

Commitment Devices

Gharad Bryan*, Dean Karlan[†] and Scott Nelson[‡] §

April 25, 2010

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 man and problematize our definition of welfare.

This paper 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 due to intra-personal conflict stemming, for example, from a lack of self control. We will exclude actions that accrue significant current benefits, or that are taken with a strategic motive.¹ So,

*Yale University

†Yale University

‡Yale University

§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 the editor, Esther Duflo, for comments. Karlan is also grateful to the National Science Foundation and Sloan Foundation for support. All errors remain our own.

¹Naturally, the phrase “commitment device” is also appropriately used to describe strategic arrangement with respect to others. However, we are limiting the scope of this paper to self-commitment devices, as defined above.

while painting your nails with Control-It Nail Biting Treatment[®] is a commitment device, buying beautiful nail polish to look pretty but that also happens to taste foul is not. And while signing up for automatic savings transfers from a checking account may be a commitment device, doing so in order to limit a spouse’s spending is not.

We will call commitment devices that call for real economic penalties for failure, or rewards for success, hard commitments. And we will call 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 saving 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 saving account labelled “send kids to college”; if someone withdraws money from that account to pay for a holiday party, he 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 of the issues related to sophistication and naïvete, 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 lab 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.

2 Definition and Motivation

In planning this paper, we had intended the introduction to include a concise one line definition of a commitment device. 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 make clear what we mean by a commitment device, and show that defining a commitment device ultimately requires some appeal to an agents motivation.

2.1 Definition

We define a commitment device as an arrangement entered into by an agent which 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, booking plane tickets, etc., are all behaviors that could be seen as “commitment” 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 which facilitate future consumption plans merely for transactional or logistical reasons, and arrangements which are desirable because they lock in (“commit”) a particular consumption path. For example, buying in bulk could be done in order 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 which are used to influence the actions of others. Consider, for example, the “doomsday device” in 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 “destroy everything” 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 wish, however, to note 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 due to spousal, familial and neighbor bargaining.²

At times we will 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

²See, for example, Anderson and Baland (2002) and Plateau (2001).

to restrict his or her own choices.³

2.2 A Formal Example: The Runner’s Dilemma

Our definition requires an understanding of the motivation behind peoples 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 \$1000 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 pre-running 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 .

Further, 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 out 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 paper we present

³Commitment contracts have generated legal discussions since if they are designed to benefit one party, the “enforcing” party must also receive consideration (i.e., some benefit) in order for the contracts to be legally binding. This requirement is typically easy to satisfy typically through the “peppercorn” principle of consideration, i.e., that even if a party receives value no more than a peppercorn, this is sufficient consideration to be a legally binding contract.

harder and more careful evidence on commitment devices but we being with a set of anecdotes:

Smoking The majority of smokers say they want to quit, often try to quit, but continuously fail in their quit attempts. Gallup polling suggests 74% of smokers would like to give up smoking,⁴ 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 a an attempt to quit.⁵ However success rates are low. Data variously suggest that fewer than 5%⁶, or 4% to 7% (Hughes et al. (2004)), of quit attempts annually are successful.

New Year's Resolutions New Year's Resolutions are (infamously) unsuccessful. While evidence is light, one study followed 3,000 resolution makers from 2007. 52% of resolution makers were confident of success, only 12% actually managed to follow through with their plans.⁷ Evidence suggests some of these failures are cyclical: roughly 60% of 2009's resolutions were repeats of failed 2008 resolutions.⁸

Weight Loss In Gallup polling from 2008, 56% of Americans said they wanted to lose weight and 30% of Americans said they were seriously trying to lose weight.⁹ More tellingly, 59% of those interviewed in 2001 said they wanted to lose weight, implying that at least 15% of Americans were still trying 7 years later.¹⁰

Various ad hoc Commitment Devices Other limited anecdotal evidence suggests ad hoc behavior that can be construed as commitment devices, including:¹¹

- Cutting up ones 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;

⁴<http://www.gallup.com/video/109033/most-smokers-us-want-quit.aspx>

⁵<http://www.cdc.gov/media/pressrel/r061026a.htm>

⁶p. 8, "Trends in Tobacco Use, American Lung Association, July 2008.

⁷http://www.quirkology.com/UK/Experiment_resolution.shtml

⁸<http://www.resolutionresearch.com/news-newyear-surveyresult.html>

⁹<http://www.gallup.com/poll/21859/close-americans-want-lose-weight.aspx>;

<http://www.gallup.com/poll/112426/Americans-Weight-Issues-Going-Away.aspx> respectively

¹⁰<http://www.gallup.com/poll/112426/Americans-Weight-Issues-Going-Away.aspx>

¹¹Many of these are cited in Schelling (1984).

- Not keeping alcohol in the house;
- Brushing one’s teeth earlier in the evening, in order to avoid late night snacking;¹²
- Drinking a lot of water before going to bed in order to make oneself get out of bed as soon as they wake up in the morning;¹³
- Going to a motel room for the day or taking longer than necessary train ride in order to get writing done;
- Leaving one’s laptop or papers at the office, so that home time is not spent working;
- Drinking through a straw in order to avoid gulping; and
- Buying long term gym memberships rather than paying by the day.¹⁴

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 and Pesendorfer (2001); and the set of dual self theories studied by Thaler and Shefrin (1981) and Fudenberg and 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.¹⁵ We also discuss extensions of these theories

¹²<http://freakonomics.blogs.nytimes.com/2010/01/21/starting-a-new-phase-of-my-career/>

¹³We owe this example to Bart Simpson in the Simpson’s third christmas episode “Miracle on Evergreen Terrace.” Lisa Simpson claims that Native Americans also used this commitment device - we cannot confirm this conjecture.

¹⁴This phenomenon has been studied in DellaVigna and 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.

¹⁵Preferences that explain a demand for commitment include: (1) preferences that change over time such as the quasi-hyperbolic model discussed below or the model of Banerjee and Mullainathan (2009); (2) preferences that depend on the set of choices such as Gul and Pesendorfer (2001), Dekel et al. (2009), Sarver (2008) and Noor (2007); (3) preferences that breach the assumption of independence of irrelevant alternatives (for example, the optimal expectations model of Brunnermeier and Parker (2005) breaches the irrelevance of independent alterantives, as shown by Spiegler (2008)

which 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 $t = \{0, 1, \dots\}$ 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 (DU) 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 individual with preferences that are time separable and additive assesses the utility from consumption over time $\{c_t, c_{t+1}, \dots, c_T\}$ according to

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

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 .¹⁶ 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 future selves.

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

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

and can generate a demand for commitment); or (4) preferences that are determined in part by past decisions such as Koszegi and Rabin (2009).

¹⁶Formally 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.

where $\beta \leq 1$. Laibson argued that this model captures the essence of hyperbolic discounting favored by psychological evidence,¹⁷ but retains the tractability of the DU model. This model has come to be called the quasi-hyperbolic discounting model.

3.1.1 Sophistication and Naïvete

The quasi-hyperbolic model leaves open the question of whether or not the decision maker knows his or her future preferences. O'Donoghue and Rabin formalize the distinction and show that it can have an important impact on behavior.¹⁸ They distinguish between two extreme types: sophisticates who know their own β , and naïfs, 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 naïvete. 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 under-commit their future selves making commitments that they do not follow through on. It is, therefore, possible to exploit partially naïve agents by charging them for the chance to under-commit.¹⁹ We will return to this general theme when we discuss market provision of commitment, since this self-awareness issue is critical to the question of take-up of commitment contracts.

3.2 Choice-Set-Dependent Utility

The main alternative to the quasi-hyperbolic model is the temptation preference model of Gul and 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

¹⁷For a review, see Ainslie (1992)

¹⁸For example O'Donoghue and Rabin (1999). A second formulation of naïvete is due to Eliaz and Spiegler (2006). The two formulations often agree but may lead to different welfare assessments - see for example Heidhues and Koszegi (2008).

¹⁹See DellaVigna and Malmendier (2004), Gottlieb (2008), Eliaz and Spiegler (2006) and Heidhues and Koszegi (2008) for more in depth discussion of this and related issues.

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 will 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's 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 (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 (1) then says that the decision maker weighs up the commitment and temptation utility in making her decision. If $u(f) + v(f) > u(s) + v(s)$ then she will choose the fish and her utility will be given by $u(f) - (v(s) - v(f))$. That is, her utility is the commitment value of the fish less 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 on the other hand the agent gives in to temptation and orders the steak, her utility is given by $u(s)$, and again we see that $u(f) > u(s)$ so 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 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 $fish \succ_1 steak$ while the hyperbolic model cannot.²⁰

²⁰ \succ_t are period t preferences.

3.3 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 , while the short run selves have preferences given by v . The long run self can exert influence on the short run self in order to change choices, but must pay a mental cost that is equal to $v(a) - \max_{c \in B} v(c)$ in order 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.²¹ 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 (Fudenberg and Levine (2006) and Thaler and Shefrin (1981)), face a distorted utility function (Benhabib and Bisin (2005)), or be addicted (Bernheim and 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 (Fudenberg and Levine (2006) and Thaler and Shefrin (1981)), have a limited ability to keep decisions consistent with a preset plan (Benhabib and Bisin (2005)) or face a stochastic process which determines whether the long run or short run self has complete control (Bernheim and Rangel (2004)).

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

3.4 Income, Poverty and Commitment

Much of the empirical work discussed in this article 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 and Shafir (2009)). Banerjee and 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

²¹Fudenberg and Levine (2006) and Benabou and Pycia (2002) formalize this claim. See also Loewenstein and O'Donoghue (2004) for a discussion of the relationship between the quasi-hyperbolic model and dual self models.

goods (z) and non-temptation 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 two agent, however, decides according to the utility function $u(x_2) + v(z_2)$. Thus, the time 2 agent values z_2 , while the time 1 agent does not. 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 and 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.

3.5 Soft Commitments

While there has been much discussion of concepts such as mental accounting, which would allow for soft commitment (for example Thaler (1985)), there is a dearth of papers which formalize these intuitions and their welfare and behavioral implications. Since 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 difficulties.

A promising approach can be found in Benabou and Tirole (2004). The authors develop a model of an agent who sets unenforced rules for himself, 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 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 his goals when in the absence of such a “rule” he 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. While this model explains personal rules, it does little to explain why formalization, for example labeling an account “the children’s college fund”, can be an effective commitment device.

Alternatively the cue based theory of Bernheim and Rangel might explain the success of some kinds of soft commitment. In Bernheim and Rangel’s 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 or cash or food altogether. A similar model could help explain the effectiveness of soft commitments. By making plans which avoid the complementary good an agent can avoid consuming the temptation good, but this is clearly a “soft” commitment, in that it is likely easy to succumb to some temptation, albeit in a different environment (e.g., shopping online rather than 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 which studies more formal, market supplied, commitment contracts and consider their effectiveness in changing behavior and outcomes.

4.1 Lab Evidence

Read et al. (1999) study simultaneous and sequential choice of virtue (long term benefits) and vice (short term benefits).²² An individual is said to make a simultaneous choice if she decides *now* what she will consume in the future, and to make a

²²Their definition of a virtue is similar to Dellavigna and Malmendier’s (2004) definition of an investment good.

sequential choice if 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 since subjects were not given the option to choose whether to select movies simultaneously or sequentially.

Trope and 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* in order 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. First, 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 and 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 “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.

Ariely and Wertenbroch (2002) studied the choices and performance of students

who had to undertake three tasks in twenty one 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 and Wertenbroch find that students in the free choice group chose to set deadlines on themselves that were before the last day possible. Further, 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 naïvete, temptation over menus, or as Ariely and 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. While similar issues arise regardless of income, evaluations tend to focus on the poor for reasons discussed earlier – a lack of commitment is more consequential for the poor (Mullainathan and Shafir 2009) and temptation spending may represent a larger share of total expenditures (Banerjee and Mullainathan (2009)).

Besley (1995) finds that local savings opportunities in West Africa offer negative interest on savings - in other words, that some people are willing to pay to have money taken out of their hands. While 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 in order 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 prefer-

ence reversals consistent with quasi-hyperbolic discounting.²³ 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 due to 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, Rotating Savings and Credit Associations (ROSCAs) may serve the role of commitment device. ROSCAs function as group savings mechanisms, where 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 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 documents 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 and Baland (2002)).

Ambec and 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 and 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 and Treich (2007) does) 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

²³See the discussion in section 5 for more details of the exact tests used and other studies which use similar tests.

key insight is that ROSCAs have two commitment features that a time-inconsistent saver could not access on 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 the agent to save more regularly than she would on her own. Because the agents are sophisticated, they know that their participation in future ROSCAs will be welfare enhancing, and so they avoid renegeing 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, the agent knows she has a meaningful commitment to continue saving through the ROSCA even when she is 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 renegeed 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, while ROSCAs in rural areas are more likely to be random. Citing data from Gugerty (2007) and from Anderson and Baland (2002), Basu argues that there is at least limited evidence consistent with that prediction.

Kast and Pomerantz (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 micro credit 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.²⁴ 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 signalling 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

²⁴Note 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 due to the negative arbitrage in which they were encouraged to engage in by their peers.

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 micro borrowers 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 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 which provide commitment devices, finding commitment products that clearly satisfy the definition we outlined in our introduction is more difficult. Here we will focus on four implementations that come closest to the pure commitment device as we defined. The examples here include a retirement savings device in the USA, a charitable giving device in Sweden, a smoking cessation contract in the Philippines, and a fertilizer coupon program in Kenya.

Benartzi and Thaler (2004) designed an employer based saving plan – Save More TomorrowTM (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 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 and 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. 78% of those offered the program enrolled, indicating strong demand for the commitment, and after 4 pay raises, 80% of those individuals were still enrolled, indicating that a soft commitment was sufficient to ensure their continued participation. Benartzi and

Thaler (2004) also argue that SMarT increased savings rates, the most compelling evidence for this comes from the third employer where SMarT was only offered to one division leading to a 1.5% increase in savings rates while savings at the other divisions stayed constant.

The success of SMarT leads us to ask which of its features are most important? While there are many dimensions along which the plan could be altered, two areas seem to us to be key. First, while 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 increase retention. If that is true, it is interesting to ask about the optimal penalty level. Second, we wonder what is the optimal rate of savings increase and whether the rate of increase should be a choice. Again we imagine that there is a tradeoff 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 increase 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 accounts discussed in Ashraf et al. (2006b). They worked with a rural bank in the Philippines to offer 700 individuals an option of opening a commitment savings account. The SEED (Save, Earn, Enjoy Deposits) program offered the choice of two commitment features: either (1) a time-based maturity, in which the account balance would become available only at a future date, or (2) 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%) takeup. Importantly, the SEED accounts were effective in increasing savings. Altogether, those who were offered accounts increased their savings held at the bank by about 80% compared to 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. While 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 in order to gain access to an account that was merely labeled as being about achieving

their goals.

Managing and overcoming addictions is another area where commitment devices may be important. Bernheim and 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 himself for smoking. Some addicts also make use of hard commitments. Alcoholics are also known to use certain metabolism-inhibiting drugs, such as Antabuse (a.k.a. Disulfiram), which temporarily modify the body processes that metabolize alcohol and produce a "highly unpleasant 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; since the addicts do not derive any utility from this current commitment, this behavior seems to be unambiguously a commitment device. While 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 and Kleber (2008)).

A commitment device to stop smoking (CARES) was tested by Giné et al. (Forthcoming). In this study a Philippine bank offered smokers an opportunity to open a CARES 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 (\$US11) pesos after 6 months, some 535 pesos more than the minimum balance. Most importantly smokers randomly offered CARES were 3 percentage points more likely to pass the 6-month test than the control group and these results held up in a 12-month follow-up study (6 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 naïvete, 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 under investment. 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%. Further, they find that 47% of farmers choose to receive the discount earlier, a result which 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 naïvete. 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 pre-purchased 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 and Thaler (1989), Loewenstein and Prelec (1992), Kirby and Herrnstein (1995), Ainslie (1992) and Benzion et al. (1989) cite early evidence to this effect.

To take one an example, the subjects in Thaler (1991) – students from the University of Oregon – were told that they had won a monetary prize which 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 3 months, \$60 in one year and \$100 in three years. The first comparison implies a three month discount rate of 0.5 while the third implies a three month discount rate of 0.85. These results reveal time inconsistency, but the data is 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.²⁵ 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 and 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.²⁶ 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.²⁷

It will be important to understand when evidence of time inconsistencies does, and does not, imply a demand for commitment. While 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 commitment devices, in particular ones demanded by those who demonstrate inconsistent preferences through surveys, none of these arguments can be a complete explanation.

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

²⁵Nearly 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, in Ameriks et al. (2007).

²⁶Intriguingly, Meier and 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.

²⁷For more recent evidence consistent with preference reversals see Tanaka et al. (Forthcoming)

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, while today and tomorrow are not similar. Such an explanation would not imply demand for a commitment device.

The second of these four arguments points out a risk confound. Keren and Roelofsma (1995), Andersen et al. (2008) and Fernandez-Villaverde and 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 e.g. she knows if she is hungry. Therefore when deciding between one apple today and two tomorrow the agent is choosing between a certain outcome (eating, given her hunger at this moment) and an uncertain outcome (eating tomorrow, in an uncertain state of hunger), while her choices for one year from now are both made under equally uncertain conditions. Fernandez-Villaverde and Mukherji (2002) show that the differing risks is sufficient to drive preference reversals, without implying a demand for commitment.

The third of these four arguments puts forward an alternative functional form that can accommodate the typical answers to time discounting questions. Benhabib et al. (2006) elicit 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 hyperbolic, quasi-hyperbolic, GP, etc. agents 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 or not preference reversals imply a demand for commitment - may hinge on several factors: (1) is an agent facing a choice where, as in Rubinstein (2003), "similarity" based decision making is appealing?; (2) is the risk confound significant enough, as in Fernandez-Villaverde and Mukherji (2002), to sway an agent's choice?; (3) 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?; and (4) is this

agent’s decision, as in Mulligan (1996), best modelled 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 earlier 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 which correlate the short term discount rates of 4 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.²⁸ While 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 2 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.²⁹ The key point is stated by O’Donoghue and Rabin (2006): “Without [present bias], economists have no coherent model of short-term impatience.”³⁰ Hence Mischel’s study of these

²⁸The reactions of the children to these experiments is worth a trip to youtube (search “Mischel marshmallow”).

²⁹The fact that the agent chooses one marshmallow over two in an hour implies $\delta \leq \frac{1}{2}$. Therefore over twenty-four hours the comparison is 1 versus $\delta^{24}x$. This implies $x \geq \frac{1}{(1/2)^{24}} = 16,777,216$

³⁰Interestingly, 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 the children are not sophisticates with regard to their temptation, since an agent that knows that they will eventually give in, should give in immediately. This sort of behavior is also reported by Skiba and Tobacman (2008) in relation to pay day loans. They note that many defaulting borrowers have already paid

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 ten 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 in which the subject has won 10 dinners to be used over the next 2 years. The subjects are first asked how they would *ideally* spread the dinners over the next two year. Next they are asked how they would *actually* spread the dinners. Those who expect to use the dinners earlier than ideal are seen as being tempted by current consumption, 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 over-consumption, and that 18.6% of the sample is 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 over delayer 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.

out many pay day loans before their default, and argue that the borrowers should therefore be modelled as partially naive agents.

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.³¹ 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 (fMRI) to measure brain activity in subjects making intertemporal decisions. They show that the limbic system (often associated with affective decisions making) is activated preferentially by decisions involving immediate reward, while decisions involving future rewards activate only the lateral prefrontal cortex (often associated with control and long term planning). McClure et al. 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.”³² 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 and Fedorikhin (1999). They randomly assigned subjects to remember a 2 or 7 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. 59 percent of those in the 7 digit treatment chose the cake, while only 37 per cent of those in the 2 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 when it is attempting to do other things.³³ 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 (for example not looking at words on the bottom of a TV 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

³¹It is not clear how we decide that the controlling part of the brain should be the part that we use for welfare comparisons.

³²At 506.

³³Strictly speaking, this explanation requires that cognition and self control centers are the same.

the first activity's self control demands. In an interesting recent study in this vein, Vohs and Faber (2007) ask subjects to write an essay about anything that came into their head, and told half of the participants not to think about a white bear while doing so. Everyone was then offered the opportunity to spend ten dollars on impulse consumption items (e.g., gum, playing cards, etc.) 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 while the control group spent only \$1.21 ($p < 0.001$). It would be interesting to see an experiment which combines the two approaches from McClure et al. (2004) and Shiv and Fedorikhin (1999), simultaneously looking at limited self control and the area of the brain that is working.³⁴

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 "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 preferences reversals. The Shiv and Fedorikhin (1999) suggests 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

³⁴Work in this direction has already begun. For example, Knoch and Fehr (2007) use rTMS to disable particular parts of the brain and then consider the ability of the subject to resist temptations. They find evidence which they believe implies that the right prefrontal cortex is responsible for our capacity to resist temptation.

the chocolate cake? It would be interesting to see whether such a measure correlate with the use of any kind of commitment device.

6 Market Provision of Commitment Devices

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

6.1 Tradeoff 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 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 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 saving requirement (i.e. a contract which requires the agent to save at least \$ x per period) is always part of the optimal contract. The resulting contract, however, is not efficient as the agents under consumes (over commits) at some shocks. This result could easily be extended suggesting that some people will be harmed *ex post* by an optimal commitment contract. Bond (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 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 imprecise relationship between consumption and ones 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 and 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 good will be efficiently provided. Leisure goods are goods with immediate benefit and long-term cost (e.g., fatty foods, cigarettes for some, etc.) while 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 subsidised 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 over consume. The intuition, as above, is that other firms can always undercut the high prices charged by the counterparty to the commitment contract. Take 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.³⁵ This market response is consistent with credit card contracts that subsidise 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, cigarettes, alcohol, soda etc., 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 and Rabin (2003)).³⁶ Gruber and Mullainathan (2005) provide empirical evidence

³⁵There is still scope for partially naïve consumers to take up commitment savings contracts.

³⁶see also Krusell et al. (2002); Krusell and Jr. (2003) for a discussion of optimal taxation where not only consumers, but also the government lack commitment.

that people are happier in states with higher cigarette taxes, and that this is true even amongst 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 put money in a bank account, an agent commits to have a certain level of credit available at a certain time. stickK.com is a website that provides such contracts to the general public. 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.³⁷ 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 non-linear combination of each, but this quickly becomes over complicated 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 agents 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 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 naive individuals may also be inefficiently served. Naive individuals are unlikely to demand commitment contracts, implying some sort of inter-

³⁷Disclosure: Karlan is President and owns equity in stickK.com.

vention is necessary to allow for commitment. Further, partially naive 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 naive 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 and Malmendier (2004) consider the issue in a model in which firms know if an agent is naive while Eliaz and Spiegler (2006) study the general case in which firms do not know an agent’s type. Both studies show that while sophisticates are served efficiently, partially naive agents are exploited by the market. Heidhues and Koszegi (2008) apply a similar model to the credit market and show that the presence of partially naive consumers implies contracts with high penalty fees for late repayment, as we see in reality. Heidhues and 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 and Laibson (2006) and Bertrand and 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 Runners Dilemma from above. Recall that Rachel signed a contract requiring her to pay \$1000 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 or not Rachel’s decision is welfare improving turns 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 ($B_1 + 1000 > C_1$), and time 0 Rachel is correct in the belief that running is beneficial ($B_0 > C_0 \Rightarrow B > C$). On the other hand, 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 $B_1 + 1000 < C_1$. In this case, Rachel will not run and will forfeit the money - probably a welfare reducing situation.

We will return to this issue below under the heading of naïvete. 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, while 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 non-running 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 is 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 and 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 is 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 Kőszegi and 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 is usually used by economists.³⁸ For example, in the quasi-hyperbolic model discussed above it is often argued that the preferences of the time-zero 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 is the “hard question” 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

³⁸An interesting recent contribution by Bernheim and 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.

“choose the planner/do-gooder self” when making welfare statements.

7.2 Conclusions

This review has attempted to demonstrate two theoretical challenges and three empirical observations. On the theory side, we have argued 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 have put forward evidence from the laboratory and field on the demand for commitment devices, 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 naive, and if naive 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 in order 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. (Forthcoming) finds long term behavioral changes occurred, as do Charness and Gneezy (2009) for incentives to exercise, but will such long term behavioral changes occur in other domains, such as weight loss or savings behavior?

References

- Ainslie, George, *Picoeconomics: The Interaction of Successive Motivational States within the Person* 1992.
- Amador, M., I. Werning, and G.M. Angeletos, “Commitment vs. flexibility,” *Econometrica*, 2006, pp. 365–396.
- Ambec, S. and N. Treich, “Roscas as financial agreements to cope with self-control problems,” *Journal of development economics*, 2007, 82 (1), 120–137.
- Ameriks, J., A. Caplin, J. Leahy, and T. Tyler, “Measuring Self-Control Problems,” *American Economic Review*, 2007, 97 (3), 966–972.

- Andersen, S., G.W. Harrison, M.I. Lau, and E.E. Rutstroem**, “Eliciting Risk and Time Preferences,” *Econometrica*, 2008, 76 (3), 583–618.
- Anderson, Siwan and Jean-Marie Baland**, “The Economics of Roscas and Intra-household Allocation,” *Quarterly Journal of Economics*, 2002, 117 (3), 963–995.
- Ariely, D. and K. Wertenbroch**, “Procrastination, Deadlines, and Performance: Self-Control by Precommitment,” *Psychological Science*, 2002, 13 (3), 219–224.
- Ashraf, N., D. Karlan, and W. Yin**, “Deposit Collectors,” *The BE Journal of Economic Analysis & Policy*, 2006, 6 (2), 5.
- , – , and – , “Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines*,” *Quarterly Journal of Economics*, 2006, 121 (2), 635–672.
- Ayduk, O., R. Mendoza-Denton, W. Mischel, G. Downey, P.K. Peake, and M. Rodriguez**, “Regulating the Interpersonal Self: Strategic Self-Regulation for Coping with Rejection Sensitivity,” *Journal of Personality and Social Psychology*, 2000, 79 (5), 776–792.
- Banerjee, A.V. and S. Mullainathan**, “The Shape of Temptation: Implications for the Economic Lives of the Poor,” Technical Report, MIT mimeo 2009.
- Basu, Karna**, “Hyperbolic discounting and the sustainability of rotational savings arrangements,” 2008. Working Paper.
- Battaglini, M., R. Bénabou, and J. Tirole**, “Self-control in peer groups,” *Journal of Economic Theory*, 2005, 123 (2), 105–134.
- Bauer, M., J. Chytilová, and J. Morduch**, “Behavioral foundations of microcredit: Experimental and survey evidence from rural India,” *Financial Access Initiative Research Paper*, 2008.
- Baumeister, R.F., T.F. Heatherton, and D.M. Tice**, *Losing control: How and Why People Fail at Self-Regulation*, Academic Press, 1994.
- Benabou, R. and J. Tirole**, “Willpower and personal rules,” *Journal of Political Economy*, 2004, 112 (4), 848–886.
- and **M. Pycia**, “Dynamic Inconsistency and Self-Control: a Planner-Doer Interpretation,” *Economics Letters*, 2002, 77 (3), 419–424.

- Benartzi, S. and R. Thaler**, “Save more tomorrow: Using behavioral economics to increase employee saving,” *Journal of Political Economy*, 2004, 112 (1), 164–187.
- Benhabib, J., A. Bisin, and A. Schotter**, “Present-Bias, Quasi-Hyperbolic Discounting, and Fixed Costs,” 2006. New York University.
- **and** –, “Modeling Internal Commitment Mechanisms and Self-Control: A Neuroeconomics Approach to Consumption–Saving Decisions,” *Games and Economic Behavior*, 2005, 52 (2), 460–492.
- Benzion, U., A. Rapoport, and J. Yagil**, “Discount rates inferred from decisions: An experimental study,” *Management Science*, 1989, pp. 270–284.
- Bernheim, B.D. and A. Rangel**, “Addiction and cue-triggered decision processes,” *American Economic Review*, 2004, 94 (5), 1558–1590.
- **and** –, “Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics*,” *Quarterly Journal of Economics*, 2009, 124 (1), 51–104.
- Bertrand, M. and A. Morse**, “Information Disclosure, Cognitive Biases and Payday Borrowing,” 2009. Working Paper.
- Besley, Timothy**, “Saving, Credit and Insurance,” in Jere Behrman and T.N Srinivasan, eds., *Handbook of Development Economics*, Elsevier, 1995.
- Bond, G.S.J.P.**, “Commitment contracts* Preliminary and incomplete,” 2009.
- Breman, A.**, “Give More Tomorrow: Two Field Experiments on Altruism and Intertemporal Choices,” *Unpublished Manuscript*, 2009.
- Brunnermeier, M.K. and J.A. Parker**, “Optimal expectations,” *American Economic Review*, 2005, 95 (4), 1092–1118.
- Charness, G. and U. Gneezy**, “INCENTIVES TO EXERCISE,” *Econometrica*, 2009, 77 (3), 909–931.
- Dekel, E., B.L. Lipman, and A. Rustichini**, “Temptation-driven preferences,” *Review of Economic Studies*, 2009, 76 (3), 937–971.
- DellaVigna, S. and U. Malmendier**, “Contract Design And Self-Control: Theory And Evidence*,” *Quarterly Journal of Economics*, 2004, 119 (2), 353–402.

- Duflo, E., M. Kremer, and J. Robinson**, “Nudging Farmers to Use Fertilizer: Evidence from Kenya,” 2009.
- Eliasz, K. and Ran Spiegler**, “Contracting with Diversely Naive Agents,” *The Review of Economic Studies*, 2006, 73 (3), 689–714.
- Fernandez-Villaverde, J. and A. Mukherji**, “Can We Really Observe Hyperbolic Discounting?,” 2002.
- Fudenberg, Drew and David K. Levine**, “A Dual-Self Model of Impulse Control,” *The American Economic Review*, 2006, 96 (5), 1449–1476.
- Gabaix, X. and D. Laibson**, “Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets*,” *Quarterly Journal of Economics*, 2006, 121 (2), 505–540.
- Galanter, M. and H.D. Kleber**, *The American Psychiatric Publishing textbook of substance abuse treatment*, American Psychiatric Pub, 2008.
- Giné, X., D. Karlan, and J. Zinman**, “Put your Money where your butt is: A commitment savings account for smoking cessation,” *American Economic Journal: Applied Economics*, Forthcoming.
- Glimcher, P.W., J. Kable, and K. Louie**, “Neuroeconomic Studies of Impulsivity: Now or Just as Soon as Possible?,” *American Economic Review*, 2007, 97 (2), 142–147.
- Goldstein, A.**, *Addiction: From biology to drug policy*, Oxford University Press, USA, 2001.
- Gottlieb, D.**, “Competition over Time-Inconsistent Consumers,” *Journal of Public Economic Theory*, 2008, 10 (4), 673–684.
- Gruber, J.H. and S. Mullainathan**, “Do Cigarette Taxes Make Smokers Happier,” *The BE Journal of Economic Analysis & Policy*, 2005, 5 (1), 4.
- Gugerty, M.K.**, “You can’t save alone: Commitment in rotating savings and credit associations in Kenya,” *Economic Development and cultural change*, 2007, 55 (2), 251–282.
- Gul, F. and W. Pesendorfer**, “The case for mindless economics,” *The Foundations of Positive and Normative Economics: A Handbook*, 2008, pp. 3–42.

- Gul, Faruk and Wolfgang Pesendorfer**, “Temptation and Self-Control,” *Econometrica*, 2001, 69 (6), 1403–1435.
- and –, “Self-Control and the Theory of Consumption,” *Econometrica*, 2004, 72 (1), 119–158.
- Heidhues, P. and B. Koszegi**, “Exploiting Naivete about Self-Control in the Credit Market,” *University of California, Berkeley*, 2008.
- Hughes, J.R., J. Keely, and S. Naud**, “Shape of the relapse curve and long-term abstinence among untreated smokers.,” *Addiction*, 2004, 99 (1), 29.
- Kahneman, Daniel, Peter Wakker, and Rakesh Sarin**, “Back to Bentham? Explorations of Experienced Utility.,” *Quarterly Journal of Economics*, 1997, 112 (2), 375–406.
- Kast, Felipe and Dina Pomerantz**, “Peers as a Saving Commitment Device: Evidence from a Field Experiment Among Low-Income Micro-Entrepreneurs,” 2009. Working Paper.
- Keren, G. and P. Roelofsma**, “Immediacy and certainty in intertemporal choice,” *Organizational Behavior and Human Decision Processes*, 1995, 63 (3), 287–297.
- Kirby, K.N. and R.J. Herrnstein**, “Preference reversals due to myopic discounting of delayed reward,” *Psychological Science*, 1995, 6 (2), 83–89.
- Knoch, D. and E. Fehr**, “The Right Prefrontal Cortex and Self-Control,” *Annals of the New York Academy of Sciences*, 2007, 1104 (1 Reward and Decision Making in Corticobasal Ganglia Networks), 123–134.
- Kőszegi, B. and M. Rabin**, “Choices, situations, and happiness,” *Journal of Public Economics*, 2008, 92 (8-9), 1821–1832.
- Koszegi, B. and M. Rabin**, “Reference-dependent consumption plans,” *American Economic Review*, 2009, 99 (3), 909–936.
- Krusell, P., B. Kuruşçu, and A.A. Smith Jr**, “Equilibrium Welfare and Government Policy with Quasi-geometric Discounting,” *Journal of Economic Theory*, 2002, 105, 42–72.
- Krusell, Per and Anthony A. Smith Jr.**, “Consumption-Savings Decisions with Quasi-Geometric Discounting,” *Econometrica*, 2003, 71 (1), 365–375.

- Laibson, David**, “Golden Eggs and Hyperbolic Discounting,” *Quarterly Journal of Economics*, 477 1997, 113 (2), 443.
- Loewenstein, G. and D. Prelec**, “Anomalies in intertemporal choice: Evidence and an interpretation,” *The Quarterly Journal of Economics*, 1992, pp. 573–597.
- **and R. Thaler**, “Anomalies: Intertemporal Choice,” *Journal of Economic Perspectives*, 1989, 3 (4), 181–93.
- **and T. O’Donoghue**, “Animal Spirits: Affective and Deliberative Processes in Economic Behavior,” 2004.
- McClure, S.M., D.I. Laibson, G. Loewenstein, and J.D. Cohen**, “Separate neural systems value immediate and delayed monetary rewards,” 2004.
- Meier, Stephan and Charles Sprenger**, “Present-Biased Preferences and Credit Card Borrowing,” *American Economic Journal*, 2010, 2 (1), 193–210.
- Mischel, W., Y. Shoda, and MI Rodriguez**, “Delay of gratification in children,” *Science*, 1989, 244 (4907), 933–938.
- Mulainathan, Sendhil and Eldar Shafir**, *Savings Policy and Decision-Making in Low-Income Households*, Russell Sage Foundation Publications, 2009.
- Mulligan, C.**, “A logical economist’s argument against hyperbolic discounting,” *University of Chicago mimeo*, February, 1996.
- Noor, J.**, “Commitment and self-control,” *Journal of Economic Theory*, 2007, 135 (1), 1–34.
- O’Donoghue, T. and M. Rabin**, “Doing it now or later,” *American Economic Review*, 1999, pp. 103–124.
- **and –**, “Studying optimal paternalism, illustrated by a model of sin taxes,” *American Economic Review*, 2003, 93 (2), 186–191.
- **and –**, “Incentives and Self-Control,” in “Advances in Economics and Econometrics: Theory and Applications: Ninth World Congress” Cambridge University Press 2006, p. 215.
- Phelps, E. S. and R. A. Pollak**, “On Second-Best National Saving and Game-Equilibrium Growth,” *The Review of Economic Studies*, 1968, 35 (2), 185–199.

- Plateau, Jean-Philippe**, *Institutions, Social Norms and Economic Development*, Harwood Academic, 2001.
- Read, D., G. Loewenstein, and S. Kalyanaraman**, “Mixing Virtue and Vice: Combining the Immediacy Effect and the Diversification Heuristic,” *Journal of Behavioral Decision Making*, 1999, 12, 257–273.
- Rubinstein, A.**, ““ Economics and psychology”? The case of hyperbolic discounting,” *International Economic Review*, 2003, pp. 1207–1216.
- Rutherford, S.**, *The poor and their money*, Oxford university press New Delhi, 2000.
- Sarver, T.**, “Anticipating regret: Why fewer options may be better,” *Econometrica*, 2008, 76 (2), 263–305.
- Schelling, T.C.**, “Self-command in practice, in policy, and in a theory of rational choice,” *The American Economic Review*, 1984, pp. 1–11.
- Shiv, B. and A. Fedorikhin**, “Heart and mind in conflict: The interplay of affect and cognition in consumer decision making,” *Journal of Consumer Research*, 1999, 26 (3), 278–292.
- Shoda, Y., W. Mischel, and P.K. Peake**, “Predicting adolescent cognitive and self-regulatory competencies from preschool delay of gratification: Identifying diagnostic conditions,” *Developmental Psychology*, 1990, 26 (6), 978–986.
- Skiba, P.M. and J. Tobacman**, “Payday Loans, Uncertainty and Discounting: Explaining Patterns of Borrowing, Repayment, and Default,” 2008.
- Spiegler, Ran**, “On Two Points of View Regarding Revealed Preference and Behavioral Economics,” *The Foundations of Positive and Normative Economics: A Handbook*, 2008, p. 95.
- Strotz, R.H.**, “Myopia and Inconsistency in Dynamic Utility Maximization,” *Review of Economic Studies*, 1956, 23 (3), 165–180.
- Tanaka, T., C. Camerer, and Q. Nguyen**, “Risk and time preferences: Experimental and household survey data from Vietnam,” *American Economic Review*, Forthcoming.
- Thaler, R.**, “Mental accounting and consumer choice,” *Marketing science*, 1985, 4 (3), 199–214.

- Thaler, R.H.**, “Some Empirical Evidence on Dynamic Inconsistency,” *Quasi rational economics*, 1991, *1*, 127–136.
- **and HM Shefrin**, “An economic theory of self-control,” *The Journal of Political Economy*, 1981, pp. 392–406.
- Trope, Y. and A. Fishbach**, “Counteractive self-control in overcoming temptation,” *Journal of Personality and Social Psychology*, 2000, *79* (4), 493–506.
- Vohs, K.D. and R.J. Faber**, “Spent resources: Self-regulatory resource availability affects impulse buying,” *Journal of Consumer Research*, 2007, *33* (4), 537–547.