THE BEHAVIORAL LIFE-CYCLE HYPOTHESIS

HERSH M. SHEFRIN and RICHARD H. THALER

Self-control, mental accounting, and framing are incorporated in a behavioral enrichment of the life-cycle theory of saving called the Behavioral Life-Cycle (BLC) hypothesis. The key assumption of the BLC theory is that households treat components of their wealth as nonfungible, even in the absence of credit rationing. Specifically, wealth is assumed to be divided into three mental accounts: current income, current assets, and future income. The temptation to spend is assumed to be greatest for current income and least for future income. Considerable empirical support for the BLC theory is presented, primarily drawn from published econometric studies.

I. INTRODUCTION

Modigliani and Brumberg’s life-cycle theory of saving [1954] and Friedman’s similar permanent income hypothesis [1957] are classic examples of economic theorizing. The life-cycle (LC) model makes some simplifying assumptions in order to characterize a well-defined optimization problem which is then solved. The solution to that optimization problem provides the core of the theory.

Attempts to test the life-cycle hypothesis have met with mixed success. As summarized by Courant et al. [1986, 279–80], “Just for all its elegance and rationality, the life-cycle model has not tested out very well... Nor have efforts to test the life-cycle model with cross-sectional microdata worked out very successfully.” Various alterations to the theory have been proposed to help it accommodate the data: add a bequest motive, hypothesize capital market imperfections, assume that the utility function for consumption changes over time, or specify a particular form of expectations regarding future income. These modifications often appear to be ad hoc, since different assumptions are necessary to explain each anomalous empirical result. This paper suggests that the data can be explained in a parsimonious manner by making modifications to the life-cycle theory that are quite different in spirit from those cited above, namely modifications aimed at making the theory more behaviorally realistic. We call this enriched model the Behavioral Life Cycle (BLC) hypothesis.

ECONOMIC INQUIRY

We are aware, of course, that criticizing the realism of the assumptions of an economic theory is hardly novel. It is trite to point out that few consumers are capable of making the present value calculations implicit in the theory. This remark, while accurate, does little to help formulate a better theory. Perhaps, as Milton Friedman might argue, households save as if they knew how to calculate the (after-tax) annuity value of a windfall gain. Therefore, in an effort to get beyond this sort of general critique, we suggest that the life-cycle model can be enriched by incorporating three important behavioral features that are usually missing in economic analyses. (1) Self-control: We recognize that self-control is costly, and that economic agents will use various devices such as pension plans and rules-of-thumb to deal with the difficulties of postponing a significant portion of their consumption until retirement. We also incorporate temptation into the analysis since some situations are less conducive to saving than others. (2) Mental accounting: Most households act as if they used a system of mental accounts which violate the principle of fungibility. Specifically, some mental accounts, those which are considered “wealth,” are less tempting than those which are considered “income.” (3) Framing: An implication of the differential temptation of various mental accounts is that the saving rate can be affected by the way in which increments to wealth are “framed” or described. Our model predicts that income paid in the form of a lump sum bonus will be treated differently from regular income even if the bonus is completely anticipated. Building upon the research done on these topics by psychologists and other social scientists (see, for example, Ainslie [1975] and Mischel [1981]), we are able to make specific predictions about how actual household saving behavior will differ from the idealized LC model.

The plan of the paper is first to present the model and to use it to derive propositions about saving behavior that can distinguish it from the standard life-cycle hypothesis. We then present the evidence we have been able to compile from existing studies on each of the propositions.

II. THE MODEL

Self-Control and Temptation: The Problem

In the Theory of Interest Irving Fisher bases his explanation of personal saving upon five characteristics: foresight, self-control, habits, expectation of life, and love for posterity. We concentrate here on the first three factors and the relationships among them. Foresight is important since retirement saving requires long-term planning. Self-control is necessary because immediate consumption is always an attractive alternative to retirement saving. Successfully dealing with self-control problems requires the cultivation of good habits. In presenting our model we begin with the concept of self-control.

* Professor, Department of Economics, Santa Clara University, and H. J. Louis Professor of Economics, Johnson Graduate School of Management, Cornell University. We wish to thank Franz Modigliani for providing many thoughtful comments on a previous draft of this paper. Thaler would also like to thank the Behavioral Economics Program at the Sloan Foundation for financial support.

Economic Inquiry
Vol. XXVI, October 1988, 609-643
How does self-control differ from ordinary choice? The distinguished psychologist William James ([1890] 1981, 1167) says that the key attribute of self-control choices is the “feeling of effort” that is present.

Effort of attention is thus the essential phenomenon of will. Every reader must know by his own experience that this is so, for every reader must have felt some fiery passion’s grip. What constitutes the difficulty for a man laboring under an unwise passion of acting as if the passion were wise? Certainly there is no physical difficulty. It is as easy physically to avoid a fight as to begin one, to pocket one’s money as to squander it on one’s cuppitudes, to walk away from as towards a coquette’s door. The difficulty is mental: it is that of getting the idea of the wise action to stay before our mind at all.

Incorporating the effort that is present in self-control contexts involves three elements normally excluded from economic analyses: internal conflict, temptation, and willpower. The very term “self-control” implies that the trade-offs between immediate gratification and long-run benefits entail a conflict that is not present in a choice between a white shirt and a blue one. When modeling choice under such circumstances the concept of temptation must be incorporated because of the obvious fact that some situations are more tempting than others. A model of saving that omits temptation is misspecified. The term willpower represents the real psychic costs of resisting temptation. The behavioral life cycle hypothesis modifies the standard life cycle model to incorporate these features. To capture formally the notion of internal conflict between the rational and emotional aspects of an individual’s personality, we employ a dual preference structure. Individuals are assumed to behave as if they have two sets of coexisting and mutually inconsistent preferences: one concerned with the long run, and the other with the short run.1 We refer to the former as the planner and the latter as the doer.2 To place the preceding concepts into a formal structure consider an individual whose lifetime extends over T periods, with the final period representing retirement. The lifetime income stream is given by \( y = (y_1, \ldots, y_T) \). For simplicity we assume a perfect capital market and zero real rate of interest. Let retirement income \( y_T \) be zero. Then lifetime wealth is defined as

\[
W = \sum_{t=1}^{T} y_t.
\]

Let the consumption stream be denoted by \( c = (c_1, \ldots, c_T) \). The lifetime budget constraint is then \( \sum c_t = W \).

The conflict associated with self-control is captured by the contrasting time horizons of the planner and the doer. The doer is assumed to be pathologically myopic, concerned only with current period consumption. At date \( t \) the doer is assumed to possess a subutility function \( U_t(c_t) \). We assume diminishing marginal utility (\( U_t(c_t) \) is concave in \( c_t \), and also nonsatiation (\( U_t \) is strictly increasing in \( c_t \)). In contrast, the planner is concerned with maximizing a function of lifetime doer utilities.

Since temptation depends on immediate consumption opportunities, we define an opportunity set \( X_t \) to represent the feasible choices for consumption at date \( t \). If free to choose from this set, the myopic doer would select the maximum feasible value of \( c_t \) (since that would maximize \( U_t \) on \( X_t \)). The planner would usually prefer a smaller \( c_t \). Suppose the planner wants to reduce consumption by exerting willpower. We assume that if exercise of willpower does diminish \( c_t \), there must be some psychic cost. If this were not the case, then exerting willpower would be effortless, and self-control problems such as overeating and overspending would not occur. The psychic cost of using willpower is represented by the symbol \( W_t \). \( W_t \) may be thought of as a negative sensation (corresponding roughly to guilt) which diminishes the positive sensations associated with \( U_t \). Total doer utility, denoted as \( Z_t \), is then the sum of the pleasure and the pain:

\[
Z_t = U_t + W_t. \tag{1}
\]

The doer is assumed to exercise direct control over the consumption choice, and, being myopic, chooses \( c_t \) in order to maximize \( Z_t \) on \( X_t \). This choice reflects the combined influence of both planner and doer. Willpower effort is effective if the maximizing values for \( Z_t \) and \( U_t \) on \( X_t \) are not the same.

Willpower effort can be applied in varying degrees. Therefore we define a willpower effort variable, denoted \( \theta_t \), to represent the amount of willpower exerted at date \( t \). The function \( \theta_t(c_t, X_t) \) gives the degree \( \theta_t \) of willpower effort required to induce the individual to select consumption level \( c_t \) when opportunity set \( X_t \) is being faced. The following assumptions characterize the significant features about willpower effort.

1. An increase in willpower effort is necessary to reduce consumption; that is, \( \theta_t(\cdot) \) is decreasing in \( c_t \).

2. Increased willpower effort is painful in the sense that reductions in consumption resulting from willpower are accompanied by reductions in \( Z_t \).

\( LW = \sum_{t=1}^{T} y_t \).
Specifically, $\frac{\partial Z}{\partial \theta}$ is negative, which together with the previous assumption implies

$$\frac{\partial Z}{\partial \theta} \cdot \frac{\partial \theta}{\partial c} > 0. \quad (2)$$

3. Increased willpower effort is not only painful, but becomes increasingly more painful as additional willpower is applied. Specifically

$$\frac{\partial}{\partial c} (\frac{\partial Z}{\partial \theta} \cdot \frac{\partial \theta}{\partial c}) < 0 \quad (3)$$

4. Willpower effort becomes less costly as retirement draws near. Formally, the value of the left-hand-side of (2) is assumed to be monotone decreasing in $t$.

To represent the idea that the planner corresponds to the rational part of the individual's personality, we associate a neoclassical utility function $V(\cdot)$ to the planner, with the arguments of $V$ being the subutility levels $Z_t$ through $Z_{T_p}$. Since $\frac{\partial Z}{\partial \theta}$ is negative, willpower costs are incorporated within the planner's choice problem.

Since willpower is costly, the planner may seek other techniques for achieving self-control. These techniques are the subject of the following section.

Rules and Mental Accounting: The Solution

One solution to the conflict between planner and doer preferences is for the planner to restrict future choices by imposing constraints which alter $X_t$. For example, placing funds into a pension plan which disallows withdrawals reduces disposable income and thus shrinks the doer's choice set. We refer to any precommitment device of the above type as a rule.

Suppose that the planner were able to choose a rule that completely precommitted future consumption to a particular path. Since the doers would have no choices to make, willpower effort would not be required. In this situation, the planner would choose $c$ to maximize $V$ subject to the budget constraint, while leaving $\theta = 0$. Denote this optimal choice of $c$ by $c^*$. The path $c^*$ is a first-best solution to the planner's problem and corresponds precisely to the life-cycle consumption path. Therefore, the LC hypothesis can be interpreted as a special case of the BLC model in which either willpower effort costs are zero, or a first-best rule is available to the planner.

The predictions of the two models diverge because neither of these conditions is likely to be met. The person with zero willpower costs is obviously a rarity, and first-best rules are generally unavailable. While pension plans and other saving vehicles are marketed, there is a limited selection available, and

$$D = (\frac{\partial Z}{\partial \theta} \cdot \frac{\partial \theta}{\partial c}) - \frac{\partial Z}{\partial \theta} > 0 \quad (4)$$

where $\frac{\partial Z}{\partial \theta}$ is evaluated at $\theta = 0$. The difference $D$ can be regarded as the net marginal cost of using willpower. We make the additional assumption that willpower effort is especially costly at low consumption levels, but essentially costless at high levels. In other words, $D$ decreases with $c$, and approaches zero for $c$, sufficiently large.

There are limits on the type of rules which can be enforced at low willpower costs. A reading of the psychology literature on impulse control (e.g., Ainels [1975]) suggests that effective rules must have the following characteristics. First, a habitual rule must exhibit simplicity since complex responses appear to require conscious thinking, whereas habitual responses are subconsciously guided. Second, exceptions must be well defined and rare, again in order to avoid the need for conscious responses. Third, the rule must be dynamically stable: habits are not easily altered. Both internal and external rules then are second-best; therefore, descriptive models of saving behavior must reflect the second-best solutions that are adopted by real savers.

While households' internal rules are idiosyncratic and context specific, there appear to be enough common elements to generate useful aggregate predictions. One of the most important elements concerns the decomposition of household wealth into a series of accounts called mental accounts. One simple and stylized version of a mental accounting system divides wealth into three components: current spending income $(I)$, current assets $(A)$, and

3. It needs to be emphasized that in our model, the planner can actually implement any budget feasible consumption plan by selecting $\theta$ appropriately. The only issue is at what cost. Precommitment offers the possibility of implementing a given consumption plan at reduced willpower cost.

4. M. King [1985] has criticized our characterization of the conflict between the planner and the doer as an agency problem on the grounds that there is no information asymmetry present. This criticism is misplaced. While in standard principal-agent models of the firm it is the information asymmetry that prevents the principal from achieving a first-best outcome, an agency problem can exist without information asymmetry if the principal has limited control over the agents' actions. That is the case we consider, for the reasons just described. The alternative bargaining formulation King suggests fails to capture some essential features of the problem such as the asymmetry between the strategies employed by the two parties. The planner precommits, the doer does not. The doer in our model generally does not engage in strategic behavior.

5. For more on mental accounting, see Kahneman and Tversky [1984] and Thaler [1985].
future income $F$. In the BLC, the marginal propensity to consume wealth is assumed to be account specific. This contrasts sharply with the traditional life-cycle model which treats the labelling of wealth as irrelevant because wealth is regarded as completely fungible in a perfect capital market. Specifically, traditional theory postulates that the marginal propensity to consume is the same for the following four events: a $1000 bonus received at work; a $1000 lottery windfall; a $1000 increase in the value of the household's home; and an inheritance, to be received with certainty in ten years, with a present value of $1000. In contrast, our behavioral enrichment of the life-cycle model assumes that households code various components of wealth into different mental accounts, some of which are more tempting to invade than others.

As explained below, the BLC theory postulates a specific set of inequalities in connection with the marginal propensity to consume from the preceding four wealth descriptions. The direction of these inequalities is not arbitrary, and we hypothesize that they evolved as a means of helping individuals to save. The decomposition of wealth into mental accounts constitutes an example of framing; see Kahneman and Tversky [1984]. In treating wealth as fungible, traditional life cycle theory makes an implicit frame-invariance assumption. The BLC model assumes frame dependence.

To illustrate how the three-account formulation works, consider a household that uses a pension rule which at each date deducts a fraction $s$ of income, and prohibits access to accumulated funds before retirement. See the appendix for a detailed discussion. The mental account balances at date $t < T$ are as follows:

1. The current income account, $I = (1 - s)y_t$.
2. The current wealth account $A$ (corresponding to cumulative discretionary i.e., nonpension) savings through date $t - 1$ is:
   \[ \sum_{r=1}^{t-1} [(1 - s)y_r - c_r] \]  
   \[ (5) \]
3. The balance in the future wealth account, $F$, is the sum of future income (after pension withdrawals have been made) and pension wealth $zW$.

Of course, this three-account formulation is a great simplification of actual mental accounting rules. In general, a more realistic model would break up the $A$ account into a series of subaccounts, appropriately labeled. Some households may have a children's education account, which would be treated as being similar to a future income account until the children reached college age. Also, there is some ambiguity in how households treat various changes to their wealth. Asset income, for example, is generally kept in the $A$ account, except perhaps dividend payments which may be treated as current income.6


ECONOMIC INQUIRY

Small windfalls are likely to be coded as current income, while larger windfalls are placed into $A$. We assume that pension wealth is framed as future retirement income, although some households might treat it more like current assets. Similarly, there will be variation in the way in which households treat home equity; some will treat home equity as if it were part of $F$ (and will not take out home equity loans), others as if it were part of $A$. There will be differences among households in the way they treat various accounts, and the model presented here can be considered a description of the representative household.

While the mental accounting system described above may seem bizarre to economists, it is remarkably similar to the accounting systems used by most private universities. A typical private university will distinguish between money in the “current” account which can be spent immediately, and money in the endowment. From the endowment, only income (somehow defined) can be spent while the principal must remain intact. The rules for allocation gifts to the different accounts are of interest. For example, small gifts from alumni that are part of the annual giving campaign are normally treated as “income,” spendable immediately. Larger gifts and those that are received as part of a “capital campaign” are put into the endowment account. Finally, a gift that is pledged but payable only at the time of the donor’s death is generally not acknowledged in either the income or endowment accounts, and will therefore create no increase in current spending.

Suppose next that the individual wants to save more than the maximum pension deduction rate offered to him, that is, he wants to engage in what we term “discretionary” saving. Then it is necessary to use some willpower effort in order to generate the associated additional savings, avoid depleting those savings before retirement, and refrain from borrowing against future earnings. The magnitude of the associated willpower effort costs is assumed to depend inversely on the temptation to spend. Some situations are more tempting than others. Irving Fisher associated great temptation with payday, since individuals are flush with cash. In our model we assume that the temptation to spend a (marginal) dollar of wealth depends on the location of that dollar in the mental accounting system, with current income being the most tempting, followed by current assets, and then future wealth.

Technically, we take the doer utility function $Z_1$ to be parameterized by the underlying mental accounting structure.7 Recall that marginal doer utility is given by \[ \partial Z_1/\partial s = \partial Z_1/\partial c \] (6) and reflects the cost of willpower effort at the margin. Figure 1 depicts the graph of $Z_1(c, s, X_t)$ against $c_t$ for a given mental accounting structure and

7. Formally, $Z_1$ is parameterized by the choice set $X_t$ where $X_t$ specifies the account balances $I$, $A$, and $F_t$. 

---

account balances. It reflects the essential structure imposed on the model. Consider the effects on $Z$, due to increments in $c$. We take the first marginal unit of consumption to be financed out of the $I$ account, with (4) reflecting the marginal utility of consumption. As consumption increases, the reduction in willpower effort contributes to higher utility, but in accordance with (3) at a diminishing rate. When the entire balance in the $I$ account is consumed, no willpower effort need be applied to this account. The next marginal unit of consumption is then financed out of the $A$ account.

We model the $A$ account as being less tempting than the $I$ account by assuming that as long as consumption from $A$ is zero, the self-control technology requires no willpower effort in connection with this account. However, any positive consumption from $A$ produces a fixed disutility penalty (representing an entry fee for invading the $A$ account). Consequently, the first unit consumed from $A$ is especially costly. Additional consumption from $A$ results in additional utility as willpower effort is reduced. Again this occurs at a diminishing rate. Similar remarks apply when the $F$ account is invaded.

To indicate how differential willpower effort costs for the various mental account balances can be incorporated into the model, attention is focused on the current income account, its balance denoted at the outset of date $t$ by the symbol $m_t$. When contemplating financing consumption from the current income account, $m_t$ measures the amount of temptation to be faced. We postulate that the greater the temptation, the greater the willpower effort required to choose any given consumption level $c_t < m_t$. Formally, it is assumed that at any given level of $c_t$, increased temptation will make the doer worse off, in the sense that

$$\frac{\partial Z}{\partial m_t} = \frac{\partial W}{\partial m_t} + \frac{\partial W}{\partial m_t} \cdot \frac{\partial \theta_t}{\partial m_t} > 0$$

and

$$\frac{\partial \theta_t}{\partial m_t} = \frac{\partial Z}{\partial \theta_t} \cdot \frac{\partial \theta_t}{\partial m_t} < 0.$$  

For example, consider an individual who plans to spend $1200 of his regular monthly take-home pay of $1500. The preceding inequalities suggest that were his take-home pay $2000, then stopping at $1200 would require greater willpower effort (cost). However we also postulate that

$$\frac{\partial^2 Z}{\partial m_t^2} > 0,$$

so that successive unit increments in the income account produce less of a negative impact. That is, given the intention to consume $1200 out of the income account, the impact on temptation of additional take-home pay of $500 (from $2000 to $2500) involves less additional willpower effort than the $500 increase from $1500 to $2000.\footnote{We make the stronger assumption that the left-hand side of (8) goes to zero monotonically as $m_t$ approaches infinity.}

Further details about the model and about the first-order conditions used to derive the predictions discussed below are presented in the appendix. In many ways, however, the key property of the model is the relaxation of the fungibility assumption of the LC model, and the introduction of the assumption that the marginal propensity to consume additions to wealth depend on the form in which this wealth is received. At a given date, the marginal propensity to consume is typically highest out of income ($I$), lowest out of future wealth ($F$), and somewhere in between for current assets ($A$). This implies that the BLC aggregate consumption function must incorporate at least three different income or wealth measures corresponding to the three mental accounts. That is, $C = f(I, A, F)$, where $I$, $A$, and $F$ now stand for their aggregate counterparts. The model suggests that

$$1 = \frac{\partial C}{\partial I} > \frac{\partial C}{\partial A} > \frac{\partial C}{\partial F} = 0.$$  

This set of inequalities and the other features of the model yield a series of testable predictions. It is to those predictions we now turn.

III. THE DIFFERENTIAL MPC HYPOTHESIS

Unfortunately, we know of no complete test of the hypothesis that the marginal propensity to consume differs across the three accounts in the way...
TABLE I
Saving Questionnaire

Sample: Santa Clara University part-time MBA students, N = 22.

For each of the following scenarios, please think about how you would actually behave. There are no right or wrong answers. Your responses are anonymous and confidential.

1. You have been given a special bonus at work. The bonus will be paid monthly over the course of a year, and will increase your take-home pay by $200 per month for twelve months.
   By how much would you expect your monthly consumption to increase during the year?
   ______ dollars per month. Median = $100 Total Consumption = $1200

2. You have been given a special bonus at work. It will be paid in a lump sum of $2400 (after tax) this month.
   By how much would you expect your consumption to increase in the following month?
   ______ dollars per month. Median = $400
   By how much would you expect your monthly consumption to increase during the rest of the following year?
   ______ dollars per month. Median = $35 Total Consumption = $785

3. You have been told that a distant relative has left you a small inheritance which has an after-tax value of $2400. You will not receive the money for five years. During that time the money will be invested in an interest-bearing account. After five years you will definitely receive the $2400 plus interest.
   By how much would you expect your consumption to rise this year as a result of this gift?
   ______ dollars per month. Median = $0

Table I suggests above. We have therefore conducted a small survey as a direct test of the hypothesis. A group of evening MBA students at Santa Clara University (most of whom work fulltime) was recruited to fill out a questionnaire. The questions are reproduced in Table I. Each question asks the respondent to estimate the marginal propensity to consume a windfall with an (approximate) present value of $2400. In question 1, the windfall comes in increments of $200 a month, and is most likely to be coded as regular income. In question 2, the windfall comes in a $2400 lump sum, which we hypothesize is large enough to be placed in the assets account, and should thus have a lower MPC. For question 3 the windfall is not payable for five years, and, as it will be coded in the future income account, should yield a very low MPC. The results support the differential MPC hypothesis. The median annual MPCs for the three questions are $1200, $785, and $0 respectively. These medians were the same for the whole sample as well as for the subset of ninety-three subjects that reported having at least $5000 in liquid assets, so liquidity constraints are not an issue.

While we find these intuitions of MBA students compelling, it is important to obtain evidence based on actual behavior. While there is no other direct test of the differential MPC hypothesis, there is some partial evidence. Courant, Gramlich, and Laitner [1986] distinguish between two types of wealth, current and future. Current wealth includes current income. They report being astonished by the difference in the estimated marginal propensities to consume from these two accounts, since no difference is expected in the LC framework. They estimated the MPC out of current assets to be very high, implying that households consume approximately 25 percent of their existing assets every year. They point out that this suggests a high positive subjective rate of time discount. Yet the MPC out of future wealth was found to be considerably lower, in fact suggesting a negative discount rate (p. 302).

In an earlier study, Holbrook and Stafford [1971] used a permanent income model which differentiates among different sources of income (labor income, capital income, transfer payments, etc.). However, the permanent income framework employed treated the timing of wealth as irrelevant (holding the present value constant). Consequently, the Holbrook-Stafford analysis did not distinguish among wealth which has been accumulated in the past, arrives as current income, or will arrive as part of future income. In our theory we assume that different sources of income are encoded into different mental accounts. Specifically, labor income is encoded into current income (f), while capital income (with the possible exception of dividend income, see Shefrin and Statman, 1984) is encoded into the dividend income upon arrival. Therefore we predict that the marginal propensity to consume from capital income is less than from labor income. This is what Holbrook and Stafford [1971, 16] found. The estimated MPC out of labor income was approximately 0.9, while the estimated MPC out of capital income was 0.7. Interestingly, the MPC out of transfer payments received by members of the household other than the head is approximately 30 percent, indicating that such income tends to be saved, rather than consumed.

Evaluation
While there is no complete test of the differential MPC hypothesis, the evidence that does exist is strongly supportive.

9. A similar study was conducted by Simon and Barnes (1971). Their results also support the differential MPC hypothesis.
IV. PENSIONS AND SAVING

Consider an individual who saves 10 percent of his yearly income for retirement. Suppose that total saving consists of 6 percent that is required to be put into a pension plan and 4 percent "discretionary" savings. What will happen to total saving if the individual is forced to increase the pension component from 6 percent to 7 percent? Putting aside issues of bequests, liquidity constraints, tax rates, vesting, and induced retirement, the LC prediction is that total saving will be unaffected. Discretionary saving should fall by the amount of the increase in the pension contribution, in order to preserve the choice of lifetime consumption plan c. Let PS be pension saving and DS be discretionary saving. Then the LC prediction is that \( dS/PS = -1.0 \).

The corresponding prediction of the BLC model follows.

**Prediction 1.** The change in discretionary saving with respect to a change in pension saving is less (in absolute value) than 1.0, and, for the young, will approach zero.

The intuitive explanation behind this statement is easily described. The representative household in our theory has a marginal propensity to consume from its income \( (i) \) account of nearly 1.0, but a marginal propensity to consume from its future wealth \( (F) \) account of 0. Therefore, when the pension plan transfers one dollar from \( i \) to \( F \), total saving rises by almost one dollar. Since expenditures are usually adjusted to be consistent with disposable income, the payroll deduction reduces the money readily available to spend. Then, once the pension contribution becomes pension wealth, it is off-limits to current consumption. The formal argument is more involved, including the saving behavior of the young, and is summarized in the appendix.

Prediction 1 illustrates the quasi-rational or second-best nature of the model. Our representative savers are not fools. They have genuine human weaknesses that act as constraints on the planner's maximization problem. People who join Christmas clubs, for example, probably know that they are giving up interest, convenience, and liquidity in return for external enforcement of willpower. They may judge that trade sensible if the perceived alternative is to have too little money for Christmas presents. But what would be downright stupid would be to join a Christmas club and then borrow against the subsequent payout. We believe few people are that silly. Similarly for pensions, we believe that people allow themselves to think of a pension contribution as a reduction in income in order that they do not defeat its primary purpose—the provision of income for retirement.

The model also predicts a positive relationship between wealth (income) and the magnitude of the offset.

**Prediction 2.** The change in discretionary saving with respect to a change in pension saving increases (in absolute value) with income or wealth.

This prediction arises because the cost of exercising willpower is taken to decline with income. Willpower becomes increasingly difficult to exercise when income (and therefore consumption) diminishes. Within the model, the prediction can be derived from the assumption that willpower is especially costly at low consumption levels, combined with inequality (8). Together these imply that the impact of a change in the account balance on the marginal utility of consumption falls as the account balance increases. Think about an individual who selects the maximum deduction rate \( s \), and augments his pension savings with additional discretionary saving (so that \( c_t(s) < f_t \)). Inequality (8) suggests that he will be less impacted by the last marginal increment \( ds \) than corresponding individuals with zero or minimal discretionary saving.

**Evidence**

The evidence pertaining to proposition I is substantial. The first work on this question was done over twenty years ago by Philip Cagan [1965] and George Katona [1965]. Cagan used a sample of respondents to an extensive survey of their members conducted by the Consumers Union. Saving was defined as the family's change in net worth over the year. Saving was then broken down into discretionary saving (DS), pension saving \( (PS) \), and other contractual saving. He obtained the surprising result that membership in a pension plan increased other forms of saving, i.e., \( dS/DS > 0 \). He attributed this result to what he called the recognition effect. Membership in a pension plan was thought to increase the awareness of the need to save for retirement and thus encourage other saving. Katona's study was much like Cagan's, and obtained similar results.

Cagan's study has been criticized in the literature, especially by Alicia 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 is discussed below. Munnell also criticized Cagan on other grounds and replicated his study using the same data. She used a different measure of saving, replaced before-tax income with after-tax income, and restricted her analysis to a subset of the observations that she thought were more reliable. She then regressed the nonpension saving rate on several variables, including a pension dummy. While she did not obtain the positive coefficients found by Cagan, none of the coefficients was significantly negative.

10. See also section VI below on nonproportionality, a closely related issue.

11. Another study by Munnell [1975] finds larger offsets. However, this study has some data limitations. The amount saved via pensions is unknown to a pension dummy must be used exclusively. More important, the results are not robust. The estimates reported for two different times did not greatly. The estimates 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 interpret the findings.
Two more recent papers on this issue have appeared in the Economic Journal. Francis Green [1981] used two British samples, the 1953 Oxford Saving Survey and a 1969 Family Expenditure Survey. Both data sets represent an improvement over those reported earlier since the magnitude of pension saving was available (rather than just a dummy variable for membership). However, the size of employer contributions was not available. Green used three definitions of "other saving:" (1) total saving minus pension saving, (2) other long-term saving, and (3) total saving plus durable purchases minus pension saving. Each was regressed on wealth, age, and pension saving. Once again the anomalous but ubiquitous positive coefficients were obtained. Breaking up the samples into homogeneous groups based on age or income had no effect.

Green also investigated the possible selectivity bias issue raised by Munnell. Before discussing his results, consider the logic of the selectivity bias argument. Suppose the true value of $dS/dPS$ is $-1.0$. How could selectivity bias yield estimates of (essentially) zero? The mean marginal propensity to save of those without pensions must exceed the mean marginal propensity to save of those with pensions by the average level of pension contributions. This seems implausible but possible. Now consider the range of pension benefits offered by various employers. It is even more implausible to think that these match up precisely with the average saving propensities of their employees. So Green reestimated his equations restricting his sample to those families with pensions. Again, all estimates of $dS/dPS$ were positive.

M. A. King and L. D. L. Dicks-Mireaux [1982] estimated the effect of pensions on wealth as part of a larger study. They used a 1976 Canadian data set. The estimated offset to saving resulting from an additional dollar of pension wealth (evaluated at the mean values for the sample) was either $-10$ or $-24$, depending on the definition of wealth used. While these estimates are of the "right" sign, they are clearly much smaller (in absolute value) than $-1.0$. King and Dicks-Mireaux [1982, 265] also report that the magnitude of the offset increases with wealth, and this supports our second proposition. Specifically, they state,

The estimated offset is an increasing function of wealth and at the mean values for the top decile group of the distribution of net worth the reduction in saving per additional dollar of pension wealth is estimated to be $1.00$ for social security and $0.40$ for private pensions.

Two additional studies utilize the most comprehensive data sets yet analyzed. Mordecai Kurz [1981] used the 1979 survey conducted by the President's Commission on Pension Policy. This data set has very good information (by survey standards) on pension wealth, including the value of employers' contributions. Kurz estimated the pension wealth offset to total wealth for three subsamples: male heads, female heads, and two-head families. The marginal effect was calculated at three different ages (thirty, fifty, and sixty) using two different measures of permanent income or wealth. He estimated the total offset to be between .39 and .47, again substantially different from the 100 percent predicted by the L. C. model.

Finally, Peter Diamond and Jerry Hausman [1984] used the National Longitudinal Survey, done between 1966 and 1976. Their estimates are not directly comparable to the others since they calculated the elasticity of the saving to permanent income ratio with respect to the pension benefits to permanent income ratio (rather than $dS/dPS$). This turned out to be $-1.4$, where a complete offset would again have produced an estimate of $-1.0$.

There is also a large related literature pertaining to the effect of social security wealth on saving. No attempt is made here to survey those studies, but one point about the debate between Robert Barro and Martin Feldstein is important. Barro has argued that individuals will not reduce their saving in response to an increase in Social Security benefits because they will want to increase their bequests to compensate their heirs for future tax increases. Whether or not this argument is plausible, notice that no similar argument applies for fully-funded pensions. Even unfunded pensions have intergenerational side effects only to the extent that pensions are imperfect substitutes for other bequeathable assets. Thus the fact that people do not offset increases in pension wealth suggests that similar findings in the Social Security arena are due to self-control reasons rather than intergenerational transfers. Thus Barro is likely to be proven empirically right, though for the wrong reasons.

Evaluation

The papers reported here used data sets spanning three decades and three countries. While the estimates of the offset vary between mildly positive (i.e., wrong sign) to nearly $-0.5$, in no case is the estimated offset close to $-1.0$. While selectivity bias could explain these results, we find that argument unconvincing, especially in light of Green's results using only pension recipients. (One could control for selectivity bias by studying the saving behavior of the continuing employees in a firm that charged pension benefits.) Other rationalizations of offsets less than unity have been made, but it is difficult to explain a zero (much less positive) offset within any neoclassical framework. We judge this particular set of results quite supportive of the BLC model.

12. See also Dicks-Mireaux and King [1984]. Using the same data set as in their earlier paper, they investigate the sensitivity of the pension and Social Security displacement effects to prior beliefs. They conclude that the estimates are relatively robust.

V. SAVING ADEQUACY

The essence of the life-cycle hypothesis is the idea of consumption smoothing. As stated earlier, if a time-dependent utility function is allowed, then virtually any intertemporal pattern of consumption can be reconciled with the life-cycle hypothesis and the theory becomes irrefutable. Operationally, the theory amounts to the prediction of a smooth consumption profile, so retirement consumption should equal preretirement consumption. Alternatively put, consumption in every period should equal the annuity value of lifetime wealth. The BLC prediction is the following:

PREDICTION 3: In the absence of sufficiently large Social Security and pension programs, retirement consumption will be less than preretirement consumption.

Prediction 3 is derived from the model using inequality (4), which is the formal representation of the principle that temptation induces impatience. The steeper the marginal utility of consumption function is at date $t$, the lower the resulting choice of $c_t$. If the $Z_t$ function is the same at all dates, then the absence of entry fees into $A$ and $F$ (meaning the opportunity to borrow against future wealth) guarantees that the individual would choose $c_t < c^*$, Pensions and Social Security serve two functions. They reduce the temptation to spend out of income, and they protect a portion of lifetime wealth which is earmarked for retirement. Of course if mandatory pensions plus Social Security were sufficient to keep retirement consumption up to preretirement levels then self-control problems are unlikely to be important. Thus the size of the pension/saving offset discussed above becomes crucial to the interpretation of saving adequacy.

Before reviewing the evidence on this issue it is instructive to begin with some simple facts. Nearly all retirement saving is done through some routinized program. The most important vehicles are Social Security, private pensions, home equity, and whole life insurance. The amount of discretionary saving done is qualitatively quite small. Diamond and Hausman [1984] found that half of the National Longitudinal Survey (NLS) sample of men aged forty-five to sixty-nine had wealth-to-income ratios of less than 1.5 if Social Security and pension wealth were excluded. Moreover, 30 percent had essentially zero nonpension wealth. Similar findings are reported by Kotlikoff, Spivak, and Summers [1982]. Just the fact that so much of retirement saving is achieved through institutionalized mechanisms can be regarded as support for our framework, since the recognition of self-control problems can be viewed as the reason why people want such institutions, but the high rates of institutionalized saving also make it difficult to interpret the results.

Several authors have addressed the saving adequacy issue directly, with a wide variety of methods and data. Blinder, Gordon, and Wise [1983] used the 1971 Retirement History Survey. Their analysis can be summarized (and simplified) as follows. Let $w = W/W_p$ be the ratio of current wealth at age $t$ to lifetime wealth, where $t$ is between ages sixty and sixty-five. Let $c = C/C_p$ be the ratio of the family's expected future person years of consumption at age $t$ to the expected total when the head entered the labor market. Then the ratio $\gamma = wc$ should be equal to unity if retirement saving is adequate. They estimated $\gamma$ to be .45.

Courant, Gramlich, and Laitner [1986] used the Panel Study of Income Dynamics to analyze families' consumption profiles. They found that real consumption increases over time until retirement, then decreases. They interpret this within the life-cycle model as implying negative subjective rates of time preference while young. Our interpretation is quite different. Consumption rises while young because real income (and thus temptation) is also rising. Consumption falls during retirement because (a) real income falls since most pension benefits are not indexed, and (b) the elderly grow to realize that their resources are inadequate and gradually adapt to a reduced standard of living.14

Kotlikoff, Spivak and Summers [1982] dealt with saving adequacy directly. Using the 1969 to 1973 Retirement History Surveys, they calculated the ratio $R_A = c_p/c^*$, where $c_p$ is the level annuity that can be purchased when old, given the present expected value of old age resources, and $c^*$ is the level annuity that can be purchased when young, based on the present expected value of lifetime resources. (They also calculated a similar ratio $R$ based on simple present values without annuities.) At first glance their results seem to support the life-cycle model. Over 90 percent of the sample had values of $R$ or $R_A$ of at least $.8$, many had ratios of unity or higher. However, it turns out that nearly all the wealth the elderly possess is in Social Security, pensions, and home equity.

Slightly more than one-third of couples reported levels of net worth that represent less than 10 percent of their total future resources. In addition, 67 percent of married couples hold less than 10 percent of their future resources in liquid wealth. Of these couples, 21 percent had no liquid wealth whatsoever (p. 1065).

The test of the life-cycle model then depends crucially on the pension and Social Security offsets. If these offsets are less than complete, then the saving adequacy cannot be attributed to rational saving behavior. The authors investigated this question and concluded that "in the absence of Social Security and private pensions, consumption in old age relative to lifetime consumption would be about 40 percent lower for the average person" (p. 1067).

14 In the absence of annuities, uncertainty about the length of life can also induce consumption to fall during retirement. Yet much of wealth is in the form of Social Security and pension annuities. Uncertainty about the length of life can also affect the level of wealth at retirement, but the direction is ambiguous. Two risks must be weighed: The risk of dying sooner than expected (and thus having saved too much ex post), and the risk of dying later than expected (and thus having saved too little). Our intuition suggests that most people will be more concerned with the former than the latter, and thus Blinder et al. (1983) should find $\gamma > 1$, if people are risk averse life-cycle savers.
Hanermesh [1984] also addressed the saving adequacy issue, but he used a different approach from Kotlikoff et al. He analyzed the spending patterns of retired households using the Retirement History Survey linked to Social Security records for information on income. The question Hanermesh asked was whether the elderly have sufficient income to sustain the levels of consumption they maintain early in retirement. He computed the ratio of consumption to annuitized income to answer this question. He found that consumption on average is not sustainable. In 1973, 54 percent of the retired households had consumption-to-income ratios exceeding 1.1. Since Social Security benefits represent nearly half of retirement income in his sample, Hanermesh also computed what the consumption-to-income ratio would be for various assumptions about the size of the saving/Social Security offset. If the offset is 50 percent then the average consumption-to-income ratio is around 1.5. If the offset is zero then the values climb to well over 2.0. Similar results would hold for pensions which are about another 30 percent of retirement income. Finally, Hanermesh found that between 1973 and 1975 the elderly reduced their real consumption by about 5 percent per year. This is a result similar to that obtained by Courant et al. The elderly respond to inadequate saving by reducing real consumption.

In comparing his measure of savings adequacy with Kotlikoff et al., Hanermesh made the point that consumption follows the inverted J-shaped age-earning profiles. "It may thus be more sensible to evaluate the adequacy of Social Security by comparing its ability to sustain consumption during retirement to consumption just before retirement rather than to average lifetime consumption" (p. 4). Clearly by this standard, saving is inadequate.

**Evaluation**

The saving adequacy issue is much more difficult to evaluate than the effect of pensions on saving. Some authors, i.e., Blinder, Gordon and Wise [1983], and Hanermesh [1984], judge saving to be inadequate, while others, i.e., Kotlikoff, Spiwak and Summers (1982), judge saving to be adequate. To the extent that saving is adequate, Social Security and pensions appear to be largely responsible. The fact that consumption seems to decline during retirement is consistent with the interpretation that saving has been inadequate. But this is also consistent with the fact that the expected age of death increases with age. Again it would be possible (in principle) to test the competing theories cleanly by studying the saving behavior of individuals who do not have access to pensions and Social Security, or for whom those institutions would be inadequate. An interesting case in point is professional athletes who earn high salaries for a short and uncertain period. We speculate that the typical twenty-four-year-old superstar spends more than the annuity value of his expected lifetime wealth.

**ECONOMIC INQUIRY**

VI. **NON-PROPORTIONALITY**

Wealth theories of saving are blind to levels of wealth. Consumption is smoothed, no matter what the level of permanent income happens to be. Friedman [1957] called this the proportionality principle. In contrast, our model predicts the following:

**PREDICTION 4: The saving rate increases with permanent income.**

We are not alone in rejecting the proportionality principle. In fact, our position was stated very well by Irving Fisher (1930, 72).

In general, it may be said that, other things being equal, the smaller the income, the higher the preference for present over future income... It is true, of course, that a permanently small income implies a keen appreciation of future wants as well as of immediate wants... This result is partly rational, because of the importance of supplying present needs in order to keep up the continuity of life and the ability to cope with the future; and partly irrational, because the presence of present needs blinds one to the needs of the future.

Our model simply formalizes and rationalizes Fisher's intuition; see (3), (4) and the discussion immediately following (4). The marginal cost of exercising willpower is very high at low consumption levels, but falls off as consumption increases. Therefore willpower costs fall off as income (and therefore consumption) increases. To the poor, saving is a luxury.

The evidence on the proportionality issue as of 1972 was reported in the very thorough and insightful survey by Thomas Mayer (1972). He also conducted five tests of his own. Only his conclusion is reproduced here:

There are many tests which disconfirm the proportionality hypothesis. What is even more persuasive, of all the many tests which have been undertaken by friends of the hypothesis, not a single one supports it. I therefore conclude that the proportionality hypothesis is definitely invalidated (p. 348).

When Friedman [1957] investigated proportionality, he found that it was violated, but argued that the observed behavior could be explained by measurement error. Those with high incomes might save more, he hypothesized, because their incomes have a large (positive) transitory component. Diamond and Hausman [1984] investigated this explanation using modern panel data. They regressed the saving to permanent income ratio on permanent income in a piecewise linear form. The results implied that for incomes less than $4770, each extra $1000 of permanent income raises the ratio by 3.3 percent; beyond $4770 it rises by 5.7 percent for each extra $1000, and beyond $12,076 it rises by 14.2 percent. The differences are all statistically significant.

**Evaluation**

The evidence against the proportionality principle is very strong. While the self-control hypothesis is only one of many possible explanations for the
observed rising saving rate, the results on the interaction between income and the pension saving offset (prediction 2) lend some support to our self-control based explanation.

VII. HYPERSONSITIVITY

One of the simple elegant features of the LC model is the way in which variability in income is handled. In each period (year) the consumer should consume the annuity value of his expected wealth. This statement applies whether or not the variability in income is deterministic or stochastic. Consumers are either implicitly or explicitly assumed to have some type of rational expectations, so permanent increases in income produce much larger responses in consumption than transitory increases because they lead to larger increases in wealth. Many factors are ruled irrelevant, for example the timing of the income across years and within a year (as long as there are efficient capital markets) and the form of the wealth (say human capital vs. home equity).

Our model yields three propositions that are significantly in conflict with the LC hypothesis in this general area. In this section the sensitivity of consumption to income is discussed. The following two sections concern the special cases of bonuses and windfalls.

PREDICTION 5. Holding wealth constant, consumption tracks income.
This prediction applies whether or not the variability is known (as with the age-earning profile) or unknown (as with a windfall). Formally the prediction is a consequence of the character of the planner's maximization problem. Recall that willpower effort costs are reduced by having consumption financed only out of the income account, with savings allocated directly to the asset accounts. In the first-best plan the entire income account is consumed at each date. In a second-best setting, this feature might still hold, even though some of the fluctuations in the income stream get transmitted to the consumption stream. It is just suboptimal to invade the asset accounts in order to smooth out consumption fluctuations which are not too large.

To evaluate the hypersensitivity issue it is instructive to compare some new evidence with some old evidence. Recall that Courant et al. found that consumption tends to follow the same hump-shaped pattern as the age-earnings profile. They rationalize this by attributing negative rates of subjective time preference (\( \rho \)) to the young. This rationalization seems implausible on the surface, and, more to the point, inconsistent with other evidence about individual discount rates. Friedman (1957) estimated \( \rho \) to be .4 (though he tended to use .33). Holbrook re-estimated \( \rho \) and found it to be closer to .5 than to .33. This implies a two-year horizon in the permanent income model. Holbrook [1966, 754] concluded that

ECONOMIC INQUIRY

...the shorter the horizon, the better is permanent income approximated by current income. When permanent income equals current income, the only significant special assumption of the PIH remaining is that of unitary-income elasticity of consumption. Therefore, the shorter the horizon, the smaller is the distinction between the PIH and what might be called the "current income hypothesis". In this sense, the evidence may be taken to indicate that it makes little difference which hypothesis is true, nearly the same conclusions follow from both.

Other authors that have tried to estimate \( \rho \) in other contexts have also found rates in excess of market interest rates (e.g., Hausman [1979], Gately [1980], Thaler [1981]). Together these results yield an inconsistency for the wealth model. Friedman's empirical results can only be consistent with a wealth model if people have very high discount rates, while the observed consumption patterns are only consistent with wealth theories if people have negative discount rates before retirement.

Recently, the hypersensitivity issue was examined by Robert Hall and Fredrick Mishkin [1982]. Hall and Mishkin derived the first truly rational expectations based model of consumption. They separated household income into three components: a deterministic component \( y_n \) which rises with age until just before retirement; a stochastic component \( y_r \) which fluctuates as lifetime prospects change and is specified as a random walk; and a stationary stochastic component \( y_s \) which fluctuates according to transitory influences and is described by a moving average time-series process.

Hall and Mishkin were particularly interested in the parameter \( \beta \), which is the marginal propensity to consume out of transitory income \( y_r \). The model predicts that \( \beta \) should be equal to the yearly annuity value of a dollar of transitory income. Therefore, \( \beta \) is determined by the expected remaining years of life and the interest rate. Hall and Mishkin gave some illustrative values for \( \beta \), which are reproduced in Table II. However, when they estimated \( \beta \) for food consumption using the Panel Study of Income Dynamics from 1969 to 1975 the estimated value for \( \beta \) turned out to be .29. This is consistent with the model only at interest rates higher than those given in the table. We take this to be a confirmation of the earlier Friedman-Holbrook estimates of discount rates in the .35 - .50 range. It is noteworthy that they obtain this result in spite of the use of food consumption as the dependent variable. Food consumption would seem to be less volatile than some other components of consumption. The high estimate for \( \beta \) surprised Hall and Mishkin, and this led them to consider whether other factors were at work. Upon closer examination they found that 20 percent of all (food) consumption is not explained by the LC model, and in consequence hypothesized that

15. An alternative lagged formulation yields a lower value of \( \beta \). Recently an alternative view of these results has been offered by Leaton [1983], Campbell (unpublished), and Campbell and Deaton [1987]. They argue that consumption is actually too smooth, rather than hypersensitive. Space limitations prevent us from discussing these interesting papers here.
it is set to a fraction of current income instead of following the more complicated optimal rule. This led them to point out that they were unable to distinguish this symptom of inability (or unwillingness) to borrow and lend from the type of behavior characteristic of consumers who simply face high interest rates.

In an earlier paper [1981] we pointed out that marginal rates of time preference greater than market rates of interest are consistent with our model if a self-imposed prohibition against borrowing (except to finance homes and other durables) is in effect. This hypothesized aversion to borrowing yields the same predicted behavior as the market imposed credit rationing suggested by Hall and Mishkin above. How then can the two hypotheses be distinguished? A data set with detailed financial information would allow the credit rationing hypothesis to be tested. First of all, capital market constraints cannot be binding for any family with significant liquid assets. Similarly, many families have equity in their homes or cash value in life insurance policies. These present easy credit sources. Finally, almost anyone with a steady job can qualify for some credit from banks and credit card companies. Any family that has not utilized these sources can be presumed to be unconstrained by the capital market. If the credit rationing hypothesis is correct, then the subset of families for whom the hypothesis can be ruled out should display hypersonsensitivity. In the absence of such tests, one can only guess at the relative importance of the two hypotheses. There is some evidence that individuals have unused credit sources. For example, Mark Warshawsky [1987] finds that many life insurance policy holders fail to take advantage of the possibility of borrowing against their insurance policy, even when the interest rate is lower than the rate at which the individual could invest. We think that it is unlikely that the average consumer is borrowed to the limit.

**Do the Retired Dismantle Enough?**

An interesting special case that has attracted considerable attention in the literature is the saving behavior of the retired. The LC model predicts that the retired will draw down on their wealth over time, that is, dissipate. The BLC prediction is more complicated. This model predicts that since annuity income is placed in the current income account, it will be spent more freely than the annuity value of other assets. Even though it makes sense for the retired to relax their rules that restrict access to savings, many households appear to have trouble making this transition. Thus, we predict that households will draw down nonannuity wealth more slowly than the LC prediction.

Most studies of this issue do not support the LC prediction. Indeed, investigators using cross-sectional data have found the puzzling result that the retired actually continue to save (see, for example, Davies [1981], Mincer, [1979] and the literature review in Bernheim [1987]). This result has been taken as strong evidence of a bequest motive. However, in a recent paper, Hurst [1987] criticized these cross-sectional studies and presented new evidence from the Longitudinal Retirement History Survey. Hurst found little support for a bequest motive since the behavior of households with living children was indistinguishable from childless households. He also found that retired households do dissipate. However, a question remains whether they dissipate fast enough to be consistent with the LC model. A key question in evaluating this is how to treat housing wealth. Hurst found that retired households dissave 13.9 percent of their total bequeathable (that is, nonannuity) wealth over the period 1969–79, and 27.3 percent of their bequeathable wealth excluding housing wealth. The former figure is clearly too low (by LC standards), while the latter figure might be considered reasonable. Hurst argued that excluding housing wealth was appropriate because of the costs of changing housing consumption levels. We are not convinced by this argument. While it is true that moving is costly, housing wealth can be reduced by borrowing. Typical retired homeowners have no mortgage, and thus could draw down on their housing wealth using the credit market. Their failure to do so must be considered at least partially a self-imposed borrowing constraint rather than credit rationing. Indeed, "reverse mortgages" have been offered in some areas with very little consumer response. Some direct evidence that retired households voluntarily maintain the equity in their homes is provided by Venti and Wise [1987]. They report that the elderly who sell one house and buy another are as likely to increase as to decrease housing equity, and conclude that the typical

16. To paraphrase a well-known expression, it is hard to teach an old household new rules.

17. The most important source of bias in the cross section, according to Hurst, is due to differential rates of survivorship. For example, the rich tend to live longer than the poor, so the older age groups have disproportionate numbers of the rich.

18. For example, Higbee [1986] contains information on a subsample of 770 respondents in the Retirement History Survey selected as having a head of household who was working in 1969, retired in 1971, and survived through 1979. For this group, the median mortgage was zero. See also, Sherman [1976, 72] who states: "the overwhelming majority of homeowners older than 65 are without mortgage debt—apparently because they paid it off before retiring."

19. With a reverse mortgage, the (usually) retired homeowner uses collateral in the house to borrow money from a bank. The proceeds of the loan are typically paid to the borrower in monthly payments. When the borrower dies or decides to sell the house, the loan is repaid.
elderly person who moves is not liquidity constrained. The elderly appear reluctant to consume out of their "home equity" account, even during retirement.

**Evaluation**

Individuals behave as if they had excessively high rates of discount. Nevertheless, much of lifetime consumption is successfully postponed. While credit markets do not permit massive borrowing against future income, we judge the hypersensitivity observed by Friedman and by Hall and Mishkin more plausibly explained by seit-imposed borrowing prohibitions than by market-imposed quantity constraints. Similarly, the elderly appear to spend money from their income account more readily than they draw down their assets, especially their housing wealth. Again this appears to represent voluntary behavior rather than capital market imperfections.

**VIII. BONUSES**

Define a bonus as a fully anticipated temporary increase in income. Our model then yields the following prediction.

**Prediction 6.** The marginal propensity to consume bonus income is lower than the marginal propensity to consume regular income.

This prediction reflects the combination of an assumption and a principle. The assumption is that bonus income, because it arrives as a large lump sum, is allocated to the A account, not the f account. The principle discussed in the theory section is that the marginal propensity to consume out of the income account exceeds that of the asset account.

The pooling of income into a lump-sum bonus increases saving in two ways. First, by lowering regular monthly income (relative to spreading out the bonus) the temptation to spend each month is reduced. Second, when monthly expenditures tend to be geared to regular monthly income. To set a higher level of monthly expenditures would require the individual either to borrow against the future bonus or draw down on the saved bonus during the year, each of which would violate typical mental accounting rules. Second, when the bonus does arrive, a considerable binge can occur and still permit an increase in the saving rate relative to normal. The binge occurs when the bonus is first placed into the A account. The portion of the bonus which is not consumed subsequently enters the A account, where it is more easily saved. Also, if the binge is spent on durables then some saving occurs in that way.

Bonuses are a nice illustration of a framing effect. In a standard economic model, a completely anticipated bonus is simply income with another name. Thus the distribution of earnings into income and bonus would be considered irrelevant. Our model offers the potential for increased explanatory power by considering variables, such as bonuses, about which the standard theories are silent.

The only evidence we have been able to find regarding bonuses comes from Japan. In Japan, most workers receive semiannual bonuses. Ishikawa and Ueda [1984] have studied the saving behavior of the Japanese and estimate the significance of the bonuses. Using a pooled cross-section time series approach, they estimated the marginal propensities to consume out of regular and bonus income respectively. Tests suggested pooling what they called normal years 1969–73, 1977–78, and treating the two recession-oil shock periods 1974–76 and 1979–80 separately. For the normal years they could reject the hypothesis that households treat the two sources of income equivalently. The marginal propensity to consume bonus income was estimated to be .457 while the corresponding figure for nonbonus income was .685. The difference is significant. The difference holds with durable expenditures included or excluded from consumption, though as should be expected, expenditures on durables respond much more to bonus income than to other parts of income. During 1974–76 the MPC out of bonuses jumped to over 1.0. This suggests that households used bonuses in bad years to smooth out consumption. The last period studied, 1979–80, returns to the pattern of a lower MPC out of bonus income.

Could the low MPC out of bonus income be explained by the permanent income hypothesis if bonuses are treated as transitory income? This explanation is dubious since the bonuses are fairly well anticipated. As one Japanese observer, Shiba [1979, 207] has put it,

> The trouble, however, lies in the interpretation of "transitory" income. Although they are called bonuses, they are fully institutionalized and workers expect bonuses as an intrinsic part of their normal income. Furthermore, workers can anticipate fairly well the level of bonus payments and thus a rational worker will treat them as permanent, rather than transitory, components of his income.

Nevertheless, Ishikawa and Ueda [1984] investigated this possibility directly using actual expectations data on bonus income. They used a sample of roughly 5000 workers who were asked to estimate six months in advance how large their next bonus would be. Later, actual bonuses received and consumption data were also collected. The authors then tested to see whether the respondents had rational expectations and whether they responded differently to permanent and transitory components of bonus income. The results indicated that expectations were not rational (bonuses were underestimated), but the MPC out of the transitory component of bonus income was approximately the same as the MPC out of the permanent component. Both were estimated to be .46.

The authors' conclusion about their findings is the same as ours. "First, the permanent income-life cycle hypothesis does not seem to apply to Japanese worker households...[and second] households distinguish bonus earnings from the rest of their income" (p. 2).
Evaluation

The results on bonuses are probably the hardest to rationalize within the LC framework. Similar tests would be possible in the United States if a sample of workers with and without bonuses were collected. Unfortunately, most data sets do not distinguish bonus income from normal wages and salaries.

IX. WINDFALLS

Predictions 5 and 6 together imply the following:

PREDICTION 7. (a) For (non-negligible) windfalls, the marginal propensity to consume is less than the marginal propensity to consume regular income but greater than the annuity value of the windfall. (b) The marginal propensity to consume out of windfall income declines as the size of the windfall increases.

The explanation of the first feature is basically identical to the argument for bonuses. The only difference is that the marginal propensity to consume from the windfall income is higher than for bonuses if the windfall is truly unexpected. This is because the individual has no opportunity to adjust his earlier saving in anticipation of the windfall. The explanation of the second feature is based on mental accounting. People tend to consume from income and leave perceived "wealth" alone. The larger a windfall, the more wealth-like it becomes, and the more likely it will be included in the less tempting Assets account. A corollary is that changes in perceived wealth (such as increases in the value of home equity) are saved at a greater rate than windfalls considered "income."

The best study we have found regarding actual windfalls was done by Michael Landsberger [1966]. He studied the consumption behavior of Israeli recipients of German restitution payments after World War II. What makes the study particularly useful for our purposes is that there is substantial variation in the size of the windfall within the sample. His sample of 297 was divided into five groups based on the windfall. A percent of family income. The family incomes and MPC out of total income were about the same for each group. However, as our theory predicts, the MPC out of windfall income increased sharply as the size of the windfall decreased. For the group with the largest windfalls (about 66 percent of annual income) the MPC was about 23 percent, while the group receiving the smallest windfalls (about 7 percent of annual income) had MPC's in excess of 2.0. Small windfalls were spent twice.

Evaluation

Ironically windfalls facilitate both splurges and saving. Windfalls are not treated as simple increments to wealth. Temptation matters.

X. POLICY IMPLICATIONS

The theory and evidence presented here suggest quite novel considerations for national policies regarding personal saving. Normally, when a government wants to alter the saving rate, it concentrates on changing either the level of income or the after-tax rate of return to saving. If the desire is to increase saving, then our analysis suggests that other seemingly irrelevant changes be considered. For example:

(1) A tax cut not accompanied by (complete) changes in withholding rates should increase saving more than an equivalent tax cut fully reflected in withholding. This follows because the underwithholding will yield refunds (like bonuses) should produce high saving rates.

(2) Since pensions increase saving, firms could be encouraged to offer mandatory (or even discretionary) pension plans. Requiring firms to have pension plans would have the additional benefit that future demands on the Social Security system might be reduced as the elderly begin to have substantial pension wealth.

(3) Similarly, firms could be encouraged to use Japanese-style bonuses as part of their compensation scheme. This form of payment is no more costly to firms (it might even be cheaper on a present value basis) and would, according to our analysis, increase saving.

XI. CONCLUSION

The LC model is clearly in the mainstream tradition of microeconomic theory. It is typical of the general approach in microeconomics, which is to use a normative-based maximizing model for descriptive purposes. The recent papers by Hall and Mishkin [1982] and Courant et al. [1986] are really advances in the LC tradition.

Our model is quite different in spirit. First of all, our agents have very human limitations, and they use simple rules of thumb that are, by nature, second-best. While the LC model is a special case of our model (when either a first-best rule exists or there is no self-control problem), our model was developed specifically to describe actual behavior, not to characterize rational behavior. It differs from a standard approach in three important ways.

(1) It is consistent with behavior that cannot be reconciled with a single utility function.

(2) It permits "irrelevant" factors (i.e., those other than age and wealth) to affect consumption. Even the form of payment can matter.

(3) Actual choices can be strictly within the budget set (as a Christmas club).

The relationship between the self-control model and the LC model is similar to the relationship between Daniel Kahneman and Amos Tversky's [1979] prospect theory and expected utility theory. Expected utility theory
is a well-established standard for rational choice under uncertainty. Its failure to describe individual behavior has led to the development of other models (such as prospect theory) that appear to do a better job at the tasks of description and prediction. The superiority of prospect theory as a predictive model, of course, is one of the explanations expected utility theory's value as a prescriptive norm. Similarly, since we view the LC model as capturing the preferences of our planner, we do not wish to question its value to prescriptive economic theory. The LC model has also served an enormously useful role in providing the theory against which empirical evidence can be judged. For example, the one-to-one pension offset was a result derived from the LC model (without bequests), and the numerous studies we cite were no doubt stimulated by the opportunity to test this prediction. Saving adequacy even more directly requires a life-cycle criterion of appropriate saving with which actual saving can be compared.

At times we have argued that the use of ad hoc assumptions, added to the theory after the anomalous empirical evidence has been brought forward, renders the LC model untenable. It is reasonable to ask whether our model is testable. We think that it is. Every one of the propositions we examined in this paper represents a test our model might have failed. For example, if the estimated pension offsets were mostly close to 1.0 instead of mostly close to zero, we would have taken that as evidence that self-control problems are empirically important. Similarly, the effects of bonuses on saving could have been negligible, implying that mental accounting has little to add. Other tests are also possible. Our theory suggests the following additional propositions.

**PREDICTION 8.** Holding lifetime income constant, home ownership will increase retirement wealth.

**PREDICTION 9.** The marginal propensity to consume inheritance income will depend on the form in which the inheritance is received.

The more the inheritance resembles “income” rather than “wealth,” the greater will be the MPC. Thus the MPC will be greater for cash than for stocks, and greater for stocks than for real estate.

**PREDICTION 10.** The marginal propensity to consume dividend income is greater than the marginal propensity to consume increases in the value of stock holdings.

We have not investigated the empirical validity of these propositions. We hope others who are skeptical of our theory will do so. Nevertheless, while we think that neither our theory nor the LC theory is empty, refutation is probably not the most useful way of thinking about the task at hand. It is easy to demonstrate that any theory in social science is wrong. (We do not believe that individuals literally have planners and doers, for example.) Negative results and counterexamples must be only a first step. This paper is intended to be constructive rather than destructive, and to show that the consideration of self-control problems enables us to identify variables that are usually ignored in economic analyses but which have an important influence on behavior.

**APPENDIX**

The model in Section II features three mental accounts—current income I, the asset balance A, and future income F. For ease of exposition we have described the structure of the account balances in Section II as if the F account is never used to finance current consumption. This is why pension wealth (sLW), against which borrowing is prohibited, can be placed into the F account without requiring an account of its own. In the discussion below, we allow the possibility that the F account may be invaded, but implicitly assume that the entry penalty attached to F is sufficiently high so as to deter frequent entry. We also assume that the entry penalty into the F account is larger than the entry penalty into the A account. Therefore, the individual would never wish to borrow from the F account while the A account has a positive balance. A more complex mental accounting formulation is required to model the issue of simultaneous borrowing and holding of liquid assets which we address in Thaler and Shefrin [1981]. For instance, some households take out automobile loans, despite having a positive balance in a saving account earmarked for their children’s education.

As in Section II the income account balance at any nonterminal date t ≤ T is defined as It = 1−t)y, although we defined the asset balance by (5) in the text, a technical qualification is required when the F account is invaded, since the value of (5) is negative in this case. When the asset account is empty at date t, then we wish to have At = 0, with any further borrowing being financed out of the F account. In this regard, denote the value of (5) by a. Then a is defined as (max{0, a}). A similar qualification is required for F, We define F, as:

\[ F = \sum_{t=1}^{T} (1-\delta)y_t + sLW + \min(B, 0) \]

where the \( \min(B, 0) \) term reflects eventual required repayment on past borrowing.

The propositions which underlie the empirical portion of the paper follow from the optimality conditions that characterize the planner’s choice of \( c_t \) and \( s_t \). The first order conditions associated with \( c_t \) concern the marginal utility to the planner from an additional unit of \( c_t \). This is given by

\[ \delta U(\beta, \delta) = \sum_{m=1}^{T} \delta U(\beta, \delta) = \sum_{m=1}^{T} \delta U(\beta, \delta) + \delta U(\beta, \delta) \cdot \delta \theta(\delta) \cdot \beta(c_t) \]

\[ (11) \]
with \( c_t(c_s) \text{ equal to one if either the } A \text{ account or } F \text{ account has been invaded at date } t \), and zero otherwise. While the first term in the above sum is the direct utility associated with \( c_t \), the second term reflects the reduced temptation effect associated with future consumption from the A account prior to T. This marginal utility is to be compared with the marginal utility of retirement consumption:

\[
\frac{\partial U}{\partial Z_t} \cdot \frac{\partial Z_t}{\partial c_t} \tag{12}
\]

The optimality conditions require that when (12) exceeds (11), consumption at \( t \) is reduced and transferred to \( T \) through increased discretionary saving. However, if (11) exceeds (12) we need to consider two cases. In the first case the account being used to finance \( c_t \) has not been drawn down to zero. Then \( c_t \) should be increased. If the financing account has been drawn down to zero, then attention needs to be paid to whether invading the next account becomes worthwhile. If not, then (11) will exceed (12) at the optimum. We refer to the condition (11) = (12) as the Fisher condition (equalization of marginal utilities) and (11) > (12) as the generalized Fisher condition. The second type of optimality condition is associated with the selection of the pension deduction rate \( s \). With \( c \) given, the impact of a marginal change in \( s \) is through the temptation effect. When \( c_t < I_t \), the net benefit at \( t \) from a marginal increase \( \Delta s \) in \( s \) is

\[
- \frac{\partial U}{\partial Z_t} \frac{\partial Z_t}{\partial \delta m_t} + \frac{\partial Z_t}{\partial \theta_t} \cdot \frac{\partial \theta_t}{\partial \delta m_t} y_t \tag{13}
\]

When \( c_t \) is financed out of the A account or F account, there is also a temptation impact due to the amount of willpower effort needed in connection with \( c_t \). It has the same general form as (13). However, this effect is small compared to the discrete effect which occurs when the increment \( \Delta s \) forces the invasion of the A (and/or F) account since this entails the entry penalty. Consequently, the choice of \( s \) will essentially balance off the lowered temptation costs in the F account against the additional entry penalties for invading the A (and/or F) account.

An implication of the model is that an increase in the pension saving rate will increase retirement savings. Consider the formal argument for this statement. Begin with the case in which no pension plan is available (so that the maximum deduction rate \( s^* \) is zero), and let a small pension plan be made available (\( \Delta s > 0 \)). Let the household contemplate increasing its deduction rate by \( \Delta s \), \( \Delta s < \Delta s^* \). Consider how total saving in our model is impacted by the marginal increase \( \Delta s \). Let \( c(s) \) be the planner's optimal choice of \( c \), given \( s \). If the pension deduction does not cause the household to become liquidity constrained, then the LC prediction is that \( c(s) \) is invariant to the choice of \( s \). Suppose that the increment \( \Delta s \) does not alter the account used to finance the representative household's marginal (i.e. last) unit of consumption at any date. For instance, if at date \( t \) the individual was consuming only out of \( I \) (prior to \( \Delta s \)), then it will continue to do so after \( \Delta s \). Recall that the remainder \( \Delta s \) in \( s \) shifts wealth into the F account from the I account. Suppose that \( c_I < I_t \), so that date \( t \) consumption is financed solely from the income account. Observe that inequality (8) implies that the impact of \( \Delta s \) is to cause a decrease in the marginal temptation to consume at level \( c_I(s) \). However, the marginal utility of retirement consumption \( c_I(s) \) remains unchanged. Therefore \( \Delta s \) causes the marginal utility of \( c_I(s) \) to fall below its retirement counterpart, thereby leading date \( t \) consumption to be decreased in response. Consequently unlike the LC prediction, \( c_I(s) \) is nonconstant in \( s \). If date \( t \) is typical then lifetime saving \( c_I(s) \) rises with \( s \). We regard this as the representative case.

There are other cases to be considered as well.

1. If consumption \( c_I(s) = I_t \) (and we continue to consider the case when \( \Delta s \) does not induce the invasion of \( A \) or \( F \) account, then date \( t \) consumption falls simply because \( I_t \) falls with \( \Delta s \).

2. When \( c_I(s) \) is financed out of the A account, then the marginal temptation hypothesis applied to \( m_t = A_t \) (or \( m_t = F_t \)) implies that \( c_I \) declines with \( \Delta s \).

3. However, when the \( c_I(s) = I_t \) and the individual is indifferent to invading \( A \), then the increment \( \Delta s \) actually induces an increase in \( c_I \) as \( A_t \) gets invaded. This situation is typical for choices of \( s \) which are greater than optimal.

Under the hypothesis that the pension deduction rate begins below the optimal levels, so that \( \Delta s \) is considered an improvement, we predict that lifetime saving (meaning retirement consumption \( c_I \)) rises with \( \Delta s \).

In Prediction 1, the zero offset for the young results from two assumptions. Assumption 4 indicates that discretionary (retirement) saving is most difficult for the young (low \( t \)). Hence the young typically have asset account balances which are (virtually) zero. We also assume that the \( F \) account is inaccessible. Leading institutions do not issue loans to young people against future pension income. Therefore the introduction of a pension causes consumption for the young to fall by the amount of the pension deduction.

REFERENCES


