. Econ.* 112(1):164–87

Benhabib J, Bisin A. 2005. Modeling internal commitment mechanisms and self-control: a neuro-economics approach to consumption–saving decisions. *Games Econ. Behav.* 52(2):460–92

Benhabib J, Bisin A, Schotter A. 2006. *Present-bias, quasi-hyperbolic discounting, and fixed costs*. Work. Pap., New York Univ.

Benzion U, Rapoport A, Yagil J. 1989. Discount rates inferred from decisions: an experimental study. *Manag. Sci.* 35:270–84

Bernheim BD, Rangel A. 2004. Addiction and cue-triggered decision processes. *Am. Econ. Rev.* 94(5):1558–90

Bernheim BD, Rangel A. 2009. Beyond revealed preference: choice-theoretic foundations for behavioral welfare economics. *Q. J. Econ.* 124(1):51–104

Bertrand M, Morse A. 2009. *Information disclosure, cognitive biases and pay-day borrowing*. Work. Pap., Univ. Chicago, Booth School Bus.

Besley T. 1995. Saving, credit and insurance. In *Handbook of Development Economics*, ed. J Behrman, TN Srinivasan, pp. 2123–207. Amsterdam: Elsevier

Bond P, Sigurdsson G. 2009. *Commitment contracts*. Work. Pap., Univ. Pennsylvania

Breman A. 2009. *Give more tomorrow: two field experiments on altruism and intertemporal choices*. Unpublished manuscript, Stockholm School Econ.

Brunnermeier MK, Parker JA. 2005. Optimal expectations. *Am. Econ. Rev.* 95(4):1092–118

Cent. Dis. Control Prev. (CDC). 2006. Decline in adult smoking rates stall: millions of nonsmoking Americans remain exposed to secondhand smoke. **http://www.cdc.gov/media/pressrel/r061026a.htm**

Charness G, Gneezy U. 2009. Incentives to exercise. *Econometrica* 77(3):909–31

Dekel E, Lipman BL, Rustichini A. 2009. Temptation-driven preferences. *Rev. Econ. Stud.* 76(3):937–71

DellaVigna S, Malmendier U. 2004. Contract design and self-control: theory and evidence. *Q. J. Econ.* 119(2):353–402

DellaVigna S, Malmendier U. 2006. Paying not to go to the gym. *Am. Econ. Rev.* 96(3):694–719

Duflo E, Kremer M, Robinson J. 2009. *Nudging farmers to use fertilizer: theory and experimental evidence from Kenya*. NBER Work. Pap. 15131

Eliaz K, Spiegler R. 2006. Contracting with diversely naive agents. *Rev. Econ. Stud.* 73:689–714

Fernandez-Villaverde J, Mukherji A. 2002. *Can we really observe hyperbolic discounting?* Work. Pap., Univ. Pennsylvania, Univ. Minnesota

Fudenberg D, Levine DK. 2006. A dual-self model of impulse control. *Am. Econ. Rev.* 96(5):1449–76

Gabaix X, Laibson D. 2006. Shrouded attributes, consumer myopia, and information suppression in competitive markets. *Q. J. Econ.* 121(2):505–40

Galanter M, Kleber HD. 2008. *The American Psychiatric Publishing Textbook of Substance Abuse Treatment*. Arlington, VA: Am. Psychiatr. Publ.

Gallup. 2006. Close to 6 in 10 Americans want to lose weight. **http://www.gallup.com/poll/21859/close-americans-want-lose-weight.aspx**
Gallup. 2008. More smokers in U.S. want to quit. **http://www.gallup.com/video/109033/most-smokers-us-want-quit.aspx**
Giné X, Karlan D, Zinman J. 2010. Put your money where your butt is: a commitment savings account for smoking cessation. *Am. Econ. J. Appl. Econ.* In press
Glimcher PW, Kable J, Louie K. 2007. Neuroeconomic studies of impulsivity: now or just as soon as possible? *Am. Econ. Rev.* 97(2):142–47
Goldstein A. 2001. *Addiction: From Biology to Drug Policy.* New York: Oxford Univ. Press
Gottlieb D. 2008. Competition over time-inconsistent consumers. *J. Public Econ. Theory* 10(4):673–84
Gruber JH, Mullainathan S. 2005. Do cigarette taxes make smokers happier. *BE J. Econ. Anal. Policy* 5(1):4
Gugerty MK. 2007. You can't save alone: commitment in rotating savings and credit associations in Kenya. *Econ. Dev. Cult. Change* 55(2):251–82
Gul F, Pesendorfer W. 2001. Temptation and self-control. *Econometrica* 69(6):1403–35
Gul F, Pesendorfer W. 2004. Self-control and the theory of consumption. *Econometrica* 72(1):119–58
Gul F, Pesendorfer W. 2008. The case for mindless economics. In *The Foundations of Positive and Normative Economics: A Handbook*, ed. A Caplin, A Schotter, pp. 3–42. New York: Oxford Univ. Press
Hamermesh D. 2010. *Starting a new phase of my career.* **http://freakonomics.blogs.nytimes.com/2010/01/21/starting-a-new-phase-of-my-career/**
Heidhues P, Koszegi B. 2008. *Exploiting naiveté about self-control in the credit market*. Work. Pap., Univ. Calif., Berkeley
Hughes JR, Keely J, Naud S. 2004. Shape of the relapse curve and long-term abstinence among untreated smokers. *Addiction* 99(1):29–38
Kahneman D, Wakker P, Sarin R. 1997. Back to Bentham? Explorations of experienced utility. *Q. J. Econ.* 112(2):375–406
Kast F, Pomeranz D. 2009. *Peers as a saving commitment device: evidence from a field experiment among low-income micro-entrepreneurs*. Work. Pap., Harvard Univ., Kennedy School
Keren G, Roelofsma P. 1995. Immediacy and certainty in intertemporal choice. *Organ. Behav. Hum. Decis. Process.* 63(3):287–97
Kirby KN, Herrnstein RJ. 1995. Preference reversals due to myopic discounting of delayed reward. *Psychol. Sci.* 6(2):83–89
Knoch D, Fehr E. 2007. Resisting the power of temptations: the right prefrontal cortex and self-control. *Ann. N.Y. Acad. Sci.* 1104:123–134
Koszegi B, Rabin M. 2008. Choices, situations, and happiness. *J. Public Econ.* 92(8–9):1821–32
Koszegi B, Rabin M. 2009. Reference-dependent consumption plans. *Am. Econ. Rev.* 99(3):909–36
Krusell P, Kurusçu B, Smith AA Jr. 2002. Equilibrium welfare and government policy with quasi-geometric discounting. *J. Econ. Theory* 105:42–72
Krusell P, Smith AA Jr. 2003. Consumption-savings decisions with quasi-geometric discounting. *Econometrica* 71(1):365–75
Laibson D. 1997. Golden eggs and hyperbolic discounting. *Q. J. Econ.* 112:443–47
Loewenstein G, O'Donoghue T. 2004. *Animal spirits: affective and deliberative processes in economic behavior.* Work. Pap., Carnegie Mellon Univ., Cornell Univ.
Loewenstein G, Prelec D. 1992. Anomalies in intertemporal choice: evidence and an interpretation. *Q. J. Econ.* 107:573–97
Loewenstein G, Thaler R. 1989. Anomalies: intertemporal choice. *J. Econ. Perspect.* 3(4):181–93
McClure SM, Laibson DI, Loewenstein G, Cohen JD. 2004. Separate neural systems value immediate and delayed monetary rewards. *Science* 306:503–7
Meier S, Sprenger C. 2010. Present-biased preferences and credit card borrowing. *Am. Econ. J.* 2(1):193–210

Mischel W, Shoda Y, Rodriguez MI. 1989. Delay of gratification in children. *Science* 244(4907):933–38
Mullainathan S, Shafir E. 2009. *Savings Policy and Decision-Making in Low-Income Households*. New York: Russell Sage Found.
Mulligan C. 1996. A logical economist's argument against hyperbolic discounting. Unpublished manuscript, Univ. Chicago
Noor J. 2007. Commitment and self-control. *J. Econ. Theory* 135(1):1–34
O'Donoghue T, Rabin M. 1999. Doing it now or later. *Am. Econ. Rev.* 89:103–24
O'Donoghue T, Rabin M. 2003. Studying optimal paternalism, illustrated by a model of sin taxes. *Am. Econ. Rev.* 93(2):186–91
O'Donoghue T, Rabin M. 2006. Incentives and self-control. In *Advances in Economics and Econometrics: Theory and Applications: Ninth World Congress*, pp. 215–45. Cambridge, UK: Cambridge Univ. Press
Phelps ES, Pollak RA. 1968. On second-best national saving and game-equilibrium growth. *Rev. Econ. Stud.* 35(2):185–99
Plateau J-P. 2001. *Institutions, Social Norms and Economic Development*. Newark, NJ: Harwood Acad.
Quirkology. 2007. *New Year's resolution experiment*. **http://www.quirkology.com/UK/Experiment_resolution.shtml**
Read D, Loewenstein G, Kalyanaraman S. 1999. Mixing virtue and vice: combining the immediacy effect and the diversification heuristic. *J. Behav. Decis. Making* 12:257–73
Resolut. Res. 2009. Personal optimism runs high despite broad uncertainty. **http://www.resolutionresearch.com/news-newyear-surveyresult.html**
Rubinstein A. 2003. Economics and psychology? The case of hyperbolic discounting. *Int. Econ. Rev.* 44:1207–16
Rutherford S. 2000. *The Poor and Their Money*. New York: Oxford Univ. Press
Sarver T. 2008. Anticipating regret: why fewer options may be better. *Econometrica* 76(2):263–305
Schelling TC. 1984. Self-command in practice, in policy, and in a theory of rational choice. *Am. Econ. Rev.* 74:1–11
Shiv B, Fedorikhin A. 1999. Heart and mind in conflict: the interplay of affect and cognition in consumer decision making. *J. Consum. Res.* 26(3):278–92
Shoda Y, Mischel W, Peake PK. 1990. Predicting adolescent cognitive and self-regulatory competencies from preschool delay of gratification: identifying diagnostic conditions. *Dev. Psychol.* 26(6):978–86
Skiba PM, Tobacman J. 2008. *Payday loans, uncertainty and discounting: explaining patterns of borrowing, repayment, and default*. Work. Pap., Univ. Pennsylvania
Spiegler R. 2008. On two points of view regarding revealed preference and behavioral economics. In *The Foundations of Positive and Normative Economics: A Handbook*, ed. A Caplin, A Schotter, pp. 95–115. New York: Oxford Univ. Press
Strotz RH. 1956. Myopia and inconsistency in dynamic utility maximization. *Rev. Econ. Stud.* 23(3):165–80
Tanaka T, Camerer C, Nguyen Q. 2010. Risk and time preferences: experimental and household survey data from Vietnam. *Am. Econ. Rev.* 100:551–71
Thaler R. 1985. Mental accounting and consumer choice. *Mark. Sci.* 4(3):199–214
Thaler RH. 1991. Some empirical evidence on dynamic inconsistency. In *Quasi Rational Economics*, pp. 127–36. New York: Russell Sage Found.
Thaler RH, Shefrin HM. 1981. An economic theory of self-control. *J. Polit. Econ.* 89:392–406
Trope Y, Fishbach A. 2000. Counteractive self-control in overcoming temptation. *J. Pers. Soc. Psychol.* 79(4):493–506
Vohs KD, Faber RJ. 2007. Spent resources: self-regulatory resource availability affects impulse buying. *J. Consum. Res.* 33(4):537–47

Annual Review of Economics

Volume 2, 2010

Contents

Errata

An online log of corrections to *Annual Review of Economics* articles may be found at http://econ.annualreviews.org