10-1-2009

Commitment Contracts

Gharad Bryan

Dean Karlan

Scott Nelson

ECONOMIC GROWTH CENTER
YALE UNIVERSITY
P.O. Box 208629
New Haven, CT 06520-8269
http://www.econ.yale.edu/~egcenter/

CENTER DISCUSSION PAPER NO. 980

Commitment Contracts

Gharad Bryan
Yale University

Dean Karlan
Yale University
Innovations for Poverty Action
Jameel Poverty Action Lab

Scott Nelson
Yale University

October 2009

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 and Richard Mansfield for helpful advice. Karlan is also grateful to the National Science Foundation and Sloan Foundation for support. All errors remain our own.

Commitment Devices

Gharad, Bryan, Dean Karlan and Scott Nelson

September 9, 2009

Abstract

We review the theoretical and empirical literature on commitment devices. A commitment device is any arrangement, entered into by an individual, with the aim of making it easier to fulfill his or her own future plans. We argue that there is growing empirical evidence supporting the proposition that people demand commitment devices and that these devices can change behavior. We highlight the importance of further research exploring soft commitments – those involving only psychological costs – and the welfare consequences of hard commitments – those involving actual costs – especially in the presence of bounded rationality.

Keywords: Consumer/household economics, institutional and behavioral economics

JEL codes: D03, D14

1 Introduction

People set goals. They promise to smoke less, to save more, to drink less, to work more, to work less, to study more, to eat less, to exercise more. They set goals with varying levels of formality. They promise themselves, make vows to their spouses, and lay bets with friends. These same people regularly fail to meet their goals. They quit smoking each year, try every new diet, 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 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 fulfil a plan for future behaviour that would otherwise be difficult due to personal conflict. We will exclude actions that are taken with a strategic motive, or which accrue significant current benefits. So, while painting your nails with Control-It Nail Biting Treatment® is a commitment device, buying pink nail polish for use over the next month is probably not. And while buying a season pass to the opera to help yourself stick with your plan to become more cultured may be a commitment device, committing to go to the opera in order to stop your husband from going to a football game is not.

We will call commitment devices that provide for harsh penalties for failure, or rewards for success, hard commitments. And we will call any device that does not have large economic consequences a soft commitment. This determination is not binary, and commitments can be placed on a spectrum from hard to soft, with the hardest commitment imposing such high costs as to remove completely a choice from a choice set and the softest commitments incurring only the smallest of psychic costs for deviation. When we make the distinction between hard and soft it is often where a hard commitment involves a monetary penalty or bonus and a soft commitment involves merely a psychological cost. This reflects the prevalent view in economics that psychological costs are largely irrelevant. An example of a hard commitment would be a commitment saving account on which interest is forfeited if a monthly installment is not made. A soft commitment would be a separate saving account labeled "Hawaiian Vacation".

Although this review paper focuses on the existing theory and empirics of commitment devices, we also discuss areas for further research. In particular we argue that there is insufficient work on the welfare implications of commitment and the use and importance of soft commitments - that is, commitments which involve only low or psychic costs.

The paper is structured as follows. In section 2 we provide a more detailed definition of commitment. Section 3 discusses circumstantial evidence which suggests that people use commitment devices and also often fail to reach their goals. In section 4 we review economic theories that can explain a demand for commitment, highlighting the multiplicity of explanations and some of the issues for welfare analysis. We also briefly review theories of bounded rationality which allow people to fail in their goals, and point out that there is very little satisfactory theory explaining soft commitments although empirically they seem more common than hard commitments. In section 5 we review evidence supporting a demand for commitment. In section 6 we look at evidence of the use of commitment in lab and field settings. Finally section 7 concludes and discusses some other active areas of research related to commitment.

2 Definition

In planning this paper, we had expected the introduction to include a concise one line definition of a commitment device. This, however, turned out to be more difficult than we thought. Many transactions in every day life restrict or alter our future choices, but not for the reasons we are interested in in this paper. This section, therefore, includes a long winded definition of a commitment device. We believe that this exercise is valuable, and shows that in many cases whether an arrangement is a commitment device *per se* depends on the individual's motivation for making the arrangement, not just the terms of the arrangement. Empirically, therefore, identifying a commitment device requires examining the motivation for behavior rather than looking exclusively at choice data.

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 arrangement is primarily aimed at helping the agent fulfill plans for their own future behavior, where they themselves are the source of the risk of failing to fulfill the plan, and not changes in prices or other agents (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 they are intended to help an agent stick to a current plan or goal which

they would otherwise find difficult to fulfill or achieve. This naturally is a difficult clause to manage given that most of economics focuses on observable choices, not motivations; it, however, seems to be necessary.

Condition (a) also rules out arrangements which simply facilitate future consumption plans, but for purely transactional and logistical reasons (hence the clause at the end about changes in prices). For example, buying in bulk can reduce future transaction costs and thus could be done in order to lock in a certain consumption path, it also could be done merely to lower future costs of going to the store repeatedly.[1]

Condition (b) rules out a well-studied set of commitment devices which are used to influence the actions of others. Consider, for example, Dr. Strangelove's "doomsday device," 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 strong evidence of demand for commitment devices due to spousal, familial and neighbor bargaining.[2] Both self-control and others-control motivations share the common prediction, that people will demand commitment devices.

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 single-handed arrangement employed by an individual to restrict his or her own choices.[3] Note that some contracts will contain

[1]Thus buying in bulk is an interesting example of a behavior that could fit three different interpretations: buying in bulk could be done merely to save transaction costs (e.g., buying a larger package toilet paper in order to avoid returning the store in a week for more), could be done as a commitment device in order to resist temptation and fulfill a plan for future behavior (e.g., long term gym memberships rather than short term), or the reverse could be true in which an individual *avoids* buying in bulk in order to avoid ongoing consumption (e.g., buying one candy bar rather than a box), Wertenbroch (1998). This highlights the earlier point, that understanding the motivation for the arrangement is key to determining whether something is a commitment device, and that merely observing the terms of the arrangement does not provide sufficient information to determine if something is or is not a commitment device.

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

[3]Commitment 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.

commitment but will compensate the committing party in advance. These could be expected when firms benefit from locking in consumers' future behavior. These contracts violate the first clause of our definition and raise a set of issues beyond the scope of this paper, and so for our purposes, compensated contracts are neither commitment devices nor commitment contracts. However, this line is murky, in that many arrangements referred to here do generate some benefit outside the contract. Again, the distinction comes down to motive: was the purpose of the arrangement to tie one's future hands, or to receive current benefit? For example, a loan generates a current benefit, the cash from the disbursal. It also commits the borrower to save, by requiring loan payments; the motivation is key.

3 Anecdotal Evidence Suggesting a Demand For Commitment

In this section we provide some anecdotal evidence of demand for commitment, and discuss some informal arrangements and ad hoc behaviors that can be construed as commitment devices. As with most behaviors, alternative theories can also account for the behaviors we discuss here. Later in the paper we present harder and more careful evidence on commitment devices.

Smoking The majority of smokers say they want to stop, and even try to stop, but continuously fail. 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 quit attempt.[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:

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

[5] http://www.cdc.gov/media/pressrel/r061026a.htm

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

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

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]

Voting Many more Americans claim to have voted than actually did vote. For example Silver et al. (1986) cite evidence that around 30% of nonvoters reported voting in studies between 1964 and 1980. At the same time, many Americans state that they intend to vote, but fail to do so. Pew Research Centers estimated pool of likely voters, which is itself a proper subset of people who report that they "plan to vote,"[11] has a 10% false positive rate[12].

Various ad hoc Commitment Devices Other limited anecdotal evidence suggests ad hoc behavior that can be construed as commitment devices, including:[13]

- 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;
- Not keeping alcohol in the house;
- Going to a motel room for the day in order to get writing done;
- Drinking through a straw in order to avoid gulping; and
- Buying long term gym memberships rather than paying by the day.[14]

[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

[11] http://people-press.org/methodology/files/UnderstandingLikelyVoters.pdf

[12] http://people-press.org/report/12/

[13] Many of these are cited in Schelling (1984)

[14] This 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.

4 Theory

Here we review theories that help explain and conceptualize the demand for commitment devices. Studying the theory has three benefits: (1) it clarifies the assumptions on which welfare evaluations rely; (2) it allows us to model market interactions and therefore to understand both optimal and market-generated commitment contracts; and (3) it allow us to leverage our understanding of commitment to explain other interesting phenomena, such as the puzzles studied in Laibson (1997).

In this review we concentrate on the models that are usually used to analyse commitment: the quasi-hyperbolic discounting model of Laibson (1997); the temptation and self control theory of Gul and Pesendorfer (2001); and the set of dual self theories first studied by Thaler and Shefrin (1981). We also review bounded rationality in these models and discuss how it can explain failures of commitment. These are the best understood models, and the ones most often used to model commitment. There are, however, many other models that can predict a demand for commitment. For example, any model that posits (1) preferences that change over time, (2) preferences that depend on the set of choices,[15] (3) preferences that breach the assumption of independence of irrelevant alternatives,[16] or (4) preferences that are determined in part by past decisions,[17] could potentially provide a rationale for a commitment contract. Models that fall in to these categories will make predictions regarding choices, psychological processes and welfare.

In all the models that follow below, we discuss a single decision maker as if he or she were a collection of decision makers over many time periods $t = \{0, 1, \ldots\}$ and refer to the time t decision maker as the time t self. We begin our discussion with an example which we hope highlights the difficulty of making welfare statements in situations where commitments are demanded.

[15]The temptation and self control model of Gul and Pesendorfer (2001) is probably the most well known of these, however other examples include Dekel et al. (2009a), Sarver (2008) and Noor (2007) to cite just a few from a large literature

[16]For example, the optimal expectations model of Brunnermeier and Parker (2005) breaches IIA as shown by Spiegler (2008). Within this model it is simple to construct an example in which a decision maker would prefer to remove a particular option from her choice set. Intuitively this occurs because removing a particular choice will free the agent to hold beliefs which justify a preferred option.

[17]See for example, Koszegi and Rabin (2009). In this model the agent gains utility from "surprising" herself when departing from a plan. She would like to commit herself not to do this.

4.1 The Runner's Dilemma

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. This contract acts to restrict Rachel's future choices, and is not explicable using standard preferences.[18]

Our aim in this section is to find a theory of decision-making that will rationalize Rachel's choice and indicate whether it is welfare improving. A relatively general approach would be to assume that the true benefit of running is B and the true cost is C. We may then model Rachel as two selves, the time 0 or pre-running Rachel, and the time 1 or running Rachel. We model these Rachels as having beliefs regarding the costs and benefits. We assume that time t Rachel believes that $B = B_t$ and $C = C_t$. Further, we assume that time 0 Rachel has beliefs about what time 1 Rachel believes. In particular we assume that time 0 Rachel believes that time 1 Rachel believes $B = \tilde{B}$ and $C = \tilde{C}$. Throughout we assume that $B_1 < C_1$ so that Rachel will actually walk if no commitment 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 $\tilde{C} > \tilde{B}$. That is, time 0 Rachel believes that it is valuable to run, but that time 1 Rachel does not.

Our welfare analysis, however, 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 naivete. 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 $\tilde{B} < \tilde{C}$, 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 will have better information about the costs of running. Further, it is also possible to argue that 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

[18] A "rational" Rachel might be indifferent between having and not having the commitment contract. However, with even a small amount of uncertainty, she would not sign the contract.

basing the assessment C_0 on a biased measure.

The difficulty here 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.[19] We take no stand on the appropriate approach, but are sympathetic to a view which suggests that combining choice information with more qualitative data, such as psychological data, can be helpful.

4.2 Changing Preferences and 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 a consumption stream $\{c_t, c_{t+1}, \ldots, c_T\}$ according to

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

where u is a one period felicity 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

[19] An 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.

of consumption across any two periods, for all τ.[20] 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 introspection implied a function that overvalues current consumption. A simple and tractable formulation that allows for such preferences was provided by Phelps and Pollak (1968) and Laibson (1997):

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

where $\beta \leq 1$. Laibson argued that this model captures the essence of hyperbolic discounting favored by psychological evidence,[21] but retains the tractability of the DU model. This representation of the decision process, referred to as the quasi-hyperbolic discounting model, does not have an obvious welfare function associated with it. It suffers from the same problem as the runners dilemma above; time 0 Rachel does not agree with time 1 Rachel, because time 1 Rachel over discounts the future by β relative to what time 0 Rachel would prefer.

Two approaches have been taken to welfare in this model. First, if a policy can be shown to be Pareto improving, in the sense that all selves prefer the policy, then it must be welfare improving. This criteria, however, does not provide a complete ordering. The second is to assume, as O'Donoghue and Rabin (1999a) do, that the time-0 exponential discounting utility function

$$U(\{c_0, c_1, \ldots, c_T\}) = \sum_{k=0}^{T} \delta^k u(c_k), \tag{1}$$

represents the "correct" preferences. We might refer to these preferences as the "normative" prefernces in that we the researchers are passing judgement, and declaring them better. Using this welfare criterion, it is possible to show that a choice to take out a commitment contract must always be welfare improving.

Four points should be noted about the quasi-hyperbolic model. First, the assertion that (1) represents the correct preferences is just an assumption, and one that

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

[21] For a review, see Ainslie (1992)

is difficult to test. Second, decisions need not reflect what is best for the individual based on the normative preferences. Therefore, revealed preference no longer holds for making welfare statements. For example, an individual deciding to smoke at time t may be maximizing time t preferences but not the normative preferences. Third, quasi-hyperbolic preferences do not allow for the possibility that successful self control may be costly – if an agent does not choose a particular alternative, then it is assumed that the alternative's presence in the choice set has no impact on welfare. Fourth, because tastes change over time some assumption must be made in order to determine how the conflict is resolved. For the moment we assume that the decision maker is aware of the conflict and in this setting, it is popular to model the decision process as an extensive form game between the various selves and to look for subgame perfect Nash Equilibria.[22] This approach has the difficulty that it may lead to serious multiplicity of results, even when sensible refinements are applied (Krusell and Smith Jr (2003)).

4.3 Choice-Set-Dependent Utility

Choice-set-dependent utility models study preferences over *sets* or menus of choices, from which an actual choice will be made at a later date. The utility of the individual thus depends directly on the set from which she will eventually choose, rather than on the actual choice that she makes. This approach was first used by Kreps (1979) to model a preference for flexibility. In order to accommodate a preference for commitment, it must be the case that utility is weakly decreasing in the size of the choice set, controlling for the choice that is made.

Gul and Pesendorfer (2001, 2004a) (GP) provide a leading example of a choice set dependent utility model, which maintains the assumption of revealed preferences and also allows for self control. Traditional choice theory assumes that if A is weakly preferred to B then A will be chosen from the set $A \cup B$ and therefore the inclusion of B in this set is irrelevant. Formally

$$A \succeq B \Rightarrow A \sim A \cup B. \tag{2}$$

In situations where the alternative B is tempting, however, this need not be the case. Intuitively, if B is tempting, then its presence in the choice set $A \cup B$ implies that the decision maker will need to apply costly self control to avoid choosing it. This leads Gul and Pesendorfer to relax the above axiom, replacing it with

$$A \succeq B \Rightarrow A \succeq A \cup B \succeq B. \tag{3}$$

[22]The case in which the decision maker is not aware of his inner conflict is discussed below under the heading of naivete and bounded rationality.

This axiom allows for the usual case, but also the case in which the individual would prefer to remove B from the set of choices (i.e. $A \succ A \cup B$). This preference could be motivated by one of two possibilities. First, it may be that B would be chosen from the set $A \cup B$ (i.e. $A \cup B \sim B$) or, second, it may be that A would be chosen, but there would be a cost of this choice relative to being committed to A (i.e. $A \cup B \succ B$).

Adding the axiom (3) to the usual axioms of choice, GP derive the following representation for preferences over choice sets:

$$U(z) = \max_{b \in z} \Big(u(b) + v(b) + \delta U(z') \Big) - \max_{c \in z} v(c),^{23}$$

where z is the choice set in the current period, u and v are VnM utility functions, $\delta \leq 1$ is the discount rate and z' is the next period choice set (which will in general depend on b). When considering consumption choices for example, z is present wealth. The representation can be given the following interpretation. u represents the agents "commitment preferences". That is, if the agent were choosing from the set $\{a\}$ then she would receive utility $u(a)$. The difference $v(a) - \max_{c \in z} v(c) \geq 0$ represents the self control cost of choosing option a from a set in which there are other more tempting alternatives, therefore v can be thought of as a measure of the decision makers temptation utility. GP provide additional assumptions which imply that period 1 choice is made so as to maximize $u + v$, which is suggested by this representation of period 0 preferences.[24]

Two key characteristics distinguish the GP model from the quasi-hyperbolic model. First, the GP model is consistent with revealed preference. For example, a time 1 self that chooses to smoke is taking the preference maximizing action given the costs of self control associated with the choice set from which she must choose. Second, disutility in the GP model is a function of the choice set and thus allows for costly self control, while disutility in the hyperbolic model is strictly a function of the actions taken. For example, consider a diner deciding between steak and fish. 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, 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.[25]

[23] GP allow for lotteries over choices and therefore, we ignore this possibility here for expositional purposes.

[24] The model has been extended in several ways, for example to allow for utility to depend on multiple tempting choices rather than the most tempting (Dekel et al. (2009b)) or to allow for choice sets to be tempting (Noor (2007)). In their presentation GP use an axiom which cannot actually be tested in their framework, Noor (2009) provides choice observable axioms for the period 1 behavior of a GP agent.

[25] $\succ_t$ are period t preferences.

Researchers differ on whether revealed preference is an important feature of a model, and there is a lively debate about the way in which it should be used.[26] On the other hand, the second feature of the model may have significant behavioral implications. For example, consider an agent that lives for three periods and has no income in the third period but can save at a positive interest rate in period 2. Suppose that this individual can choose in period 1 to signup for "social security" which removes a small portion of her period 2 income and transfers it to period 3, but at a zero interest rate. We might think this would help an agent that demands commitment. However, a quasi-hyperbolic individual will not signup for social security as it will merely crowd out saving in period 2.[27] A GP decision maker, however, may sign up, as she values the smaller choice set in period 2 (Gul and Pesendorfer (2004b)).

4.4 Dual Self Models

The GP representation is 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.[28] 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

[26] See for example, the papers in Caplin and Schotter (2008)

[27] Some conditions must be met for this statement to be true. In particular the agent cannot be too myopic and must have concave utility in each period.

[28] Fudenberg 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.

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.

Dual self models, by specifying the process through which decisions are made, can help us to widen the sorts of contracts that we consider. For example, the Cue based theory of Bernheim and Rangel indicates that commitment contracts may more usefully restrict the choice of locations rather than the actual choice of consumptions. For example, if someone has trouble controlling their consumption while at a mall, it is potentially better to commit to not go to the mall, rather than committing to a lower level of consumption through high savings. This will avoid the problem of self control, but also vitiate any concerns stemming from committing to particular saving levels in the presence of an uncertain income.[29]

4.5 Sophistication, Naivete and Bounded Rationality

4.5.1 Two Formulations of Naivete

Throughout the above discussion we have implicitly assumed that a decision maker making a decision at time t is aware of the actions and utility that will accrue to future decision makers. There is, however, some evidence that agents are not fully aware of their self-control abilities, or their likely future actions. Addressing these facts requires that we incorporate bounded rationality into our models. Importantly, this addition clarifies some situations in which agents demand too much or too little commitment.

There are two popular approaches to incorporating bounded rationality. The first was introduced in a series of papers by O'Donoghue and Rabin.[30] They consider the quasi-hyperbolic model discussed above and associate with each individual three parameters $(\hat{\beta}, \beta, \delta)$. $\hat{\beta}$ is the decision maker's estimate of her own degree of dynamic inconsistency β. So an individual for whom $\hat{\beta} = \beta$ is not boundedly rational. O'Donoghue and Rabin refer to these individuals as sophisticates. On the other hand, individuals for whom $\hat{\beta} = 1$ believe that their future selves will behave optimally and do not recognize their inconsistency. O'Donoghue and Rabin refer to these

[29]This issue is discussed in more depth in Amador et al. (2006) who show that in the presence of uncertainty, minimum saving requirements are often optimal contracts.

[30]For example O'Donoghue and Rabin (1999a).

individuals as Naifs. Finally those individuals for whom $\hat{\beta} \in (\beta, 1)$ are considered to be partially naive.

Eliaz and Spiegler (2006) adopt a different approach to modeling boundedly rational or naive agents. They assume that an agent's current preferences regarding future consumption are given by g and that her future preferences will be given by f. Naivete is captured by assuming that the agent believes that with probability p her future preferences will be g and with probability $1-p$ her future preferences will be f. Thus $p = 0$ is sophistication and $p = 1$ is complete naivete.[31]

Both formulations of partial naivete allow for agents to under-commit their future selves, and thus make commitments that they do not follow through on. Thus, it is possible to exploit partially naive agents by charging them for the chance to under-commit.[32] The two formulations, however, seem to differ in their interpretation. For example, an individual who is partially naive under the O'Donoghue Rabin definition will believe with certainty that they will act in a particular way and then be surprised by their future actions. They can therefore be properly seen as making mistakes if they sign a contract that they think will commit them, but which does not. On the other hand, an Eliaz/Spiegler agent will commit to a contract while fully aware that they may not fulfill their commitment, but believing that the choice to commit is the superior one given the information that they have. This is harder to view as a mistake, even if we posit that the agent has an objectively incorrect probability distribution over her future actions.[33]

The models also differ in their behavioral and welfare implications. For example, Heidhues and Koszegi (2008) construct a model of credit card borrowing and consider the impact of naivete on welfare. They show that in the O'Donoghue/Rabin formulation there is a discrete loss in utility associated with the move from full sophistication to any naivete. On the other hand in the Eliaz and Spiegler (2006) formulation there is no such discontinuity. This implies that there are different welfare predictions from the credit market regulations they consider. The discrepancy occurs because the O'Donoghue/Rabin consumer that is very slightly naive is confident that she will carry through with her future plans - but will in fact be wrong about this. On the

[31] A similar formulation appears in Dekel et al. (2009b) in their discussion of uncertainty regarding temptation in a GP style model of choice set dependent utility. Their formulation, however, is one of uncertainty rather than bounded rationality as the agent is presumed to have a correct prior over the future utility functions.

[32] See 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.

[33] As we will see below Eliaz and Spiegler (2006) make use of a model in which there is an objective probability q that the agent will have future preference v. To the extent that $p \neq q$ the agent can be seen as objectively mistaken.

other hand, the Eliaz/Speigler consumer puts a very high probability on the action that she will actually take. Research addressing the appropriate way to mode naivete would therefore be helpful.

4.6 Soft Commitments

The theory discussed so far does not do a good job of explaining soft commitments. While there has been much discussion of concepts such as mental accounting, which would allow for soft commitment, there is a dearth of papers which formalize these intuitions and their welfare and behavioral implications. Soft commitment seems to be an area in which a better understanding of the actual decision process becomes important. For, example, it is possible to include soft commitments in the quasi-hyperbolic model, but the process does not seem compelling. Specifically, there are an infinite number of equilibria in the intrapersonal game that is usually used to model infinite time quasi-hyperbolic decisions. This suggests that soft commitments could be used as a form of equilibrium selection. This argument, however, only raises the question of why these are not seen merely as cheap talk. There are many modeling options which could be pursued and we believe the returns would be substantial.

5 Evidence for Particular Theories that Lead to Demand for Commitment Devices

In this section we review evidence from economics, psychology and neuroscience on the theories discussed above. We first discuss laboratory evidence on time discount rates, and discuss competing interpretations and alternative methods of elicitation. Then we will examine studies that compare laboratory elicitations of discount rates or similar parameters to "real world" behavior either in the short term or long term. Lastly we will discuss the neuroeconomics evidence and its implications 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 later. They were asked how much they would require to make them indifferent between receiving the prize now, or later. The subjects received 3 $\times$ giving 3 prize values and 3 waiting times. 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 the second a three month discount rate of 0.7 and the third a three month discount rate of 0.85. These results seem to imply hyperbolic or quasi-hyperbolic discounting, but are also compatible with other explanations, such as decreasing temptation over time.

Second, perhaps 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 Filippino microfinance clients exhibit preference reversals in which he client is more patient in the future time period than in the current time period. 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. These findings support some kind of break with the time consistent exponential discounting model, but can be reconciled with both the quasi-hyperbolic and GP models presented above.

However, four arguments suggest interpreting these laboratory results differently, such that they do not imply a potential demand for commitment devices. The first argument, from Rubinstein (2003) uses a boundedly rational model which focuses on cognitive and procedural aspects of choice rather than on utility or preferences. Rubinstein argues that a "similarity" based decision making process can explain evidence of preference reversals. The idea is that any choice over time can be thought of as choosing between pairs (x, t) where x is a reward and t is a time. Rubinstein

argues that the agent first looks at a choice to determine if there is similarity between the x's or t's. If there is similarity in one dimension the decision is made on the basis of which outcome is larger in the other dimension. It seems intuitive that today and tomorrow are not similar and that a year from now and a year and a day are similar. Suppose, we ask how many apples in one year make an agent indifferent to 2 apples in a year and one day, and the agent says y. Then we can argue that $y < 2$ and also that y is similar to 2. On the other hand, if we ask how many apples today makes you indifferent to 2 apples tomorrow and the agent says z, then we can argue that $z < 2$ and that z is not similar to 2. Combining these observations it seems reasonable to argue that $z < y$, and therefore we see a preference reversal. This model of the cognitive process generates the above cited laboratory results, but does not imply demand for a commitment device.

The second 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, she knows if she is hungry, how much money she has etc. 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, and Andersen et al. (2008) show how simultaneously estimating risk preferences provides better theoretical estimates of the time preference parameters, and tends to decrease the evidence for decreasing discount rates. Again, as with Rubinstein's cognitive model, this theory or confound does not predict that reversals will correlate with use of commitment device.

The third argument puts forward an alternative functional form that can accommodate the typical answers to time discounting questions but does not require hyperbolic or quasi-hyperbolic discounting. Benhabib et al. (2006) elicit discount rates for a group of NYU students. 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. Instead, each self will over consume by \$4 – hardly worth bothering to sign a commitment contract over. One obvious explanation for the findings in Benhabib et al. (2006) is that some goods are tempting and some are not. In this case, we may think that the rewards offered by Benhabib et al. (2006) – money – were not particularly tempting. The GP model can easily capture the different temptation characteristics of different goods, and so can suitable generalizations

of the hyperbolic discounting function.

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.

These arguments share the common implication that the observed reversals in time discounting should not predict a preference for commitment. Which interpretation is correct is, however, an empirical question. If the above laboratory questions predict usage of commitment devices then that lends support to the time inconsistent or GP models relative to the alternative models. Ashraf et al. (2006b) test this. 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 aimed to help clients who wanted to save more but found themselves not achieving their goals. SEED accounts offered the choice of two different withdrawal-side commitment features: either (1) a time-based maturity, in which the account balance would become available for withdrawal only at a specified future date, or (2) an amount-based maturity, in which the account balance would become available only once the account-holder reached a specified goal of total savings balance. 202 clients (28% of those offered) opened a SEED account, despite the fact that the account offered no increase in interest rate in exchange for the loss of liquidity. Ashraf et al. (2006b) report that clients who exhibit preference reversals were 12.5% points more likely to open a SEED account, the effect being driven largely by female borrowers.

5.2 Long Term Evidence on Discount Rates

Walter Mischel has a series of studies which correlate the short term discount rates of children with their long run life outcomes. The studies began by considering the ability of 4 year olds to delay gratification. 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.[34] While this setup does not document preference reversals, it does document very high discount rates for children who prefer the smaller amount of candy, as did a significant portion of the sample. 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.[35] The key point is stated by O'Donoghue and Rabin (2006a): "Without [present bias], economists have no coherent model of short-term impatience." The children with particularly high discount rates are, therefore, likely to be suffering from self control problems, and their behavior can be a laboratory for studying long term impacts of a lack of self control.

Mischel and Mischel (1983) note several behaviors which allowed the children to wait longer for the larger reward – Mischel's rules. Children were able to wait longer when their attention was shifted, when they could not see the reward, when they were in a cold state, rather than a hot state,[36] and when they were told to be task oriented rather than reward oriented.[37] Moreover, in several studies, *some* children worked out how to focus on something other than the marshmallow (e.g., hiding the marshmallow or concentrating on some other game), thus creating their own soft commitment device of a sort. This behavior lends support to a view in which some people have self control problems, and within this group a subset of people are able to develop strategies to deal with them. Understanding who falls in to which group (i.e., is it a result of nature or nurture, a result of education, etc.,) would be interesting and useful.[38]

Perhaps most telling, the children's behavior in these experiments predicts future

[34]The reactions of the children to these experiments is worth a trip to youtube (search "Mischel marshmallow").

[35]The 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 $x \geq \frac{1}{(1/2)^{24}} = 16,777,216$

[36]A hot state was created by telling the children to concentrate on the taste of the reward, and a cold state by telling children to think about the abstract qualities of the reward.

[37]Thinking about their aim, rather than what the rewards would be like.

[38]Interestingly, the children often waited for a significant amount of time before giving in to the temptation and consuming the candy. This sort of behavior cannot be explained with a sophisticated model. 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 out many pay day loans before their default. This sort of behavior is hard to explain with a calibrated quasi-hyperbolic model and leads them to conclude that pay day borrowers are best modeled as partially naive agents. The experiments, therefore, provide support for boundedly rational models of self control.

life success ten years later. Those who were better able to delay gratification are more attentive, better able to concentrate, exhibit greater fustration 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 put self control mechanisms in place go on to live more satisfactory lives, by some measures, and Suggests that commitment contracts may be able to limit the impact of this lack of childhood development.

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 idea 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 consumer accumulating 25% more. Again the study shows the *potential* for commitment to have real impacts in people's lives and also seems to validate the use of these questions as providing an accurate measure of self control.

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. There are two parts to the neuroscientific dual self theory of the brain, which can be mapped onto the economic dual self theories discussed above. First, the brain has two sources of motivation with conflicting desires. Second, the ability of the "controlling" center to alter behavior is limited.[39] Both of these propositions have found support in neuroscientific studies.

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

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."[40] 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)) The approach, however, does not shed light on how the two areas interact and is consistent with most of the models discussed in our theory section.

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 obvious 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.[41] 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 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

[40] At 506.

[41] To be strictly consistent with the dual self approach, this experiment requires that cognition and self control centers are the same.

bear, and to put a dash on the top of their paper every time they did. 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.[42]

While supporting a dual self model, it is not clear how much support these last two findings lend to the demand for commitment. Specifically, many commitment devices do not take options off the table, but rather make one of the options more costly; however this need not change the amount of self control needed to resist. The temptation that a chocolate cake presents to the primitive brain would seem to be independent of its cost, so long as that cost is pushed into the future. The impact depends 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.[43] However, under this interpretation, raising the cost of the cake would be welfare reducing, assuming that the agent has to pay the cost of self control.

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 chocolate cake *while* remembering 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

[42]Work 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. While this is suggestive, it does not seem to be able to answer our questions about how the self control is affected and in particular whether someone remembering a 7 digit number has exhausted the ability for lateral prefrontal cortex to control the limbic system when resisting a chocolate cake.

[43]It is not clear how the GP model should approach this question. It provides no definition of temptation and therefore we are free to assume that free chocolate cakes are more tempting than \$10 chocolate cakes. However, once we consider models of brain function, this seems to be a less realistic modeling choice.

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

6 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.

6.1 Lab Evidence

Read et al. (1999) study simultaneous and sequential choice of virtue (long term benefits) and vice (short term benefits).[44] An individual is said to make a simultaneous choice if she decides *now* what she will consume in the future, and to make a sequential choice if she must decide later, 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 to watch on three different days. The choices were made from a list of highbrow and lowbrow movies. In the simultaneous treatment they chose all three ahead of time and in the sequential treatment they chose on each day. In the simultaneous treatment 44% of subjects chose 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 shows the success of a commitment device in changing behavior, but does not address the demand for commitment since subjects were not given the choice to choice simultaneously or sequentially, but rather randomized into the different groups. Thus the welfare implication is unclear.

[44]Their defintition of a virtue is similar to Dellavigna and Malmendier's (2004) definition of an investment good.

Trope and Fishbach (2000) study self imposed penalties in the context of a medical test. Subjects received a glucose test, and were told that the results would help inform them how to maintain the blood glucose level optimal for their cognitive functioning. 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 – the information – and a short term cost – the pain of avoiding sugary foods. Subjects were also asked how much they would be willing to pay if they did not successfully fast; and importantly, they were required to pay this amount. 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, as subjects facing greater temptation imposed higher costs on their own potential failure.

In the same paper, Trope and Fishbach (2000) studied whether subjects were willing to make a bonus 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 bonus 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 student workers 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 also indicate that providing the ability to commit can improve performance. The results also show that agents may not be able to demand the optimal commitments, perhaps because of some form of naivete, temptation over menus, or as Ariely and Wertenbroch argue, because the decision makers can see the normative appeal of removing all restrictions on themselves.

6.2 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.

Besley (1995) finds that local savings opportunities in West Africa offer negative interest on savings. This finding shows that some people are willing to pay to have money taken out of their hands. The SafeSave program in Bangladesh shows similar results (Rutherford (2000)) offering a deposit collection and borrowing service with effective negative interest rates on savings. While one interpretation of these findings 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. The most direct evidence on this proposition, however, is unclear. Ashraf et al. (2006a) asked time discounting questions of a group of individuals marketed a deposit collection service. They see a *lower* uptake of deposit collection services among those who exhibited preference reversals. This offer was made subsequent to the commitment savings account SEED, discussed above, 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. The key counterfactual is to understand 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 hy-

pothesis, 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.

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 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.

6.3 Formal Commitments in the Field

Although naturally many firms offer products which are similar to commitment devices, pure commitment devices are rare. Impure devices range from certificates of deposits (since they lock in savings and create a penalty for withdrawal) to buying in bulk, but for the reasons discussed in the introduction we are focusing on a tighter

definition of commitment devices. 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 purchase of fertilizer in Kenya.

Benartzi and Thaler (2004) designed a saving plan – Save More Tomorrow™ (SMarT) – motivated by three aspects of behavioral theory. First, naive agents will tend to procrastinate on taking action, believing they will take actions later. Second, quasi-hyperbolic or GP consumers (and perhaps others) will be tempted to save less than optimal at any given time. Third, commitments need not bind, as the inertia associated with naivete will stop withdrawals. The first two ideas suggest a saving system where enrolment is by default – counteracting procrastination – and the decision to increase saving affects only future selves – counteracting temptation.

In SMarT, employees are offered the chance to signup to increase their *future* automatic paycheck deduction for retirement savings. Employees can decide to cancel or reverse the increase at any time at no cost beyond that of filling out a form. Future increases were timed to coincide with the employee's future pay raise. At one firm, only 3 out of 162 participants opted out of the SMarT program before their second pay raise, and by their fourth pay raise over 85% of participants still remained in the program. At a second firm, savings rates for individuals who signed up for the SMarT program rose by roughly 2%, while other employees "did not change their savings rate much." At a third firm, those individuals who signed up for the SMarT program increased their savings rate by about 1.5%, while savings rates at the firm's other divisions, where SMarT was not offered, remained mostly steady. This work presents real world evidence on the demand for commitment, and its effects, and once again shows that soft commitments can be very effective in practice.

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 above. SEED accounts 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. Some addicts also make use of hard commitments. Alcoholics are also known to use certain metabolism-inhibiting drugs, such as Antabase (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, these informal devices meet our definition of "commitment device," and have found some success (Goldstein (2001)). However, 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. (2008). 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. Treatment on the treated estimates imply a 30% to 65% quit rate for this population relative to the control group. Notably, these results held up in a 12-month follow-up study (6 months after the smokers were allowed to withdraw their money) where the treatment on the treated effect was found to be a 31% to 53% increase in the probability of passing the follow-up urine test. Those who successfully quit had higher balances at stake than those who did not quit, with the successful-quit average balance nearly four times the average balance of those who did not successfully quit.

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 naivete, 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 naivete. 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?

7 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 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. (2008) 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?[45]

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 Intrahousehold 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.

[45]We have not touched upon a growing theoretical literature which discussed market provision of commitment contracts (DellaVigna and Malmendier (2004), Gottlieb (2008), Eliaz and Spiegler (2006) and Heidhues and Koszegi (2008)) , optimal provision of contracts (Amador et al. (2006) and O'Donoghue and Rabin (1999b)) and potential government intervention through taxation (Gruber and Koszegi (2001); Gruber and Mullainathan (2005); Krusell et al. (n.d., 2002); Krusell and Jr. (2003); O'Donoghue and Rabin (2003, 2006b)) or direct regulation (Heidhues and Koszegi (2008)).

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.

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 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.

Besley, Timothy, "Saving, Credit and Insurance," in Jere Behrman and T.N Srinivasan, eds., *Handbook of Development Economics*, Elsevier, 1995.

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.

Caplin, A. and A. Schotter, *The foundations of positive and normative economics: a handbook*, Oxford University Press, USA, 2008.

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.

Dekel, Eddie, Barton Lipman, and Aldo Rustichini, "Temptation Driven Preferences," *The Review of Economic Studies*, 2009, *73* (6), 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.

Eliaz, 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.

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," *Mimeogr., Dep. Econ., Yale Univ*, 2008.

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. and B. Koszegi, "Is Addiction Rational"? Theory and Evidence*," *Quarterly Journal of Economics*, 2001, *116* (4), 1261–1303.

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.

_ **and** _ , "Self-control, revealed preference and consumption choice," *Review of Economic Dynamics*, 2004, *7* (2), 243–264.

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.

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.

Kreps, D.M., "A representation theorem for" preference for flexibility"," *Econometrica: Journal of the Econometric Society*, 1979, pp. 565–577.

Krusell, P. and A.A. Smith Jr, "Consumption-savings decisions with quasi-geometric discounting," *Econometrica*, 2003, pp. 365–375.

_ **, B. Kurusçu, and A.A. Smith Jr**, "Equilibrium Welfare and Government Policy with Quasi-geometric Discounting," *Journal of Economic Theory*, 2002, *105*, 42–72.

_ **, B. Kuruscu, and NJ Princeton**, "Temptation and taxation."

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.

Mischel, H.N. and W. Mischel, "The development of children's knowledge of self-control strategies," *Child Development*, 1983, pp. 603–619.

Mischel, W., Y. Shoda, and MI Rodriguez, "Delay of gratification in children," *Science*, 1989, *244* (4907), 933–938.

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.

_ , "Temptation, Welfare and Revealed Preference," 2009.

O'Donoghue, T. and M. Rabin, "Doing it now or later," *American Economic Review*, 1999, pp. 103–124.

_ **and** _ , "Incentives For Procrastinators*," *Quarterly Journal of Economics*, 1999, *114* (3), 769–816.

_ **and** _ , "Studying optimal paternalism, illustrated by a model of sin taxes," *American Economic Review*, 2003, *93* (2), 186–191.

_ **and** _ , "Incentives and Self-ControP," in "Advances in Economics and Econometrics: Theory and Applications: Ninth World Congress" Cambridge University Press 2006, p. 215.

_ **and** _ , "Optimal sin taxes," *Journal of Public Economics*, 2006, *90* (10-11), 1825–1849.

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 E ect 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.

Silver, B.D., B.A. Anderson, and P.R. Abramson, "Who overreports voting?," *The American Political Science Review*, 1986, pp. 613–624.

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.

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.

Tracy, B., *Goals!: How to Get Everything You Want–Faster Than You Ever Thought Possible*, Berrett-Koehler Publishers, 2004.

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.

Wertenbroch, K., "Consumption self-control by rationing purchase quantities of virtue and vice," *Marketing Science*, 1998, pp. 317–337.