

NBER Working Paper Series

AN ECONOMIC THEORY OF SELF-CONTROL

H. M. Shefrin* and Richard Thaler**

Working Paper No. 208 revised

CENTER FOR ECONOMIC ANALYSIS OF HUMAN BEHAVIOR
AND SOCIAL INSTITUTIONS

National Bureau of Economic Research, Inc.
204 Junipero Serra Boulevard, Stanford, CA 94305

revised July 1978

Preliminary; not for quotation.

NBER working papers are distributed informally and in limited number for comments only. They should not be quoted without written permission of the author.

This report has not undergone the review accorded official NBER publications; in particular, it has not yet been submitted for approval by the Board of Directors.

Thaler's research was supported in part by a grant to the National Bureau of Economic Research from the Kaiser Family Foundation. The authors thank Tom Russell for his helpful comments on an earlier draft. Of course, he should not be held responsible for any remaining weaknesses.

*University of Rochester and University of Santa Clara.

**Cornell University and National Bureau of Economic Research.

AN ECONOMIC THEORY OF SELF-CONTROL

H. M. Shefrin and Richard Thaler

Abstract

Although many economists, most notably Strotz, have discussed dynamic inconsistency and precommitment, none have dealt directly with the essence of the problem: self-control. This paper attempts to fill that gap by modeling man as an organization. The Strotz model is recast to include the control features missing in his formulation. The organizational analogy permits us to draw on the theory of agency. We thus relate the individual's control problems with those that exist in agency relationships.

Richard Thaler
Graduate School of Business
and Public Administration
Cornell University
Ithaca, NY 14853

The idea of self-control is paradoxical unless it is assumed that the psyche contains more than one energy system, and that these energy systems have some degree of independence from each other.

Donald McIntosh [1969]

Introduction

Although economics has been applied to more and more kinds of problems in recent years, the title of this paper may still raise more than a few eyebrows. Why do we need an economic theory of self-control? What is wrong with our traditional theories? Why should self-control be of interest to economists? This introduction will try to answer these questions and provide some information about the contents of the rest of the paper. The answers to the questions involve several points.

1. A significant portion of consumer behavior is characterized by the presence of self-control problems. It shouldn't be difficult to visualize the importance self-control might have in understanding savings behavior. This alone would justify its investigation. Yet we believe that problems of self-control will be present to some degree in all consumer decisions that involve intertemporal tradeoffs. This paper is a first step toward a positive theory of intertemporal choice.

2. Our current theories are inadequate. A simple result from standard consumer theory is that a consumer cannot improve his welfare by restricting his choices. Yet people engage in such activities frequently. In fact, savings behavior is nearly dominated by institutions that reduce the individual's flexibility. Some examples

are Christmas clubs, mandatory pension plans, social security, payroll savings plans, whole life insurance policies, and piggy banks. Christmas clubs have always been a real puzzle for economists. For many years Christmas clubs paid no interest and were seemingly dominated by other forms of saving such as simple savings accounts. It is obvious that their popularity was due to their value as a self-control device. We claim that the popularity of the other institutions cited is also related in varying degrees to the same attribute.

3. Our proposed theory is simple, uses standard economic tools, and yet describes actual behavior. Our basic advance is simply the recognition of self-control as a problem. Our model explicitly deals with consumer choice as a control problem. As the quotation above suggests, self-control seems to necessarily imply the existence of a controller and a controllee. Individuals are thus assumed to behave as if they possessed two separate sets of preferences. These two aspects of their personality are referred to as the planner and the doer. By modeling behavior in this way we seem to have taken the only reasonable course.¹ Once this formulation is adopted it becomes apparent that individuals share many of the control problems found in organizations. The resulting analogy can be exploited to test the model. If individuals are like organizations facing control problems, then they should adopt many of the same strategies for solving these problems. In fact, we observe that they do.

The plan for the paper is as follows. The economic literature that has come closest to discussing self-control is that which has followed the classic paper by Strotz. We thus begin our paper in Section I with a

statement of Strotz's theoretical problem. Section II contains our justification for adopting a two-self model. We draw heavily in this section on the work of Donald McIntosh. Section III presents our formal model. We show that standard economic tools can be used even if the unified self is dropped. Section IV describes a control problem in an organization to illustrate the techniques that can be used to reduce the costs arising from conflicts of interest. Section V shows how the same techniques are used by individuals to deal with their problems of self-control. Section VI discusses some empirical implications of our model. Section VII contains a short summary.

I. Dynamic Inconsistency and Self-Control

Why do people impose constraints on their future behavior? This is the question which motivated us to write this paper, and it is a question which has attracted the attention of economists since the seminal paper by Strotz [1955]. Strotz's answer hinges on the observation that most people's tastes change over time in a systematic way. While Strotz's discussion contains many useful insights into the problem, he fails to deal directly with self-control, and this we feel renders his model inadequate. We will begin by summarizing his argument.

Strotz considers an individual with a nonnegative stock $K(0)$ of an exhaustible resource at time 0. The individual must decide on a consumption plan $c(\cdot)$ by which he will deplete the resource over a finite time interval $[0, T]$. Let $U[c(t), t]$ be the utility accruing at a time t and $\lambda(t, \tau)$ the rate of discount applied to $U[c(t), t]$ when viewed from time τ . At time 0, the individual is assumed to choose a plan $c^*(\cdot)$ to maximize

$$\int_0^T \lambda(t, 0) U(c(t), t) dt \quad (1)$$

subject to

$$\int_0^T c(t) dt \leq K(0) \quad (2)$$

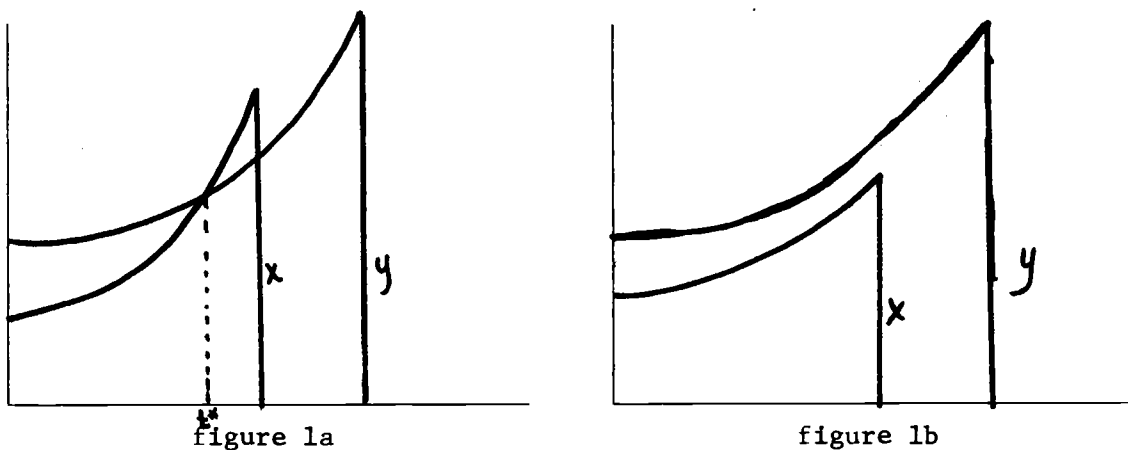
In following a plan $c(\cdot)$, the amount of resource used up between times 0 and τ is just $\int_0^\tau c(t) dt$. Consequently, the amount of the resources remaining at time τ is simply

$$K(\tau) = K(0) - \int_0^\tau c(t) dt \quad (3)$$

The plan $c^*(\cdot)$ that maximizes the above problem is called the original plan. Notice that $\lambda(t, 0)$ is the rate of discount applied to $U[c(t), t]$ in the determination of $c^*(\cdot)$. Of course, the individual's

tastes may have changed at τ because he will discount at rate $\lambda(t, \tau)$ and not $\lambda(t, 0)$ (if $\lambda(t, \tau)$ is not independent of τ). Will the individual want to alter the original plan when his tastes change? That is, will he want to reallocate the remaining amount $K(\tau)$? One of Strotz's conclusions is that the individual will not alter the original plan if $\lambda(t, \tau)$ is exponential in $|t-\tau|$; otherwise he will. If the original plan is altered, then the individual is said to display dynamic inconsistency.

This result is illustrated in figure 1. Suppose an individual must choose between a small reward x at time t and a larger reward y at time t' . Time is measured horizontally with "today" at the origin. The present value of either alternative is given by the height of its respective curve at any point in time. Dynamic inconsistency is shown in panel a.



If the decision between x and y is made any time before t^* , y will be chosen. However, between t^* and t (the time just before x would be received) x will be preferred to y . Thus, if the choice is made before t^* , the individual will want to change his mind once t^* arrives. If the discount function is exponential, then the curves will never cross (as in panel b) and choices will be dynamically consistent.

What will an individual do if he recognizes this inconsistency? One strategy Strotz suggests is that of precommitment.

Today it will be rational for a man to jettison his "optimal" plan of yesterday, not because his tastes have changed in any unexpected way nor because his knowledge of the future is different, but because today he is a different person with a new discount function--the old one shifted forward in time. Yet, it is also rational for the man today to try to ensure that he will do tomorrow that which is best from the standpoint of today's desires.²

Since the individual knows his preferences will change he precommits behavior to enforce his current preferences on himself later. Strotz claims that this is rational. Yet, while the change of preferences hypothesized by Strotz is necessary for precommitment to be rational, it is not sufficient. Consider the plight of a man with the following preferences: at 3:00 PM he prefers fish over meat for dinner, while at dinner time, he prefers meat over fish. If he does his shopping at 3:00 PM and knows that his preferences will change, what should he buy? For it to be rational to buy fish, we must add another condition which is implicit in Strotz's reasoning: namely, that the earlier preferences are judged in some sense to be "right."

Another example will illustrate the point. Suppose an individual observes that whenever he goes to a restaurant the combination of his hunger and the pleasant aromas emerging from the kitchen induce him to order more than he can eat. He is aware of this tendency to order too much, but never seems to overcome it. However, if he called in his order at 3:00 PM when less hungry and away from the aromas, he would order the correct amount. In this example, precommitment seems entirely sensible since it overcomes a systematic bias which the individual recognizes.

In fact Strotz's subsequent discussion demonstrates that he does view the early preferences as "right." Dynamic inconsistency can occur whenever

the discount function is not exponential, but Strotz has in mind a particular shape of the discount function:

Special attention should be given, I feel, to a discount function... which differs from a logarithmically linear one in that it "over values" the more proximate satisfaction relative to the more distant ones....My own supposition is that most of us are "born" with [such] discount functions...³

If the instantaneous discount rate is plotted over time, then Strotz is hypothesizing a function which has the shape illustrated in Figure 2. In choices between "now" and "later" the later option will be heavily discounted,

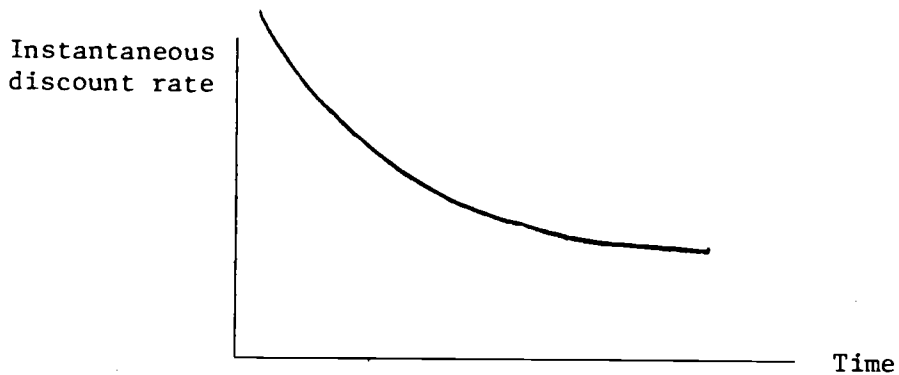


figure 2

but the rate applied to two "later" options will be much lower. Strotz is not very explicit about why he supposes we are born with such discount functions, but he cites Böhm-Bawerk who is explicit.

It is one of the most pregnant facts of experience that we attach a less importance to future pleasures and pains simply because they are future... To goods which are destined to meet the wants of the future, we ascribe a value which is really less than the true intensity of their future marginal utility... Which of us has not been surprised to find that under the pressure of some momentary appetite, he was not able to refuse some favorite dish or cigar which the doctor had forbidden--knowing perfectly that he was doing an injury to his health, which, calm consideration would tell him, was much more considerable than the pleasure of that trifling indulgence?... Any one who knows himself, and keeps his eyes open to what is going on around him, will find this fact of the underestimate of future pleasures and pains exhibited under a thousand forms in the midst of our civilized society. Of the fact then there is no doubt.⁴

Three reasons are given for this phenomenon: want of imagination, defect in will, and the uncertainty of life. It is defect of will, however, which leads Böhm-Bawerk to suggest a discount function of the shape illustrated in Figure 2.

I should like to call special attention, further, to the fact, that the undervaluation which results from these causes is not at all graduated harmoniously, in the subjective valuation of the individuals, according to the length of the time that intervenes... On the contrary, the original subjective undervaluations are, in the highest degree, unequal and irregular. In particular, so far as the undervaluation is caused by defects of will, there may be a strong difference between an enjoyment which offers itself at the very moment, and one which does not; while, on the other hand, there may be a very small difference, or no difference at all, between enjoyment which is pretty far away, and one which is farther away.⁵

We agree.

While Strotz never mentions self-control explicitly, he does discuss the tendency of lower income people to "gorge themselves with food after pay-day; overheat their homes when they have money for a bucket of coal;... go on sprees on pay-day; engage in heavy installment buying..."⁶ etc. all of which he says "can be explained as a failure to cope intelligently with the problem of the intertemporal tussle."⁷ What is the "intertemporal tussle" if not the lack of self-control?

Finally, consider Strotz's remarks on consumer sovereignty:

I would have confidence in the judiciousness of a person today...to decide how much to save and how much to spend for the rest of his life starting a couple years from now... The real decisions to worry about are those where an immediate or proximate satisfaction is gained at the expense of still-more-future costs. Precommitments may be regarded as either good or evil, depending upon whether the period of precommitment begins now or later.⁸

These remarks, while insightful and sensible, cannot be justified by Strotz's model. What in Strotz's model leads him to the conclusion that some precommitments are good and others are evil? Why are decisions whose impact only

begins a couple years from now considered good? The remarks seem to be based not on the model (in which it is always rational to do what is best given today's preferences) but rather on Strotz's value judgment that the high discount rates observed in the short run are inappropriate. To make the model complete, however, it is only necessary to have the individual share Strotz's value judgment. Since delay of gratification is more difficult as the object of desire draws nearer, observed discount rates become high in the short run. Sophisticated individuals will recognize this internal, systematic bias and, like the man in the restaurant, they will rationally take steps to reduce the costs of this bias. The contribution of this paper is to explicitly recognize these costs, and to show how they can be incorporated into an economic model of intertemporal choice.

II. The Incoherent Self

Strotz hypothesizes a non-exponential discount function which produces dynamic inconsistency. To complete the model all that is necessary is to incorporate the conflict which this inconsistency will inevitably produce. Conflict arises because the individual recognizes his own weaknesses. Plans made in advance are consistently broken because "temptation" becomes too great. What the person knows to be his best long run interests conflict with his short run desires. To model such conflict it is useful to assume the individual acts as if he were influenced by two separate sets of preferences, in essence a two-self model. This idea is also advocated by Donald McIntosh [1969] in a book which has strongly influenced our presentation. In his chapter entitled, "The Psychology of Rational Action," McIntosh evaluates utility theory from a psychological perspective. We will very briefly summarize his ideas.

Standard utility theory models man as a sophisticated, maximizing agent. How does man acquire the skills necessary for him to act rationally? McIntosh notes that individuals are born not with coherent purposes but with drives or needs. The satisfaction of these needs requires the establishment of certain relationships with the outside world. These relationships are established as the individual matures. During this time, a coherent idea of selfhood develops; however, a considerable amount of psychic conflict is always present. This necessitates that some impulses toward drive satisfaction be blocked because of the existence of a multiplicity of drive mechanisms.

The needs of the individual are not one but many; they are present now, but they will also be present in the future.

Short-term satisfactions must be weighed against longer-run results... In a word, self-control is needed.

The idea of self-control is paradoxical unless it is assumed that the psyche contains more than one energy system, and that these energy systems have some degree of independence from each other.⁹

The last sentence is precisely the position we take in this paper.

McIntosh continues by citing two situations in which traditional utility theory cannot be applied. The first situation, which he calls discontinuity of purpose, refers to intertemporal preferences that change over time. He calls the second situation incoherence of purpose. Incoherence refers to the presence of psychic conflict at a single point in time--which is to say that the individual is influenced by more than one preference system. This distinction cannot be overemphasized. In particular, the reader should note that Strotz's model is based only on discontinuity of purpose. Yet, Strotz's remark that "the real decisions to worry about are those where an immediate ...satisfaction is gained at the expense of...future costs," is entirely consistent with McIntosh's statement about independent energy systems. The major contribution of this paper is the incorporation of incoherence of purpose into a formal model of individual behavior.

We model incoherent purpose by treating an individual as if he contained two distinct psyches which we will denote the planner and the doer. These terms help stress the analogy between an individual with a self-control problem and an organization with a principal-agent problem. We will exploit that analogy extensively in Sections IV and V. The planner/doer framework is developed in a formal model in Section III, but we should point out here that it is consistent with McIntosh's view of the individual as a combination of a system of coherent drives together with a coherent idea of self. The doer corresponds to McIntosh's drives, being concerned only with "short-term

satisfactions," while the planner corresponds to the coherent ideas of the self, it being concerned with the tradeoff between "short-term satisfactions" and "longer-run results." Before turning to the model, there remain two final concepts in McIntosh's discussion which will prove quite useful. They are external autonomy and internal autonomy. Quite simply, external autonomy refers to the best the individual can do for himself when he is fully coherent while internal autonomy denotes the best he can do when some degree of psychic conflict is present.

III. Modeling Incoherent Purposes

In this section we will elaborate on our description of the planner and the doer, and will present a formal model of an individual with incoherent purposes. We have several reasons for developing a mathematical formulation of our model. First, the formal treatment makes explicit the role of the planner and the doers. Second, by structuring our model to closely resemble that of Strotz, we facilitate comparisons of our theory to the existing literature.¹⁰ Third, the formal model makes transparent the close correspondence between our theory and the theory of agency. We draw on this correspondence in Section 4. Fourth, we show that the recognition of incoherence and inconsistency does not preclude the use of the traditional tools of economic theory. Finally, we hope to show that economic analysis can be used to gain insights into what many may consider a purely psychological phenomenon.

We will cast our model in a discrete time framework. Consider an individual with a fixed income stream $y = [y_1, y_2, \dots, y_T]$. The individual is assumed to choose a nonnegative level of consumption c_t in period t ; call $c = [c_1, c_2, \dots, c_T]$ a consumption plan.

What features does McIntosh cite which are absent in Strotz? First, McIntosh associates a set of independent energy systems or drives to the individual; these correspond to the incoherent parts of the self. Second, the self also has a coherent part. It is this part that in McIntosh's terms attempts to balance off "shorter-term satisfactions" against "longer-run results." Third, there remains present within the individual a degree of psychic conflict. In other words, the individual is never completely integrated.

How can these features be introduced into the Strotz framework?

We introduce these features by viewing the individual as an organization unto himself. The organization consists of $T+1$ components: T distinct doers (one for each period) and a single planner. The period t doer is assumed to exercise direct control over the period t consumption level c_t . Doer t corresponds to an independent energy system. Its associated drive is represented by a utility function $Z_t(\cdot)$. $Z_t(c_t)$ denotes the degree of immediate or "short-term satisfaction" that accrues if c_t is consumed in period t . Assume that Z_t is a strictly increasing concave function.

Strotz certainly recognized some of these features as is evidenced by the following quotation:

The individual over time is an infinity of individuals, and the familiar problems of interpersonal utility comparisons are there to plague us. The interpersonal aspect of the intertemporal problem becomes clear if we think of a similar problem involving a family of brothers where each has a utility functional depending not only on his own utility but upon a weighted sum of the utilities of all of them. Suppose the oldest brother always has the power to allocate the annual proceeds of an estate, but with it being foreknown that each year one brother will die off, the oldest next.¹¹

The basic difference between our model and Strotz's would seem to be that Strotz considered an individual to be a system of doers with no planner. In addition, each Strotz doer has some concern for the other doers whereas our doers are completely selfish.¹²

In our model incoherence results from the imposition of a present value budget constraint¹³

$$\sum_{t=1}^T c_t \leq \sum_{t=1}^T y_t = Y \quad (7)$$

Notice that this implies the existence of a perfect capital market. The multiplicity of drive mechanisms $[Z_1, Z_2, \dots, Z_T]$ are in mutual conflict as a result of (7). Consequently, the coherent part of the self must be identified with the individual's ability to express consistent preferences over the achievement of his various drives. The planner effectively fills this role. The planner's preferences are represented by a utility function $V(Z_1, Z_2, \dots, Z_T)$.

An individual is represented by his energy system $[V, Z_1, \dots, Z_T]$. We have now incorporated the first two McIntosh features described above. The third and final aspect centers on psychic integration. If the individual were fully integrated, then the planner would choose a consumption plan to

$$\text{maximize } V(Z_1, Z_2, \dots, Z_T)$$

subject to

$$\sum_{t=1}^T c_t \leq Y$$

In McIntosh's terminology, this refers to the achievement of external autonomy. However, McIntosh asserts that external autonomy is not possible in general because of the absence of full psychic integration. The best the individual can hope to do is achieve internal autonomy. The impediment to the achievement of external autonomy lies, of course, with the doers. It is the period t doer, not the planner, that exercises control over the period t decision c_t . In particular, doer t is oriented towards

achieving maximum short-term satisfaction Z_t , not longer-run gain V . In fact, an unrestrained doer 1 would borrow $Y - y_1$ on the capital market and therefore choose $c_1 = Y$; the resulting consequence is naturally $c_2 = c_3 = \dots = c_T = 0$. Such action would suggest a complete absence of psychic integration.

What can the planner do to exert some control over the doers? In general he has two instruments he can use. First, he can impose rules on the doers behavior. The rules alter the doers behavior by changing the constraints. One possible rule would be to purchase an annuity which allocates each doer a specific consumption level. A less restrictive rule would be to simply forbid borrowing.

The second instrument available to the planner is to alter the doers utility function directly. The goal would be to produce a new function which has an internal maximum. In Figure 3 the original doer utility function is labelled Z^1 . We assume that the only way to reduce the marginal

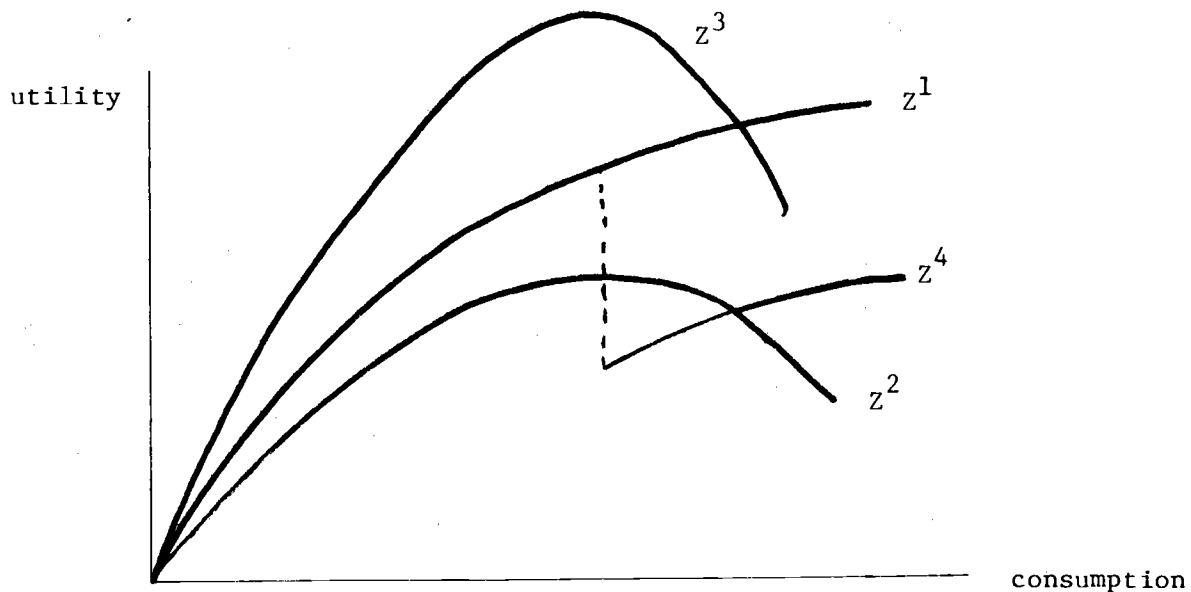


figure 3

utility necessary for an internal maximum is to reduce the overall level of arousal and thus total utility. So the altered utility function will resemble Z^2 . Altering the utility function to Z^3 such that satiation is reached earlier while utility is increased is considered infeasible, as is introducing a discontinuity at the desired level of consumption, such as Z^4 .

To model the modification of the doer's utility function, we introduce a modification parameter $\theta = (\theta_1, \theta_2, \dots, \theta_T)$. Henceforth, Z is assumed to be a function of two arguments, c_t and θ_t . If $\theta_t = 0$, then the doer is completely unrestrained. As θ_t increases, both Z and $\frac{\delta Z_t}{\delta c_t}$ are reduced as in Z^2 in Figure 3. To provide a specific example, θ might be thought of as a guilt parameter. The higher is θ_t , the more guilt the doer feels for any level of c_t .

Define $c_t^*(\theta_t)$ to be the consumption level doer t chooses to maximize $Z_t(c_t, \theta_t)$ when the planner picks θ_t . If sufficient modification has taken place so that Z_t has an internal maximum, then $c_t^* < Y_t = Y - \sum_{s < t} c_s^*$.

We can now write down the planner's problem in the discretionary mode. Let $Z(c^*(\theta), \theta) = [Z_1(c_1^*(\theta_1), \theta_1) \dots Z_T^*(c_T^*(\theta_T), \theta_T)]$. Then the planner wants to solve

$$\begin{aligned} & \underset{\theta}{\text{maximize}} \quad V(Z(c^*(\theta), \theta)) \\ & \text{subject to} \quad \sum_{t=1}^T c_t^*(\theta) \leq Y. \end{aligned}$$

The solution to this problem entails the usual sort of marginal conditions. In this case the planner will increase θ_t until the marginal loss to the planner from the resulting decrease in doer t 's utility is equal to the marginal gain to the planner from the increases in utility to all future doers. Both the gain and the loss have two components. Doer t is worse off from a rise in θ_t because he consumes less and because he enjoys each unit of consumption less. Similarly future doers gain both because there is more income remaining and because less future modification will be employed.

In the more general problem, the planner can choose not just θ but also the set of formal constraints (or rules) he wishes to impose on the doers. We will present details of this more general problem in our next paper on this subject. For our present purposes we will just move on to discuss what actual devices individuals use in dealing with self-control problems. We will begin by examining what organizations do when facing similar problems.

IV. Rules vs. Discretion: An Organizational Analogy

We have modeled an individual facing a self-control problem as an organization with a principal-agent or conflict of interest problem. If this analogy is apt, then we should find that the devices used by organizations to minimize the costs imposed by conflicts of interest are also used by individuals in intertemporal choice situations. In this section we will develop a specific example of an organizational control problem. Its applicability to individuals is examined in Section 5.

Consider the case of a bank that is run by an owner-manager. One of the functions which the owner serves is that of loan officer. He determines which applicants should be granted loans. Two kinds of procedures are used. The applicant fills out a report which is coded and run through a "credit scoring" program which predicts the probability of default. In our terminology the credit scoring procedure is a rule. This rule can also be supplemented or supplanted by the judgment of the owner based on information gathered in a personal interview. The interview judgment process is labeled discretion.

Since the interviews are more costly to conduct, the owner would rely completely on the credit scoring procedure if it were equally effective. However, the interview may permit the owner to gather information which is difficult to obtain on a written application (such as "appearance") or is difficult to process quantitatively (such as nervousness in response to particular questions). The owner will utilize the interview if the extra precision in granting loans results in sufficient extra profits to cover the extra costs. Now assume that the owner hires an employee to process loan applications. Suppose further that the owner

used both the credit scoring procedure and the interview because the additional accuracy of the interview more than made up for the extra cost. Finally assume that the employee is just as skillful as the owner in judging loan applications. While the production function does not appear to have changed, it may now be profitable to abandon the interview. The reason for this is that the incentives facing the employee may not be the same as those facing the owner. The equal skill assumption means that if the employee were the owner he would make decisions identical to (or as good as) the current owner's, but as an employee who does not get to keep the profits of his section, his decisions may differ. He may become careless, lazy or even dishonest. This will create an incentive to adopt rules. The situation described is now analogous to that of an individual facing a problem of self control. He (like the owner) knows what he should do, but can't get himself (the employee) to do it.¹⁴

Agency costs can be defined as the difference between the profits of the firm that would occur if every agent had the same objective function as the owner and the actual profits (or the difference between external autonomy and internal autonomy).¹⁵ Organizations will choose a mix of rules and discretion designed to minimize agency costs. Rules are used to reduce the opportunities for the employee to misbehave. Discretion must be combined with appropriate incentives. Some specific strategies for the bank example will illustrate the basic techniques.

If no interviews are permitted then the bank is using a pure rule strategy. Two techniques can be used to introduce a limited amount of discretion. First, the bank might have the credit scoring program produce

a cardinal score (rather than a zero or one). This score might be the estimate probability of default, p . A pure rule would be to grant loans only if $p < p^*$. Limited discretion could be added by allowing the program to be overruled over a certain range. So, for example, all loans with scores less than $p^* - \delta$, would be approved and all with scores greater than $p^* + \delta$ would be rejected, and over the range $p^* - \delta$ to $p^* + \delta$ the employee would decide.

Second, certain classes of decisions may be removed from discretion. Thus, the owner might feel that some loans are particularly likely to be associated with large divergence between the interests of the employer and the employee. Examples might be loans to friends and relatives, or very large loans. These particular decisions might be removed from the employee's discretion. Which decisions to remove will depend not just on the probability of malfeasance but also on the cost of determining when the rule should apply. So abolishing discretion for "attractive clients" might be advisable in theory but infeasible because of the difficulty in defining an attractive client. Once any discretion is permitted the bank will try to reduce the agency costs by creating an environment in which the employee's interests are as close as possible to the employer's. Three basic methods can be used to alter the employee's incentives. First, the employer can monitor the employee's inputs, or the decisions themselves. A random sampling of interviews with appropriate penalties for incorrect decisions is an example. Second, the employer can monitor outputs. Setting up the loan department as a profit center with the employee's salary determined by profits would be a strategy of this type. Third, the owner can try to alter the employee's interests through moral suasion. Many profit-sharing

arrangements may be based on this concept. Most profit-sharing systems provide very little actual incentive to any employee since his share of marginal profits is minute. But, by "giving the employees a share of the profits" the firm changes the way employees think about the firm. Though moral suasion and social norms are typically sneered at by economists, they explain a great deal of otherwise mysterious behavior such as tipping in strange out-of-town restaurants and cleaning up campgrounds.

The preceding analysis applies to any form of organization. The organization should select some combination of rules and incentives to minimize agency costs. The actual mix selected will depend on the relative costs of each strategy. This observation has an interesting implication for our understanding of the workings of government bureaucracies and other non-profit organizations. A characteristic all such organizations share is that output is difficult to measure. Because of this, the various incentive strategies open to profit-making firms will be either awkward or impossible to implement. The Army cannot, for example, set up a platoon as a profit center! This implies that such organizations will be forced to rely on rules to a greater extent than will firms. Since bureaucracies cannot measure output they are forced to use rules to approximate the goals of the organization and then monitor the adherence to these rules.

V. The Techniques of Self-Control

Walter Mischel, a prominent psychologist who has conducted experiments on self-control for over a decade, suggests that self-control should be thought of as a two-stage process. In the first stage the individual must choose to wait for the more preferred but delayed outcome. In our model this corresponds to the preferences of the planner. The second stage is the execution of the delay. During this stage it helps if "the person can convert the difficult aversive 'self-control' situation into one which he can master more easily."¹⁶ Thus the second stage entails getting the doer to follow the plan. This section investigates the techniques of self-control--those strategies which the individual uses to make the difficult easier. The techniques we describe are those suggested by our model and by the bank analogy from the previous section.

We have several objectives we hope to meet in this section. First, we want to describe the kinds of behavior we should expect to observe in self-control situations. It should be stressed that these are rational, maximizing solutions to choice problems in a second-best world. In fact, we show that individuals use the same strategies to deal with self-control problems that profit maximizing firms use to deal with conflicts of interest. The strategies are rational because the costs of self-control are real (just as agency costs are real). Second, we can explain some kinds of behavior which seem irrational. By rationalizing seemingly irrational behavior we can make positive statements about when and where we should expect to observe it. Third, we will try to indicate some of the weaknesses in the neoclassical theory of intertemporal choice, again pointing to the specific situations in which the theory is most likely to fail.

We will proceed as follows. The two basic instruments the planner can use in our model are rules and discretion. Rules operate by altering the constraints imposed on any given doer. Discretion must be accompanied by some method of altering the incentives or rewards to the doer. Obviously, these incentives must have short run payoffs. Both rules and discretion can be implemented either with external help or purely internally. Furthermore, many strategies involve a combination of both rules and discretion. We will describe each technique in turn: discretion (external and internal), external (or pure) rules, and combined strategies (internal rules-of-thumb and opportunity manipulations). For each strategy we indicate its particular advantages and disadvantages and illustrate with everyday examples usually taken from savings or dieting behavior. Where possible we point to specific institutions which have arisen to meet the demands for externally imposed self-control aids. We conclude the section with a discussion of the desirable characteristics of rules-of-thumb.

We hope this section will answer the following questions: Why do people impose rules on themselves? What happens when no rule is used? What characterizes a good rule? What determines the choice between rules, discretion, and various combinations?

Discretion

The essential feature of discretion is that the doer is allowed to choose in an unconstrained way. Thus if behavior is to be altered, the incentives facing the doer must be modified. We will describe how this can be done.

If external help can be used then the simplest strategy is to rearrange the short-run rewards for a specific activity. An extreme example is the

drug antabuse used by alcoholics. The effect of this drug is to make the user sick immediately after taking a drink. Less extreme but more salient to academics is the practice of agreeing to give a paper at a conference to see that it will in fact get written. Again, this works because as the conference draws near, failure to write the paper will result in immediate costs (abuse from the organizer). Similarly, some college basketball players who have signed professional contracts before graduating have asked for special clauses to be included in their contract specifying that a large bonus (\$10,000) will be paid immediately upon graduation. Since the returns to finishing will only begin after their basketball career is over this acts to shift some of the reward forward. Finally, some people simply make a bet with a friend: "I will pay you \$200 if I smoke another cigarette."

The same basic idea can be done purely internally. Thus some people deny themselves some reward until they finish some unpleasant task. An ingenious variation on this idea is Ainslee's "private side bet."¹⁷ The essence of many self-control problems (smoking, dieting, saving) is that each particular instance of restraint has only a trivial long-run gain. Just one cigarette (donut, spree) will have no significant effect on lifetime health (weight, wealth). Yet, at the time, the utility of the immediate reward may loom very large. Compare the utility of a tantalizing dessert with the loss in utility of weighing an extra gram! Yet if all decisions are made on this basis, the individual will never restrain, and the cumulative effects will be significant. To overcome this Ainslee suggests that the individual might tie all acts of a specific form of restraint together. In essence he bets with himself that he will

never eat donuts. He will want to keep this bet as long as the utility from eating donuts is less than the utility of the weight loss associated with not eating donuts. If the scheme works it is because the individual perceives that breaking the rule even once will jeopardize the entire routine. Thus the cost of eating a donut would be viewed not just as the weight gain from one donut but the weight gain from a lifetime of donuts. If this perception is correct it seems like it must be based on two aspects of the technology of restraint. First, good behavior can be habit-forming. (The costs of restraint decline with continual practice.) Second, such habits are very easily broken. Both aspects seem accurate descriptions of human behavior.

Two other internal incentive alteration techniques are worth mentioning: monitoring and moral suasion. Self-monitoring is simply the process of keeping track of various activities. "Weight-watchers," a diet club, uses this technique extensively, as its name suggests. Not only are outputs monitored through weekly weigh-ins (with appropriate reinforcement) but members are instructed to weigh and count carefully everything they eat. It seems that in many self-control situations simply keeping track helps cut down on the undesirable activity.

Moral suasion can also be effective if short-run incentives can be so altered. Essentially, moral suasion involves the adoption of some norm. An example would be to view saving as a goal in and of itself.¹⁸ If the planner can convince the doer to adopt this norm then the doer will save in spite of myopic preferences. Saving will provide positive current utility. Guilt can work the same way, though in the opposite direction. If overeating is considered bad in and of itself then the problem pointed out by Ainslee is minimized.

Pure Rules - External Enforcement

...but you must bind me hard and fast, so that I cannot stir from the spot where you will stand me...and if I beg you to release me, you must tighten and add to my bonds.

Strotz begins his article with the above well-known quote from The Odyssey. The solution which Ulysses adopts to his self-control problem is an externally enforced precommitment. In general, precommitment can be a very effective self-control strategy. Examples are pension plans, fat farms (where you pay not to be fed) and even wiring one's jaw closed. The advantage of these strategies is that once in place they require little or no self enforcement. In terms of our model, if a pure rule is used then θ can be left at zero. This implies that the level of utility will not have to be reduced. In some cases, precommitment may be the only way to partake in an activity. Gamblers, for instance, often bring only as much cash as they are willing to lose, and like Ulysses instruct their friends in advance not to loan them additional funds no matter what they may say later. Institutions which sell precommitment services are very common in the areas of dieting, saving, smoking and drinking. Of course, economically, the savings institutions, especially pension plans, are most important.

Strotz finds precommitment so compelling that he writes: "What needs to be explained is not that people do precommit their future actions, but that the practice is not still more wide-spread."¹⁹ Strotz offers one reason, uncertainty as to future tastes and opportunities. "Because of risk and uncertainty, people are also willing to pay for options permitting them a greater range of choice at future dates..."²⁰ Though this may in fact be the most significant defect of precommitment, there is another reason why alternative strategies will be used.

Pure rules require external help which may be either unavailable or, equivalently, too expensive. The expense may arise because actual physical restraint is required, as in fat farms, and so the precommitment is only feasible for short intervals. Alternatively, the expense may result from the difficulty in defining the legitimate exceptions which occur in an uncertain world. Of course Strotz may have failed to recognize the full range of alternatives to precommitment which also help explain its limited use.

Combined Strategies

In the bank analogy, the owner could combine rules with some discretion by limiting either the range over which the discretion could be exercised or by limiting the domain of discretion to relatively "safe" decisions. Both combinations are used by individuals as well. We will refer to these strategies as internal rules and opportunity manipulation, respectively.

Internal Rules

Our continual observations upon the conduct of others, insensibly leads us to form to ourselves certain general rules concerning what is fit and proper either to be done or to be avoided.

Adam Smith [1759]

We have identified two extreme modes of savings behavior. Pure discretion is characterized by the individual deciding in each period how much to save without any self-imposed constraints. A pure rule eliminates all choice, perhaps with the help of an annuity. Obviously there are many alternative strategies which lie somewhere between these extremes. We call these intermediate cases internal rules because generally they are self-enforced rules-of-thumb, rather than externally enforced precommitments. Two points should be stressed here. First, almost any external rule can be

used internally. Second, almost any external rule can be defeated at some cost. Purchasing an annuity, for example, serves no purpose if the individual is free to borrow at will. Thus the distinction between internal and external rules is somewhat blurred. Which is chosen depends on a comparison of costs. External rules usually have lower self-enforcement costs but higher monetary costs and less flexibility.

In the context of our savings model we can identify some likely rules-of-thumb between pure discretion and a pure rule. Each rule alters the budget constraint facing the doer. A natural first departure from pure discretion would be to introduce a ban on borrowing. We maintain the notation that y_t is current income and Y_t is the present value of remaining future income and we introduce S_t as the accumulated saving up to period t $\left(S_t = \sum_{t=1}^{t-1} (y_t - c_t) \right)$. Then on a no-borrowing regime the budget constraint is simply $c_t \leq y_t + S_t$. A somewhat weaker rule which seems common is to prohibit borrowing except for specific purchases: say houses and automobiles. The existence of such rules-of-thumb may explain some apparently irrational behavior. Many students are eligible for subsidized loans at very favorable rates and yet fail to borrow. Why? It is unlikely that their rates of time preference are less than the 3% interest charged on some of these loans, and in any case the money could be placed in a savings account for a sure gain. While some students do exactly this, many do not. Similarly, countries with high rates of inflation and regulated interest rates (like Israel) provide strong incentives to borrow which many fail to take advantage of. An aversion to borrowing per se seems the best explanation of these anomalies.

Another simple rule-of-thumb is a prohibition on dissaving combined with limits on borrowing. If borrowing is banned then the budget constraint

becomes $c_t \leq y_t$. If borrowing is permitted up to some level, say B, then the constraint becomes $c_t \leq y_t + B$. This rule-of-thumb may explain why many people borrow and lend simultaneously in spite of a substantial difference in interest rates. A particularly striking example is the institution of the "pass book loan." This institution allows consumers to borrow money using their savings account balance as collateral. While the interest rate paid is less than for an unsecured loan, it is still higher than the rate paid on the savings account. Thus the transaction seems dominated by the alternative of simply withdrawing some of the savings. Self-control seems to play a crucial role in explaining this institution. People find it attractive either because they don't want to dissave or because they value the regularized repayment plan associated with the loan, or both.

Each of the above rules may be combined with a savings plan. An example of such a plan would be to save at least $s\%$ of income in each period, with the discretion to save more but not less. Here the budget constraint becomes $c_t \leq (1 - s) y_t$ until retirement. Institutions such as "payroll savings" where a certain amount is deducted each pay period toward the purchase of savings bonds or some other savings instrument are examples of this kind of plan. Since the plan can be stopped at any time and the savings are very liquid it is obviously an internal rule.

Mandatory pension plans and social security are nearly pure rules since both dropping out and withdrawing funds are difficult or impossible. However, they must still be complemented with no borrowing rules in order to be effective.

Opportunity Manipulation

...meanwhile I took a large round of wax, cut it up small with my sword, and kneaded the pieces with all the strength of my fingers. The wax soon yielded to my vigorous treatment and grew warm...I took each of my men in turn and plugged their ears with it...

The Odyssey

Since Ulysses' crew needed to have their hands free to row, he had to employ a different strategy for them, so he filled their ears with wax. Since the Sirens were seductive, pure discretion was also not feasible. In general, seductive goods will require special consideration, as in the bank example. This implies that c_t should be considered a vector, so we denote the level of good i consumed in period t by c_{it} . If θ is also good specific then we can use $\frac{-\delta c^*_{it}}{\delta \theta_{it}}$ as a measure of how seductive good i is. If a good is highly seductive then in utility terms discretion will be very costly. The best alternative may be to avoid the good altogether. Thus dieters refuse invitations to lavish dinner parties rather than go and try to stick to their diet. In the absence of outright prohibition, variations in $\frac{\delta c^*_{it}}{\delta \theta_{it}}$ imply that the individual will appear more impatient with goods that are more seductive. In other words, the discount function implicit in observed behavior will be good-specific. Similarly if $\frac{\delta c^*_{it}}{\delta \theta_{it}}$ depends on the situation then even the implicit discount functions for specific goods will be dynamically unstable.²¹

A concept related to the seductive good is the addictive or habit-forming good. One way to define an addictive good is that it becomes more seductive as it is consumed. A possible measure of the degree to which a good is addictive would then be $\frac{\delta}{\delta c_{it}} \left[\frac{\delta c^*_{it+1}}{\delta \theta_{it+1}} \right]$. It is easy to see that addictive goods present a special problem in self-control, especially if the good is seductive to begin with.²² Drugs such as heroin, alcohol, and

tobacco are the obvious examples. The current doer receives all the benefits of consuming an addictive good while the costs, in terms of future attempts to control behavior as well as harmful side-effects are all imposed on future doers. Pure rules may be an attractive strategy, and could explain some of the support for legal prohibitions we have observed over the years. For those who find the good initially seductive, no other strategy is likely to be successful. We conjecture that most people who manage to avoid becoming smokers simply found the practice originally distasteful. Very few people seem to avoid the temptation to "just try it." The costs of addiction to heroin, on the other hand, are high enough to scare most people away.

Of course "habits" can be both good and bad. For some goods it seems likely that some modification today will make the good less seductive tomorrow. This possibility underscores the value of giving children "proper training."

The correspondence between the individual's problem of self-control and the agency problem faced by an organization is illustrated in Table 1 which summarizes and organizes the ideas presented in Sections IV and V.

Table 1
Summary of Control Techniques for Organizations and Individuals

Category	Illustrations from Bank Analogy	Illustrations from Individual Behavior
Pure Discretion	<ul style="list-style-type: none"> ● Employee makes decisions using interview and credit scoring program as he sees fit 	<ul style="list-style-type: none"> ● Unconstrained deer maximization
Methods to Alter Incentives	<ul style="list-style-type: none"> ● Check employee decisions through random sampling and reinterviews ● Make the loan department a profit center ● Moral Suasion - profit sharing 	<ul style="list-style-type: none"> ● Self-monitoring <ul style="list-style-type: none"> - Weight Watchers ● External incentive alteration <ul style="list-style-type: none"> - Antibuse - Agreeing to give a paper at a conference ● Internal devices <ul style="list-style-type: none"> - private side bets - saving for its own sake
Pure Rules	<ul style="list-style-type: none"> ● Use credit scoring program exclusively 	<ul style="list-style-type: none"> ● Externally Enforced Rules - <ul style="list-style-type: none"> Precommitment - pension plans as exclusive form of saving - fat farms
Partial Rules	<ul style="list-style-type: none"> ● Employee discretion over a limited <u>range</u> of credit scores ● Employee discretion over a limited <u>domain</u> of decisions 	<ul style="list-style-type: none"> ● Rules-of-Thumb <ul style="list-style-type: none"> - no borrowing - no dissaving - savings plan ● Opportunity Manipulation <ul style="list-style-type: none"> - not using credit cards - keeping an empty refrigerator

Desirable Characteristics of Rules-of-Thumb

The preceding discussion has argued that people will sometimes use internal rules of-thumb to resolve dynamic inconsistencies rather than external precommitments or pure discretion. What form will these rules-of-thumb take? We can identify several characteristics that are desirable:

1. Simplicity. A rule-of-thumb can only work if the doer decision process is somehow by-passed. If a rule is too complicated and the doer has to figure out how to apply the rule then its value will be lost.
2. No Exceptions. The ideal rule has no exceptions. If exceptions are necessary they should be very well defined cases. Again, if the doer has to decide whether or not the rule applies, its value is severely diminished.
3. Dynamic Stability. This is closely related to no exceptions. Frequent changes in the rule are comparable to many exceptions.
4. Plausibility. If the doer is convinced that the rule is reasonable and even in his best interest then enforcement costs will be reduced.
5. Myopic Desirability. Any rule, internal or external, must be initially adopted by the current doer. To get over this hurdle it may be necessary to sweeten the pot. Thus some Christmas Clubs offer a free gift for joining, and the Army has sometimes used the strategy of the pre-enlistment bonus. As Thomas Aquinas said, "Lord make me good, but not now."

VI. Implications

Time Preference and the Rate of Interest

A simple result from the neoclassical theory of intertemporal choice is that individual rates of time preference will be equal to the interest rate. In our model this result does not apply. First we agree with Böhm-Bawerk and Strotz that most individuals' discount functions are not exponential. Thus, it makes no sense to talk about a single discount rate. Second, rather than seeking the equality between the marginal productivity of waiting and the marginal rate of time preference, individuals in our model adopt rules (both internal and external) to overcome their high rates of time preference in the short run. The difference between our approach and the standard approach is well illustrated by a study by Kurz, Spiegelman and West [1973] (hereafter "KSW").

KSW used survey techniques to try to measure the rates of time preference of the participants of the Denver Income Maintenance Experiment. They asked a sample of participants a series of questions of the following sort: What size bonus would you demand today rather than collect a bonus of \$100 in one year? Several different forms of this question were asked. KSW state the standard theoretical result, and argue that different rates of time preference among groups imply different rates at which groups can borrow. To assure that everyone in the sample was in equilibrium in the loan market (rather than being rationed) they asked people whether they could borrow \$500 to make an installment purchase or \$1000 in cash and only included those who said they could. The results were striking. For Whites, the mean rates of time preference implied by their answers varied from 36% to 76%. For Blacks the rates varied between 40% and 122%. Different questions

elicited different mean rates but the consistent and surprising finding was that all the rates were much higher than observed interest rates. (Subjects were asked at what rate they could borrow and most replied in the 5-20% range.)

The conclusion drawn by KSW is that: "The attempt to measure the rate of time preference with an interview technique is upward biased."²³ While this conclusion may be correct, the model presented in this paper presents a viable alternative. To us the high implicit rates of time preference are not an anomaly. Most of the questions used offered a choice of "bonuses" or "gifts." Subjects might reasonably view such choices as "exceptions" or "special cases" and thus not governed by rules-of-thumb designed to control myopic behavior. In fact KSW even provide some evidence for rules-of-thumb on borrowing since 81.3% of the sample reported that they would not borrow \$1000 in cash even if they could do so at their perceived market rates of interest (generally less than 20%). The results for Blacks seem hard to explain only with the upward bias hypothesis. Why should the bias be 20 points higher for Blacks? On the other hand, if the responses are regarded as true measures of underlying time preference, then the results are consistent with the hypothesis advanced by Banfield [1970] that Blacks have higher discount rates.

Pensions and Saving

A topic of current interest is the effect of social security on saving. Since this effect is difficult to measure directly, some attention has been given to the related issue of the effect of pension plans on individual saving. Our model has a clear prediction here. Since a pension plan reduces the cost of delaying consumption, we would predict that the

introduction of a mandatory pension plan will increase total saving. Put another way, if saving is divided into pension saving and other saving then the introduction of positive pension saving will produce less than a complete offset in other saving.

It is difficult to test this prediction for several reasons. Most important, there is no clear neoclassical prediction to which it can be compared. Some writers have described the neoclassical prediction as a complete offset in other saving--i.e., no net effect. But as Feldstein [1977] has forcefully argued, this prediction is too simple-minded. For example, since pensions receive preferential tax treatment, the rate of return to pension saving is higher. This produces an ambiguous effect on the rate of savings accumulation. Similarly the fact that pension benefits are typically in the form of annuities may produce either more or less saving depending on which of two effects is larger. Feldstein also points out that even if the individual wants to completely offset his pension saving it may be impossible (or very expensive) for him to do so. Feldstein concludes that:

It seems reasonable to believe at this time, on the basis of the theoretical analysis and the available evidence, that private pensions increase the total asset accumulation of the covered employees but by less than the full actuarial value of the pensions.²⁴

Since our theory just predicts that the offset will be less than the neoclassical theory would predict (and that prediction is ambiguous) we can do little more than present the estimates that are currently available and let the reader judge for himself.

The first studies of this issue were done in the 1960's by Philip Cagan [1965] and by George Katona [1965]. Although they used different data sets, they produced similar conclusions. We will briefly summarize Cagan's results.

Cagan used a sample of respondents to a survey conducted by Consumers Union of its members. Saving was defined as the family's change in net worth over the year. Saving was then broken down into discretionary saving, pension saving and other contractual saving. The issue Cagan wished to investigate was the effect of pension saving on the two other categories. His results are summarized in Table 1. As can be readily seen, Cagan got the surprising result that membership in a pension plan increases other forms of saving. He attributes this increase to what he calls the recognition effect. That is, membership in a pension plan increases awareness of retirement needs, and thus increases other forms of saving.

Table 2

Average Ratio of Components of Saving to Income and
Estimated Rates of Substitution Between Them for
Covered and Not-Covered Households

Sample Households	Average Ratios (per cent)				Rate of Substitution			
	Discretionary Saving	Pension Saving	Other Contractual Saving	Total	Other Contractual Saving with Pension Saving	Discretionary Saving with	Other Contractual Saving	
Excluding those with gains or losses over \$1,000								
Covered	2.8	3.2	11.6	17.6	0.25	0.10	-0.23	
Not covered	2.3	--	9.1	11.4	--	--	-0.28	
Excluding those with gains or losses over \$1,000 and extreme saving ratios								
Covered	2.8	2.8	5.9	11.5	0.28	0.21	-0.73	
Not covered	2.1	--	5.7	7.7	--	--	-0.77	

Source: Cagan [1965] p. 21

Cagan's study has been criticized in the literature, especially by Munnell [1974]. The most troublesome problem is one of which Cagan was aware: selectivity bias. Put simply, people with a taste for saving may be more likely to work for firms which offer a pension plan. This may be a serious problem, but to our knowledge no one has dealt with it directly. Munnell also criticizes Cagan on other grounds and replicates his study using the same data. However, she uses a different measure of saving, replaces before-tax income with after-tax income, and restricts her analysis to a subset of the observations that are thought to be more reliable. Her results are summarized in Table 3.

Table 3

Effect of Pension Coverage on Ratio of
Nonpension Saving to Income

Sample	Simple Model	Model including Socioeconomic Variables
Total	-0.005 (0.7)	-0.008 (1.2)
Ages:		
30-40	-0.004 (0.4)	-0.006 (0.7)
40-54	0.001 (0.1)	-0.002 (0.2)
55-65	-0.02 (0.9)	-0.03 (1.2)

Source: Munnell [1974] p. 92 (numbers in parentheses are t-values)

The numbers in Table 3 are the coefficients of a pension dummy in a regression in which the dependent variable was the nonpension saving rate. As can be seen, the positive effect found by Cagan is no longer present, but no significant offset is found either. The only remotely significant effect

was in the 55-65 age group where those in a pension plan had saving rates lower than nonmembers by 3%. Even here the t value is only 1.2. A reasonable conclusion to draw from this study would be that (except for those near retirement) pensions increase saving by their entire value. For those near retirement the increase may be less than the total amount going into the pension plan.

These results are clearly consistent with our model. If the true offset is nearly zero then this is clearly less than any prediction based on a neo-classical model. Even the finding of a positive offset for older workers is consistent with our theory. As retirement gets closer it will become more salient and saving will become easier. (Many find it easier to diet in the spring since bathing suit weather is getting close!)

Another more recent study by Munnell [1976] finds larger offsets. However, this study has severe data limitations. The amount saved via pensions is unknown, so a pension dummy must be used exclusively, and the sample is limited to men over age 45. Furthermore the results are not robust. The estimates are reported for two different time periods and differ greatly. In fact the estimate for the latter period implies that those having pensions reduce their other saving by an amount three times the average value of pension contributions in the U.S. in that year. Also, the results change dramatically when an alternative specification is used. These problems make it difficult to evaluate Munnell's new findings.

Perhaps future work on this issue will provide further evidence that can be used to test our model. At the present time, however, we find the results supportive.

VII. Summary

We now briefly recapitulate our argument. We have investigated intertemporal choice as a problem in the economic theory of self-control. The paradox of dynamic inconsistency discussed by Strotz and others is obviously a self-control phenomenon, as Strotz's discussion seems to imply. But as the quotation that began our article states, self-control itself is paradoxical unless some kind of multi-self model of man is adopted. We have introduced self-control to the formal model of intertemporal choice by modeling man as an organization--with a planner and many doers. Conflict occurs because the doers are myopic (selfish). This conflict is fundamentally similar to the agency relationship between the employer and the employees, and individuals use many of the same strategies that organizations adopt to deal with their "conflicts of interest." These strategies can involve doer (employee) discretion while their incentives have somehow been altered or they may entail the implementation of precommitment (a rule) to avoid the doer (employee) decision process altogether.

The close correspondence between the solutions to control problems adopted by organizations and individuals provides strong support for our model. Though our model has been nontraditional, our tools have been strictly traditional. Formally, our model closely resembles that used by Ross [8] in his study of the theory of agency. Finally, we note that ours is a theory of rational behavior, just as Ross's theory is of profit maximizing behavior.

Footnotes

¹There are precedents in the literature. Adam Smith [1759] used a two-self model in his Moral Sentiments. His impartial spectator corresponds to our planner. Böhm-Bawerk [1891] discusses self-control in his Positive Theory of Capital. We cite some passages in Section 1. More recently Schelling used the notion of precommitment in a game theoretic setting in his Strategy of Conflict [1960] and has very recently proposed a two-self model quite similar to ours (Schelling [1978]). Buchanan [1975] has also discussed the rationality of self-imposed rules in a collective choice context. The planner/doer model was first proposed in an early draft of Thaler [1977].

²Strotz [1955] p. 173.

³Strotz [1955] p. 177.

⁴Böhm-Bawerk [1891] p. 253.

⁵Böhm-Bawerk [1891] p. 257.

⁶Strotz [1955] p. 178.

⁷Strotz [1955] p. 178.

⁸Strotz [1955] p. 179.

⁹McIntosh [1969] p. 122.

¹⁰In addition to Strotz, this literature includes Blackorby, Nissen, Primont and Russel [1973], Hammond [1976], Peleg and Yaari [1973], Pollak [1968], and Yaari [1977].

¹¹Strotz [1955], p. 179.

¹²It is not necessary for our doers to be completely selfish. We make this assumption because it is the extreme case and it makes the analysis simpler. However, the model can easily be generalized to allow for some consumption externalities among the doers. The only essential feature is that the doers value current consumption relatively more than the planner.

¹³For simplicity the interest rate has been suppressed. Its inclusion would add nothing of substance to the analysis.

¹⁴Note that while rules may be adopted either to save on decision-making costs or to mitigate a control problem, it is usually quite easy to distinguish between the two cases. Could institutionalized saving be explained simply as a device to reduce decision-making costs? No. If this were the case then individuals who joined such plans would expect to save the same amount, on average, as they would without the plans. But clearly people join such plans expecting to save more. Has anyone ever joined a Christmas Club because he always had too much money saved up for gift giving?

¹⁵For a discussion of agency costs see Jensen and Meckling [1976]. For a more formal treatment see Ross [1973].

¹⁶Mischel [1974] p. 287.

¹⁷See Ainslee [1975].

¹⁸This idea has been suggested by Scitovsky [1976]. The importance of norms in controlling individual behavior is also stressed heavily by Adam Smith [1759], p. 326.

¹⁹Strotz [1955] p. 173.

²⁰Strotz [1955] p. 173.

²¹This point was also made by Böhm-Bawerk. "All three causes of our underestimate of future utility...manifest themselves in extremely different degrees in different individuals, and even in the same individual at different times, according to differences of temperament and mood." Böhm-Bawerk [1891] p. 257.

²²Scitovsky [1976] has argued that food is both seductive and addictive. It is seductive in that one is tempted to start even when not hungry and addictive in that once started it is difficult to stop. He suggests that the institution of meals is used to deal with these problems. By eating only at meals one can eat until satiation is reached $((\delta Z_t = 0)/(\delta c_t))$. Fancy or expensive foods are similarly controlled through the institution of feasts.

²³Kurz, Spiegelman and West [1973] p. 30.

²⁴Feldstein [1977] p. 17.

References

- Ainslee, G., "Specious Reward: A Behavioral Theory of Impulsiveness and Impulse Control," Psychological Bulletin, Vol. 82, No. 4, 1975.
- Banfield, E. The Unheavenly City (Boston: Little Brown) 1970.
- Blackorby, C., Nissen, D., Primont, D., and Russel, R., "Consistent Intertemporal Decision Making," Review of Economic Studies, Vol. 40, No. 3, 1973.
- Böhm-Bawerk, E., The Positive Theory of Capital translated by W. Smart (London: MacMillan & Co.) 1891.
- Buchanan, J., "The Samaritan's Dilemma" in E. Phelps ed. Altruism, Morality and Economic Theory (New York: Russell Sage Foundation) 1975.
- Cagan, P., The Effect of Pension Plans on Aggregate Savings (New York: National Bureau of Economic Research) 1965.
- Feldstein, M., "Do Private Pensions Increase National Saving" National Bureau of Economic Research Working Paper #186, 1977.
- Hammond, Peter, "Changing Tastes and Coherent Dynamic Choice," Review of Economic Studies, Vol. XLIII(1), No. 133, 1976.
- Jensen, M., and Meckling, W., "Theory of the Firm: Managerial Behavior, Agency Costs, and Ownership Structure," Journal of Financial Economics, Vol. 3, 1976.
- Katona, G., Private Pensions and Individual Saving Survey Research Center, Institute for Social Research, The University of Michigan 1965.
- Kurz, M., Spiegelman, R., and West, R., "The Experimental Horizon and the Rate of Time Preference for the Seattle and Denver Income Maintenance Experiments: A Preliminary Study" SRI International Research Memorandum #21, November, 1973.
- Munnell, A., The Effect of Social Security on Personal Savings (Cambridge: Ballinger) 1974.
- _____, "Private Pensions and Saving: New Evidence" Journal of Political Economy Vol. 84, October 1976.

- McIntosh, D., The Foundation of Human Society, Chicago, University of Chicago Press, 1969.
- Peleg, B., and Yaari, M., "On the Existence of a Consistent Course of Action when Tastes are Changing," Review of Economic Studies, Vol. 40, No. 3, 1973.
- Pollak, R., "Consistency Planning," Review of Economic Studies, Vol. 35, No. 2, 1968.
- Ross, S., "The Economic Theory of the Agency: The Principal's Problem," American Economic Review, Vol. LXII, 1973.
- Schelling, T., The Strategy of Conflict (Cambridge: Harvard University Press) 1960.
- _____, Economics, or the Art of Self-Management, forthcoming American Economic Review, May 1978.
- Scitovsky, T., The Joyless Economy, Oxford University Press, 1976.
- Smith, A., The Theory of Moral Sentiments (London: A. Millar) 1759.
- Strotz, R., "Myopia and Inconsistency in Dynamic Utility Maximization," Review of Economic Studies, Vol. 23, No. 3, 1955-56.
- Thaler, R., "Toward a Positive Theory of Consumer Behavior," Graduate School of Management, University of Rochester, Working Paper No. 7717, 1977.
- Yaari, M., "How to Eat an Appetite-Arousing Cake," Center for Research in Mathematical Economics and Game Theory, Hebrew University, Research Memorandum, No. 26, June 1977.