Time-Inconsistent Preferences in Adam Smith and David Hume

Ignacio Palacios-Huerta

Recent research in economics and psychology reveals that people and animals may evaluate choices inconsistently at certain times or in some areas of their lives. This inconsistency is often characterized, for example, by choosing the smaller and earlier of two alternative cash prizes when both are near, but changing to the larger, later prize as they draw more distant. The persistence and robustness of such observations have led some economists and psychologists to think "that the problem may not come from some extraordinary condition that impairs the normal operation of intentionality, but rather from the process by which all people, perhaps all organisms, evaluate future goals" (Ainslie and Haslam 1992, 58).

Evidence supports the postulate that the *basic* temporal discount function of human beings and some animals is such that preference reversals may naturally arise unless some actions are taken to avoid them. Preference reversals have been explained by the tendency for discounting to increase as the time delay diminishes. Indeed, several experiments and

real-life experiences indicate that a discount curve more concave than an exponential curve may govern the subjects' choices in the situations studied. Discount functions may thus be approximated by a hyperbola; that is, rewards t periods in the future are discounted by $1/(k_1 + k_2 \cdot t)$, where the k_i's are constants.[1] A function more concave than the usual exponential curve β^t produces intertemporal conflict, causing preferences to change between a given pair of alternatives as time elapses. This is because such a function discounts more heavily than the exponential function for events in the near future, but less heavily for events in the distant future. Therefore, preferences of decision makers with hyperbolic discount functions are dynamically inconsistent.

Over the last few years a large literature has developed that studies various "behavioral anomalies." In particular, there is a growing body of literature that studies the behavior of economic agents with hyperbolic discount functions, as well as the implications of such behavior.[2] In fact, some economists and psychologists argue that hyperbolic discounting may explain some of the behavioral anomalies that have been documented during the last decades.[3]

However, despite the apparent novelty and sense of discovery with which this behavior is treated in the literature, the analysis of dynamically inconsistent behavior is not new. This behavior was analyzed by Hume ([1739] 1978) and Smith ([1759] 1976), and later by Malthus ([1826] 1986), Jevons (1871), Böhm-Bawerk ([1891] 1971), Marshall ([1890] 1920), and Pareto (1909) in their discussion of intertemporal trade-offs. Although foreseeable changes of preferences were not formalized analytically until Strotz (1956), and later further developed in the articles of the psychologist Ainslie (1975, 1986, 1992), the

1. This observation was made by Herrnstein (1961) with regard to animal behavior, and later by De Villiers and Herrnstein (1976) in relation to human subjects. Although some small corrections to the hyperbolic structure have been suggested by further evidence, the basic shape has remained unchallenged. Gul and Pesendorfer (2001) propose an alternative approach where temptation, rather than a preference change, may help explain the observed evidence.

2. See Laibson 1997, Brocas and Carrillo 2000, Caillaud and Jullien 2000, and the references therein. Strotz (1956) was the first economist to study dynamically inconsistent preferences in a formal model. Pollak (1968), Peleg and Yaari (1973), and Goldman (1980) extended Strotz's work, arguing that when preferences are dynamically inconsistent, dynamic decisions should be modeled as an intrapersonal game among different temporal selves (i.e., "me today" is modeled as a different player from "me tomorrow"). Elster (1984), Schelling (1984), Thaler and Shefrin (1981), and Posner (1997) also discuss the importance of conflicting selves within the same person.

3. See, for example, Ainslie 1992, Loewenstein and Prelec 1992, and Rabin 1999.

pervasiveness of this behavior was studied in previous centuries. The purpose of this article is to examine the insights established by Adam Smith and David Hume within the context of dynamically inconsistent behavior. In particular, it shows how their analyses of this behavior remain novel, despite much progress in the literature during the last decade. The essay also provides a formal assessment of their contributions.

The implications of a hyperbolic discount function are important for economics as well as for all other social sciences. One of the main implications is that it implies an overvaluation of an imminent reward relative to a larger, later reward. Consequently, in order to attain their long-range goals, individuals prefer to compensate for their inconsistencies by *restricting* in part their own current and future choices. The most apparent way for an individual to forestall his change in optimal choices or preferences is to adopt some type of *commitment* device: "He may 'precommit' his future behavior by precluding future options so that it will conform to his present desire as to what it should be. Or, alternatively, he may modify his chosen plan to take account of future disobedience, realizing that the possibility of disobedience imposes a future constraint on the set of plans that are attainable" (Strotz 1956, 166).

How do individuals commit their future behavior? Current literature has focused on various commitment mechanisms that allow individuals to restrict future options, and hence to overcome the solicitations of certain passions and impulses. These include irrevocable contracts, compulsory savings plans, self-imposed deadlines, Christmas clubs, pension plans with early withdrawal penalties, drugs that suppress the appetite, and other constraints. An individual may also decrease the attractiveness of a particular activity, good, or object by avoiding the acquisition of information about it, by avoiding direct sensory contact with it, and when faced with the problem of Ulysses' sirens, by asking his fellow sailors to tie one to the mast and to put wax in their own ears.[4] Commitment mechanisms can also be generated within certain social groups and institutions like firms, schools, religious groups, families, marriages,

---

4. Carrillo and Mariotti (2000) formally examine the effects of avoiding the acquisition of information about certain goods and activities. The example of Ulysses is formally equivalent to placing an alarm clock on the other side of the bedroom so that one can party or stay up late tonight and still force oneself to wake up early tomorrow; to going window shopping without carrying any means of payment so that no money is spent for sure; to buying only one unit of a consumption good when buying three units is cheaper per unit; and to other similar commitment strategies that preclude deviation from today's optimal choices, even if some passions, impulses, and temptations arise to the contrary.

and communities. The reason is that contacts and relations with spouses, family members, friends, and members of the same firm, religious group, or community may implicitly and explicitly generate a system of incentives (e.g., reputation, credibility) that help to precommit future actions in an intertemporal setting.

Modeling the individual as a sequence of different selves, Laibson (1997) solves for the subgame perfect equilibrium of the problem of a hyperbolic discounter who has access to a "commitment technology" that consists of an illiquid asset which has to be sold one period before its proceeds arrive. This technology is analytically equivalent to an irreversible decision whose effect will take place one period after the decision is made. In this sense, it may also include certain contracts (e.g., marriage), goods (e.g., housing), and investment in human capital assets whose high transaction costs may make them serve as effective commitment mechanisms over time. This type of illiquid asset or irreversible choice formally captures the essential features of all of the commitment mechanisms mentioned above.

An important characteristic common to all of the mechanisms examined in the recent literature is that they are *external* commitments. The reason is that they do not act both directly and exclusively through preferences. Instead they operate through the "budget constraint," by limiting or precluding some choices, or indirectly through preferences as in the case of reputation or credibility or when taking a drug to suppress one's appetite, instead of having or developing the willpower to behave in a certain manner or not to eat in excess.

Interestingly enough, these kinds of commitments are not the type of "solutions" to the problem of intertemporal inconsistencies that are emphasized by Adam Smith and David Hume in their analysis of certain "lower" passions. As will be shown, their arguments can be interpreted as being concerned with a class of mechanisms that can be denominated *internal* commitments, since they act both directly and exclusively through preferences. They operate as follows. It is clear that an imminent reward can be countered effectively when a deliberate action is taken to increase its cost or to exclude it from the set of potential choices, as is the case with the external commitments emphasized by current research. But imminent rewards, Smith and Hume argue, can also be countered effectively by competing rewards, even nonmaterial ones (sentiments), provided they too are imminent. Human preferences may indeed have a natural way of generating such competing rewards to avoid valuing

current material payoffs "excessively." In particular, these rewards include *current* effects of both *past* events and *future* events. When past events (through memory, habits, culture, traditions, and the like) and imagination of future events (current savoring and dread of future experiences) have an impact on current utility, they at least partially commit current and future choices, and hence help individuals to behave in a time-consistent fashion. Imagination of future experiences as well as "habits, traditions, and other preferences that are directly contingent on past choices partly control, and hence commit, [current and] future behavior in predictable ways. Indeed, habits and the like may be very good substitutes for long-term contracts and other explicit commitment mechanisms" (Becker 1992, 338). These nonmaterial competing rewards act as commitment mechanisms whose effect is direct and exclusive through preferences.[5] In this sense they can be defined as *internal* commitments.

Adam Smith and David Hume study the implications of passions for the formation of human preferences. In particular, they study how sentiments derived from past and future experiences may be developed in order to overcome the overvaluation of imminent rewards that is induced by certain passions or lower-order preferences. In other words, they study how certain sentiments can make human reward mechanisms more sensitive to future rewards and penalties in selected instances. They rightly viewed such sentiments, including moral sentiments, as a fundamental part of the discounting mechanism.[6]

This article is concerned with their analysis of these internal commitments. Interestingly enough, the role of sentiments has been ignored by recent developments in the literature on time-inconsistent discounting and other behavioral anomalies.[7] Indeed, neither the role of past experiences nor the role of future ones has been incorporated into formal

5. Hirschman (1977) argues how *material* competing rewards can act as commitment mechanisms that can effectively counter some passions of man. He examines how the most important economists and philosophers of the seventeenth and, especially, eighteenth centuries were greatly concerned with how some passions of man "conspired" against the general progress of mankind. He discusses in detail the intellectual transformation that occurred, wherein the pursuit of material interests—so long condemned as the deadly sin of avarice— was assigned the role of containing some unruly and destructive passions of man. This idea that some passions could be usefully restrained by a contrary impulse (material interests) is considered a relevant determinant of the emergence and development of commercial society.

6. See also Malthus [1826] 1986 on moral restraints as internal commitment mechanisms.

7. The only exception in the literature is the recent work by Benabou and Tirole (2000). Their analysis, however, involves no endogenous formation of preferences or sentiments.

intertemporal models with time-inconsistent discounting.[8] As will be shown, the theory of sentiments as internal commitments developed by Smith and Hume provides a simple and robust theory of the endogenous formation of time-consistent behavior and, therefore, of the endogenous formation of preferences. Various authors consider that "most of us are born with hyperbolic discount functions" (Strotz 1956, 177) and that "consistent behavior is sometimes acquired, to a greater or lesser extent, as a *skill*" (Ainslie 1992, 57). This article will demonstrate how Smith and Hume establish the process by which consistent discounting behavior may be acquired as a skill, and how this skill is embedded in human preferences. In this sense, the analysis can be interpreted as indicating the type of fundamental features that preferences may have so that, even if future utility is discounted in a hyperbolic way and intertemporal inconsistencies are induced by certain lower passions, no changes in optimal choices will occur as time elapses. In other words, it examines a theory where intertemporally consistent behavior can be "learned," developed, or acquired when initially endowed with a time-inconsistent discounting structure.[9]

The rest of the article is organized as follows. Section 1 presents the basic problem of intertemporal consistency in the simplest possible analytical form. Section 2 shows how Adam Smith and David Hume

8. The analysis of temporal nonseparabilities, however, is not absent in the modern economics literature. During the last couple of decades, the important role of habits, traditions, and other forms of past experiences has been successfully incorporated, for example, into models of consumption, asset pricing, addiction, and others (see Becker 1995, 1996; Campbell and Cochrane 1999; and the references therein). Imagination of future events is also a source of utility, as waiting for consumption or events to occur can often be gratifying, pleasurable, painful, or distressing. In this sense, imagination also partly controls—and hence commits—present and future behavior. However, except for the contributions of Loewenstein (1987) and Meghir and Weber (1996) to consumption theory, the role of pleasure and pain derived from future events has not been incorporated into formal models in economics. See also the applications of Epstein and Hynes (1983) to some basic problems in dynamic economic analysis, and the discussion in Elster and Loewenstein 1992. Kimball (1987) and Hori (1992) take the past and the future into the current utility function in the context of intergenerational altruism.

9. At this point, it is important to attend to the fact that the way in which David Hume and Adam Smith consider the possibility of the control of a "lower" set of passions and impulses by a "higher" set was already explored in the previous century by French Jansenist theologians and jurists such as Pierre Nicole and Jean Domat. David Hume became familiar with this work during his stay in France in the 1730s. John Locke translated some of Pierre Nicole's work into English (e.g., Locke 2000). Pierre Nicole's analysis on the weakness of will is notably relevant. However, he is more pessimistic than Adam Smith and David Hume, as he often emphasizes, in his analysis of the weakness of man, his view of man's inability to allow reason to control passions.

emphasize the natural formation of certain sentiments as internal commitments to resolve the problem of intertemporal inconsistencies. It then offers a basic analysis of the extent to which current sentiments derived from past experiences and from imagination of future experiences can mitigate or fully solve an individual's time-consistency problem. Section 3 provides an extension of their analysis to the case of choice under uncertainty. This extension confirms Adam Smith's insight that individuals systematically overvalue good states of nature and undervalue bad ones, a phenomenon that can be thought of as rational wishful thinking. Section 4 concludes with a summary and some final remarks. An appendix provides formal proofs of the results in section 2.

1. The Basic Problem of Time Inconsistency

The basic problem of a time-inconsistent discounter can be formulated in an intuitive way using a quasi-hyperbolic discounting structure, as is typically done in the current literature. Following the working definition of Laibson (1997), consider the basic problem of an individual who maximizes the following preferences:

$$u(c_0) + \delta \sum_{t=1}^{T} \beta^t u(c_t), \tag{1}$$

with $u' > 0$, $u'' < 0$, $\beta \in (0, 1)$, $\delta \in (0, 1)$, for deterministic consumption sequences beginning at time 0, $\{c_t : t = 0, \ldots, T\}$, and subject to a given intertemporal budget constraint. Note that the discounting structure $\{\delta\beta^t : t = 1, \ldots, T\}$ is quasi-hyperbolic: immediate future period $t = 1$ is discounted by $\delta\beta$ whereas immediate future periods after period 1 are discounted by just β. This function approximates a hyperbola and delivers the intuition for the problem of time-inconsistent discounting in a simple fashion (Laibson 1997).

At time $t = 0$ this representative individual will select his optimal consumption sequence. This sequence is obtained by equating the marginal rates of substitution (*MRS*'s) between consumption at adjacent times to one plus some rate of interest r which, without loss of generality, is considered to be constant:

$$\frac{u'(c_0)}{\delta\beta u'(c_1)} = \frac{u'(c_1)}{\beta u'(c_2)} = \cdots = 1 + r. \tag{2}$$

At time $t = 0$, the individual will consume the amount c_0 selected from these first-order conditions. The question is, when time elapses from $t = 0$ to $t = 1$, will he consume the amount c_1 selected from these conditions? The answer is no. When he arrives at period 1, he will be tempted to consume *more* than originally planned. Note that when the problem is evaluated at time $t = 1$ the *MRS*'s change: the *MRS* between consumption at times $t = 1$ and $t = 2$ as seen from period 0 was $u'(c_1)/\beta u'(c_2)$, whereas as seen from period 1 is $u'(c_1)/\delta\beta u'(c_2)$. Therefore, the optimal consumption sequence from the perspective of time 0 is *not* optimal anymore at $t = 1$. More precisely, because the *MRS*'s are always equal to $1 + r$, c_1 becomes *more urgent* when $t = 1$ arrives than as of period 0. At $t = 1$, the individual will want to rebalance his optimal consumption plan and consume a *greater* amount of c_1 than originally planned at 0 (thereby decreasing marginal utility at time 1) by substituting away from all future consumptions after period 1 (which in turn slightly increases the corresponding marginal utilities) and reestablishing the equality of the different *MRS*'s to $1+r$. This time-consistency problem only arises because time has elapsed from 0 to 1 in the presence of a nonexponential discounting structure. Thus, the parameter δ may be interpreted as an impulse that arises when period 1 arrives and that modifies the *MRS* that period. The same inconsistency problem occurs in period $t = 2$ with respect to the optimal consumption sequence planned at $t = 1$, in period 3 with respect to the plan in period 2, and in all subsequent periods with respect to the previous period.

Knowing that this deviation from today's optimal choices will arise in the future unless some actions are taken to avoid it, the individual has an incentive to constrain his future choices by adopting some kind of "commitment technology" in order to force future selves to behave as it is optimal from today's perspective. As discussed earlier, the modern literature has focused on external commitment mechanisms, that is, on self-imposed restrictions that will limit consumption above and beyond the one originally planned and that will preclude yielding to impulses.

However, Adam Smith and David Hume focused on the endogenous formation of certain internal commitments or sentiments. These sentiments, as will be shown, can be interpreted as referring to certain characteristics that the utility function $u(.)$ may exhibit so as to eliminate or reduce the extent of time-inconsistent choices over time, *even* if the future is discounted in a nonexponential fashion and passions and impulses arise over time. In particular, they argue that current sentiments derived

from past and future experiences are naturally developed in order to over-come, at least in part, the overvaluation of imminent rewards. These sentiments make the human reward mechanism more sensitive to future rewards and penalties in selected instances. They refer to specific instances given that certain "lower"-order passions appear to be dependent upon activity, goods, and situation. The reason is that high rates of discounting are more often observed, for example, when a person comes into direct sensory contact with an object or in situations in which a person is driven by appetites like hunger and sexual desire. Such sentiments, including moral sentiments, are a crucial part of the discounting mechanism.

Next, we turn to the specific *internal* commitment technologies or sentiments emphasized by Adam Smith and David Hume that human nature can develop, consciously or unconsciously, in order to partially or fully avoid the changes in optimal choices induced by passions and other impulses over time.

2. Recursive Preferences and Temperance in Adam Smith and David Hume

In order to understand how sentiments can be formed to curb and restrain certain lower-order preferences, it is first convenient to understand how Adam Smith views the nature of human preferences. Although he devotes most of *The Theory of Moral Sentiments* to this matter, in the following quote he precisely addresses the aspect that is the focus of this paper.

In his description of the different accounts that were given on the nature of virtue, Smith's thoughts on the propriety of conduct agree with Plato's account of the nature of virtue. Plato argues that the soul is composed of two different faculties or orders, and Smith describes it as follows:

> The first is the judging faculty, the faculty which determines not only what are the proper means for attaining any end, but also what ends are fit to be pursued, . . . the faculty by which we judge of truth and falsehood, . . . that by which we judge of the propriety or impropriety of desires and affections. [Second,] the different passions and appetites which are so apt to rebel against their master, [are divided into] two different classes or orders. The first consisted of those passions, which are founded in pride and resentment, or in what

schoolmen called the irascible part of the soul; ambition, animosity, the love of honour, and the dread of shame, the desire of victory, superiority, and revenge. The second consisted of those passions which are founded in the love of pleasure, or in what the schoolmen called the concupiscible part of the soul. It comprehended all the appetites of the body, the love of ease and security, and all the sensual gratifications. It *rarely* happens that we break in upon that plan of conduct, which the governing principle prescribes, and which in all our cool hours we had laid down to ourselves as what was most proper for us to pursue, but when prompted by one or other of those two sets of passions; either by ungovernable ambition and resentment, or by importune solicitations of present ease and pleasure. But though these *two orders of passions are so apt to mislead us, they are still considered as necessary parts of human nature:* the first having been given to defend us against injuries, to assert our rank and dignity in the world, to make us aim at what is noble and honourable, and to make us distinguish those who act in the same manner; the second, to provide for the support and necessities of the body. The first set of passions was of a more generous and noble nature than the other. They were considered upon many occasions as the auxiliaries of reason, to check and restrain the inferior and brutal appetites; the irascible part of our nature is called in to assist the rational against the concupiscible. (267–68; emphases added)

This analysis emphasizes the fact that impulses or passions leading to an overvaluation of the present are a natural, inherent, and *necessary* feature of human preferences. This is in sharp contrast with current literature on this matter that considers *all* passions and impulses as faulty faculties. Clearly, the pervasiveness of impulse and control problems that time-inconsistent discounting implies raises the question of how natural selection could have produced a psychological reward mechanism that gives "too much" weight to immediate material rewards. Some biologists, psychologists, and economists argue that it is not difficult to imagine why natural selection might have favored a reward mechanism that gave heavier weight to current payoffs than one that gave uniformly greater weight to the future: organisms that do not survive the current moment and do not get past immediate threats to survival may simply have no future. This explains why time-inconsistent discounting is considered to be dependent upon activity, goods, and situation. Indeed,

high rates of discounting are typically observed, for example, when driven by appetites like eating, drinking, and sexual desire or when coming into direct sensory contact with an object. The same scientists, however, have also observed certain experimental and real situations in which it is not optimal to discount the future hyperbolically. They ask then, why can't the reward mechanism discriminate more finely and give more weight to future payoffs *only* in those cases where it would be advantageous to do so? Surprisingly enough, in response to such questions, it is often simply pointed out that "natural selection and the solutions that we see in nature, while sometimes elegant, are more often clumsier than those a purposeful engineer would have come up with" and that "there may be sharp limits on the extent to which nature can fine-tune the reward mechanism" (Frank 1988, 90). Ainslie (1992) and Herrnstein (1997) suggest, for example, that heavy discounting might have been favored because of cognitive limitations.

Smith, however, argues that while the irascible and concupiscible faculties are apt to mislead us, they are also *necessary* parts of human nature. Moreover, the first set of passions (the irascible part of our nature) are considered as the *auxiliaries of reason* to assert our dignity in society, to determine and pursue noble and honorable ends, and to assist the rational against the concupiscible. Only the "inferior and brutal appetites" appear in his view to be "importune" and "misleading," although to some extent they are still considered necessary for the support and necessities of the body. In terms of the formal model studied in the literature described in the previous section, Smith's description of the irascible and concupiscible passions can be interpreted as the extent to which discounting is not exponential, whereas the judging faculty, our reason or governing principle, can be interpreted in a temporal setting as an exponential discounting structure under which neither passions nor impulses arise over time.

Smith then examines the relationship between the different faculties. The composure and harmony of the soul in which the different faculties were in perfect concord with one another, when neither the irascible nor the concupiscible passions ever aimed at the gratification of which reason did not approve, constituted to Smith the virtue of "temperance." Before analyzing how temperance gets to be formed—and thus how human nature may be perfectly equipped to "deal" with passions that induce dynamic inconsistencies—Smith continues to analyze some aspects of the nature of human preferences by discussing Epicurus's ancient system

in which virtue consists of prudence. He expands upon this system, but agrees with Epicurus in the intrapersonal nature of preferences. He makes an important distinction between the effects of the rewards experienced by the body and those experienced by the mind. This distinction, which can be interpreted as an analysis of the recursive nature of preferences with respect to past and future experiences, will later be crucial for his analysis of sentiments as internal commitment mechanisms:

> But the pleasures and pains of the mind, though ultimately derived from those of the body, were vastly greater than their originals. The body felt only the sensation of the *present* instant, whereas the mind felt also the *past and the future, the one by remembrance, the other by anticipation*. . . . When we are under the greatest bodily pain, . . . we shall always find, if we attend to it, that it is not the suffering of the present instant which chiefly torments us, but either the agonizing remembrance of the past, or the yet more horrible *dread* of the future. The pain of each instant, considered by itself, and cut off from all that goes before and all that comes after it, is a trifle, not worth the regarding. Yet this is all that the body can ever be said to suffer. In the same manner, when we enjoy the greatest pleasure, we shall always find that the bodily sensation, the sensation of the *present* instant, makes but a *small* part of our happiness, *that our enjoyment chiefly arises* either from the cheerful *recollection of the past,* or the still *more joyous anticipation of the future,* and that *the mind always contributes by much the largest share of the entertainment.* (295–96; emphases added)

Using modern terminology, Smith's argument here can be interpreted as indicating that temporal discounting is not only done but *mostly* done *within* the utility function, rather than "outside" the utility function as is usually considered in most formal models by assuming time-separable preferences. In particular, the pleasures and pains of the mind—that is, the effects of past and future events upon current utility—are the chief determinants of our current happiness and misery. In formal terms, the argument is that utility is not intertemporally separable but rather that past and future experiences, jointly with current ones, provide current utility. Unfortunately, perhaps because of analytical convenience, formal models in economics and other social sciences typically assume that preferences are time-separable, that is, that only current experiences provide current utility. These models can only capture the experiences of the body, not those of the mind. As a result, as Becker (1992, 327) remarks,

"the assumption of time separability has discouraged economists from grappling with issues of considerable importance—including addiction, work habits, preference formation, why children support their elderly parents, preference solutions to the problem of future commitments, and the evolution and stability of institutions."

But how can these time nonseparabilities be the mechanism through which passions and impulses can be mitigated or overcome? How do they operate? Having analyzed the fundamental features of human preferences, Smith is then ready to analyze how the mind, in the case of conflict between reason and imminent rewards ("lower" passions and impulses), can maintain its superiority:

> Since our happiness and misery, therefore, depended chiefly on the mind, if this part of our nature was well disposed, if our thoughts and opinions were as they should be, it was of little importance in what manner our body was affected. Though under great bodily pain we may still enjoy a considerable share of happiness, *if our reason and judgment maintained their superiority. We may entertain ourselves with the remembrance of the past, and with the hopes of future pleasure*. . . . That careful and laborious and circumspect state of mind, ever watchful and ever attentive to the most *distant consequences* of every action, could not be a thing pleasant or agreeable for its own sake, but upon account of its tendency to procure the greatest goods and to keep off the greatest evils. *To abstain from pleasure too, to curb and restrain our natural passions for enjoyment,* which was the office of temperance, *could never be desirable for its own sake.* The whole value of this virtue arose from its utility*, from its enabling us to postpone the present enjoyment for the sake of a greater one to come, or to avoid a greater pain that might ensue from it.* (296–97; emphases added)

Thus, the main insight from his analysis is based upon one fundamental postulate: both past and future experiences have an impact on immediate well-being and, hence, these two intertemporal nonseparabilities in human preferences may act as commitment devices, internal to the individual, since they partly control and commit current behavior. The possible impacts of past experiences (e.g., memory, habits, durability, custom, culture) and future ones on current utility lead to a potential triple effect of experiences: first through imagination, then through direct experience, and finally through memory, custom, and other forms

of durability in preferences.[10] Obviously, the manners in which memory and imagination influence current utility are qualitatively different. If we take the view of recent literature and consider an individual at different times as different selves, then when memory influences behavior one changes current behavior to alter the memories of future selves. The behavior is in the present and its effect on utility is in the future. On the other hand, when imagination influences behavior one takes actions in the present that will affect the future in order to alter experiences imagined by the present self. The behavior is in the future, and its effect on utility is in the present. Thus, the positioning of actor and beneficiary are reversed for memory and imagination.

Smith, therefore, offers two possible solutions to curb and restrain our natural lower passions from enjoyment: current effects of past and future experiences. However, he does not explicitly state whether he believes that these solutions would simply *mitigate* or could potentially be able to *fully* restrain all lower passions and impulses. This aspect, however, may be formally evaluated next.

First, consider the current effects of past experiences alone. Unlike the future, the past is not alterable, and even though we have some capacity to represent it as we wish and to direct our thoughts toward or away from it, its effect on the present is for the most part determined by prior choices. At the same time, our current self will become the past for future selves. Therefore, knowing that our present experiences will affect our future (as memories, habits, and the like), we may take actions in the present to change our future stock of past experiences (e.g., "Let's do it; we will never forget it"). If past experiences have an effect on current utility, utility at time t can be expressed without loss of generality as a recursive function of current experiences c_t and past utility:

$$u_{(t)} = u_{(t)}\left(c_t, u_{(t-1)}\right). \tag{3}$$

Despite the fact that Smith builds on, and corrects, David Hume's account of the way in which sentiments evolve (Haakonssen 1981), it is Hume ([1739] 1978, 422) who initially appears to offer an answer to the question of whether past and future experiences can preclude deviations from optimal choices. He concurs with his friend Smith on the general role of past experiences such as memories, habits, and custom.

---

10. Bentham (1789) and Jevons (1871) later used this concept of utility to understand intertemporal trade-offs.

And, in addition, he considers that past experiences are the main mechanism through which passions are modified and converted into current feelings:

> Nothing has a greater effect both *to increase and diminish our passions*, to convert pleasure into pain, and pain into pleasure, than custom and repetition. Custom has two *original* effects upon the mind, in bestowing a *facility* in the performance of any action or the *conception* of any object; and afterwards a *tendency* or *inclination* towards it; and from this we may account for all its other effects, however extraordinary. (422; emphases in original)

The relevant question is whether this formulation of preferences helps one simply mitigate or fully overcome the effects of passions and impulses. Consider these preferences within the basic problem stated in section 1. It can be formally shown that this specification of preferences alone *mitigates* the time-consistency problem of a nonexponential discounter but *cannot* fully solve it (see the appendix). The intuitive reason why past experiences alone cannot completely preclude all future deviations from today's optimal choices is that the dynamic problem of an individual subject to lower passions is in the future, not in the past. By "valuing" the past more (than when current utility is a function of current consumption alone), fewer consumption units will be allocated to future periods (the consumption path is steeper). As a result, changes in optimal consumption will still occur at every date since future selves will still be tempted to increase consumption above that originally planned. Note, however, that the effects of past experiences on current utility will *mitigate* or *reduce in part* the size of the change in optimal choices (with respect to the case in which the past has no effects on current utility). This is because a relatively steeper consumption path implies lower consumption levels in the future, and thus less drastic changes in planned choices over time.

Second, let us consider the state of mind "ever watchful and attentive to the most distant consequences of every action": the anticipation of future experiences through imagination. If future experiences are imagined and provide some utility in the present, utility at time t can be represented without loss of generality as a recursive function:

$$u_{(t)} = u_{(t)}\left(c_t, u_{(t+1)}\right). \tag{4}$$

It can be formally proven that this specification of utility may provide a *complete* resolution to the time-consistency problem under some conditions. In particular, future consumption choices will conform to present desires—and hence passions and impulses will be completely restrained—if and only if future utility is *currently* experienced at a marginal rate $(1 - \delta)\beta$; that is, if and only if

$$\frac{\partial u_{(t)}}{\partial u_{(t+1)}} = (1 - \delta)\beta. \tag{5}$$

This important formal result means that even though future utilities are discounted "outside the utility function" by a time-inconsistent discounting structure, the *MRS*'s will not change over time and, as a result, choices will be intertemporally consistent. The intuition for why anticipation of the future leads to the resolution of the intertemporal consistency problem is simple: self t and self $t + 1$ are to a certain extent "enemies," as self $t + 1$ will experience impulses inducing him to consume more at $t + 1$ than what self t wants for him now. Current utility derived from imagination of the future just makes them somewhat more "friendly," enough so that self $t + 1$ will *not* consume more than what self t wants him to consume by taking consumption from all future periods. The marginality condition indicates the rate at which this friendship needs to be built into preferences, that is, how "extra friendly" toward his future selves an individual needs to be in order to have time-consistent preferences.[11]

This analysis means that exponential discounting with intertemporally separable preferences is *observationally equivalent* to hyperbolic discounting when future utility is *currently felt* or *experienced* at a marginal rate of $(1 - \delta)\beta$. This equivalence result means that all the virtues, properties, and tractability of intertemporal separable preferences with exponential discounting can be recovered if we currently derive utility from future utility at a certain marginal rate. Even if future utilities are "externally" (outside the utility function) discounted in a time-inconsistent fashion, the individual will then behave *as if* he were a time-consistent agent with intertemporally separable preferences. Of course, this

---

11. This condition is derived from solving for the subgame perfect equilibrium *within preferences* through backward induction (see the formal proof in the appendix). When discounting is strictly hyperbolic, rather than quasi-hyperbolic, the condition will involve deriving utility in the present not just from the adjacent period but from each and all future periods at a rate that declines with the distance of the period.

is an "observationally equivalent" result in terms of choices or observed behavior, not in terms of actual feelings or experiences. The difference is that part of the present utility experienced by an individual with the "external" hyperbolic discounting structure is derived from his future utility, that is, he experiences some sentiments now associated with the future through imagination: $u_{(t)} = u_{(t)}(c_t, u_{(t+1)})$. The purely exponential discounter, however, experiences no passions and thus needs to derive no current utility from future utility in order to be time-consistent. In consequence, the valuation of future utilities may effectively be time-consistent if it operates through two components: the hyperbolic component, which operates "outside" utility, and the current effects of future experiences, which operate "within" the utility function. As Smith's arguments imply, these two components may cause the future to be discounted in a time-consistent fashion, allowing reason and judgment to maintain their superiority over lower passions.

Lastly, in general, both past and future experiences may simultaneously have an impact on current well-being. In this case, utility can be formulated as recursive in both past and future utilities:

$$u_{(t)} = u_{(t)}\left(u_{(t-1)}, c_t, u_{(t+1)}\right). \tag{6}$$

Formally, the marginality condition that is necessary and sufficient in order not to yield to passions and impulses over time is similar to the condition obtained in the case in which future experiences alone have an effect on current utility (see the appendix). From previous analysis, it is clear that if there are no cognitive or other costs involved in imagining the future, imagination of the future will be a *necessary* and possibly *sufficient* condition in order not to experience changes in optimal choices when prompted by the solicitations of lower-order passions. Past experiences, however, would not need to play any role. Alternatively, if there are some costs involved, imagination will still be a necessary and possibly sufficient feature of human preferences, but past experiences will also play a role if either (1) it is less costly to imagine past experiences than future ones, or (2) if they are a necessary input for the development of the imagination of the future.

Note that neither Smith nor Hume appears to resolve the question of whether these time nonseparabilities simply mitigate the time-consistency problem induced by certain lower passions or whether they are capable of solving the problem completely. Adam Smith seems to treat

both of these time nonseparabilities largely in a similar way, although he appears to be inclined toward the state of mind "ever watchful and ever attentive to the most distant consequences of every action," that is, toward the imagination of the future. As we saw in the quote above, Hume initially seems to favor custom and repetition as the main mechanism through which we can diminish our passions. Later, however, he appears to support the imagination of the future over the imagination of the past. He explains how it is naturally easier to imagine the future than the past:

> When the object is past, the progression of the thought in passing to it from the present is contrary to nature. . . . On the other hand, when we turn our thought to a future object our fancy flows along the stream of time, and arrives at the object by an order, which seems more natural, passing always from one point of time to that which is immediately posterior to it. This easy progression of ideas favours the imagination, and makes it conceive its object in a stronger and fuller light, than when we are continually oppos'd in our passage, and are oblig'd to overcome the difficulties arising from the natural propensity of the fancy. . . . An equal distance in the past and in the future, has not the same effect on the imagination. (Hume [1739] 1978, 431–32)

This passage favors the imagination of future events over memory. It also means that if past experiences are going to play a role in diminishing our passions it will not be because it is cognitively cheaper to imagine the past than the future. Hume, as a result, can be interpreted as favoring custom and repetition first, and then the imagination of the future over the imagination of the past. This, in principle, would appear to be incorrect given the superior effects of the imagination of the future over *any* effects of past experiences (including custom and repetition) shown in our previous analysis. However, this is not necessarily the case. The reason is that other mechanisms through which past experiences have an effect upon current utility such as habits, custom, and repetition do not need to be imagined, since they are already internalized in preferences. Also, as mentioned above, another possibility is that past experiences are an input into the development of imagination, possibly through a learning-by-doing mechanism. Hume ([1739] 1978, 9) considers very early in his *Treatise of Human Nature* that this is the case:

> Neither the lively nor the faint ideas [of memory or imagination] can make their appearance in the mind unless their correspondent impressions have gone before to prepare the way for them.

This passage is thus consistent with the ideas that custom bestows a facility and a tendency or inclination toward imagination, and that past experiences are a necessary input for the development of the imagination of the future in a learning-by-doing fashion—mainly through custom, habits, and repetition, and possibly through memory as well. As a result, his arguments can be interpreted as supporting the idea that utility may be formally represented by a function $u_{(t)} = u_{(t)}(u_{(t-1)}, c_t, u_{(t+1)})$, as Smith's analysis also implies, where past experiences $u_{(t-1)}$ are an input into the development of imagination. Formally, this would simply imply that past experiences $u_{(t-1)}$ and future experiences $u_{(t+1)}$ are complements in this function.

Unfortunately, not much systematic empirical evidence is available in the literature about how *internal* commitment mechanisms are actually developed. Recent work by Becker and Mulligan (1997) provides some evidence on how individuals can invest in "imagination capital." The investments that they discuss in the context of the endogenous formation of time discounting are concerned with individuals' efforts to reduce the discount on future utilities, rather than to overcome passions. Most of them, however, can also be interpreted as investments in sentiments (that is, in deriving current utility from imagining future events), and thus as investments in the formation of the utility function itself.[12]

3. Discussion

We have seen how Adam Smith and David Hume develop a theory of sentiments where passions could be usefully restrained by a contrary impulse. In particular, they emphasize how *nonmaterial* competing rewards may act as commitment mechanisms that can effectively counter some natural impulses of man. Because uncertain environments are the rule rather than the exception, it is relevant to discuss an extension of the analysis to the case in which choices are made under uncertainty.[13]

12. Interestingly, many of the arguments that these authors make are similar to Chalmers's (1830, 1832) plan of moral and religious education to improve the poor by making them more prudent and longsighted.

13. I am grateful to Gary S. Becker and Kevin M. Murphy for having suggested this extension.

Interestingly enough, this natural extension will confirm Adam Smith's ([1776] 1976) related insight that individuals *systematically* tend to over-value good states of nature and undervalue bad ones. Consider the general problem of time-inconsistent discounting under uncertainty:

$$\max u_{(t)} + \delta \sum_{i=1}^{T} \beta^i E_t u_{(t+i)}, \tag{7}$$

where past utility and *expected* future utility provide current utility:

$$u_{(t)} = u_{(t)} \left[u_{(t-1)}, c_t, E_t u_{(t+1)} \right], \tag{8}$$

where $E_t u_{(t+i)} = \sum_{s_{t+i}=1}^{S_{t+i}} \pi_{t+i}(s_{t+i}) \cdot u_{(t+i)}$ denotes the expected utility at time $t+i$ conditional on information at time t, and $\pi_{t+i}(s_{t+i})$ denotes the probability of state of nature s_{t+i} at time $t+i$, with $s_{t+i} \in [1, S_{t+i}]$. The maximization problem is subject to an intertemporal budget constraint. In addition, there is an updating rule in subjective probabilities $\pi(.)$ that may depend, for instance, on past realizations of utility levels or consumption choices in a Bayesian fashion.

When uncertainty about the future and the formation of subjective beliefs π are introduced into the analysis, two important new effects arise. First, subjective beliefs will now enter into the marginality condition under which a hyperbolic discounter will be able to solve, now in *expected* terms, his intertemporal problem of changes in optimal choices as time elapses. The reason is that it is *expected* future utility, rather than certain future utility, that is *felt* and thus discounted today. An individual with these preferences and this discounting structure will hope to be time-consistent at most *on average*. As a result, we may still observe changes in optimal choices, yet these changes are not intentional or irrational, as they are not expected to occur on average but may generally occur in individual cases.

Second, while it is true that when markets are complete (all possible states can be reached through markets) changes in optimal choices will not occur on average, when markets are incomplete they will occur, *even* on average. Incomplete markets can be interpreted as a situation in which some future possible states cannot be reached through trades in markets, and may include cases in which objective probabilities are not precisely defined. In these cases, individuals will find it optimal to bias their subjective probabilities of certain states of nature. The reason is that if $u_{(t)}$ is *concave* in expected future utility, then individuals will

optimally, consciously or unconsciously, decrease (increase) to some extent the subjective probabilities of those bad (good) states of nature (whose marginal utility is high (low)) where markets are not present, so that the expected payoffs in terms of utility are identical across states. The rational incentive in this case is to believe that bad states of nature are somewhat *less* likely to happen than objective probabilities, if available, dictate. Under incomplete markets, this effect is induced by the concavity of utility in expected future utility and the fact that individuals compute their expectation at time t with an endogenously chosen probability measure $\pi(.)$. This generates a process of thought or feeling that may be termed "rational wishful thinking."[14]

This analysis under uncertainty confirms Adam Smith's insight that individuals systematically tend to overvalue good states of nature and undervalue bad ones and thus tend to be "unrealistically" optimistic about future life events:

> How little the fear of misfortune is then capable of balancing the hope of good luck, appears still more evidently in the readiness of the common people to enlist as soldiers, or to go to sea, than in the eagerness of those of better fashion to enter into what are called the liberal professions. . . . though they [soldiers] have scarce any chance of preferment, they figure to themselves, in their youthful fancies, a thousand occasions of acquiring honor and distinction which never occur. (Smith [1776] 1976, 29–30, quoted in Becker and Mulligan 1997, 756)

The empirical evidence presented in Weinstein 1980, 1984 strongly supports his arguments. More importantly, note that these subjective valuations or self-serving biases are rational and crucially depend upon both the concavity of current utility in expected future utility and the hyperbolic discounting structure. This insight, which has not been incorporated into the modern economics literature, may have important implications for the analysis of risk-taking behavior within the context of current research on various behavioral anomalies.[15] In particular,

14. The same implications are derived if markets are complete but it is costly to transfer resources across possible states of nature or to gather relevant information about them. Likewise, a similar situation occurs when current consumption generates an uncertain negative externality toward future selves (e.g., as in the cases of the consumption of tobacco, alcohol, cholesterol, or fatty food).

15. Related "belief manipulation" has also been studied in interpersonal circumstances, although not in cases of time-inconsistent discounting. Rabin (1995) studies the implications of

phenomena such as weakness of will, shortsightedness, and disregard for future consequences are not necessarily signs of irrationality. As shown in this analysis, they are rational when current utility is concave in expected future utility and the discount factor is exponential on average (that is, it is not certain or stable since one of its components is uncertain and depends on the probabilities of future states, while the other component is hyperbolic).[16]

4. Concluding Remarks

This article points out a novel commitment mechanism that was suggested by Adam Smith and David Hume in their analyses of reason and passions. Their analyses show how sentiments may be endogenously formed within preferences to solve or mitigate intertemporal inconsistencies induced by lower-order passions.

Time-inconsistent discounting is considered to be a robust regularity in experimental psychology and behavioral economics. This feature, it is argued, is part of the hard-wiring of most nervous systems, and is "the process by which all people evaluate future goals" (Ainslie and Haslam 1992, 58). However, there can be little doubt that this statement, if left unqualified, may seem to conflict with much of the theoretical and empirical evidence in economics that supports models that use exponential discounting to explain much of the human behavior we observe in real situations. The argument in Adam Smith and David Hume is that passions, impulses, and imminent rewards are not always irresistible. They can be countered effectively not only when a deliberate action is taken to increase their cost or to exclude them from the set of potential choices, but also by competing rewards, even nonmaterial ones, provided they too are imminent. As Becker and Mulligan (1994, 11) discuss, "It may well be true [that people undervalue future utilities], but people train themselves to reduce and sometimes more than fully overcome any tendency to overvaluation . . . partly by spending time and other resources

different moral dispositions for the production of self-serving biases and the preferences for less information in moral reasonings when actions are expected to have bad consequences for other people (not other selves). The moral struggle is derived from the concern for other people in the evaluation of the extent of social harm of individual actions. See Prelec 1991 for a discussion of the analogies between intrapersonal and interpersonal choices in the moral realm.

16. See Becker and Mulligan 1994, 1997 for related discussions of this issue and Skog 1997 for a descriptive account of erratic but unbiased fluctuations in discount factors with intertemporally separable preferences.

to produce 'imagination' capital that helps them better appreciate the future." This imagination capital can take the form of *imminent* feelings or *sentiments*, that is, of investments in forming current utility itself. Smith and Hume study the natural formation of sentiments within the context of lower-order preferences. While certain passions induce time-inconsistent choices, the reason and judgment of what Smith refers to as the "impartial spectator" can curb and restrain them through sentiments associated with past and future experiences. Their analyses represent a robust theory of the endogenous formation of time-consistent discounting where sentiments, as imminent rewards, can be viewed as a mechanism to fine-tune our natural discounting mechanism, to make it more sensitive to distant rewards and penalties in selected instances.

Some economists and psychologists consider that we are initially endowed with hyperbolic discount functions and that consistent behavior is acquired as a skill. In this sense, Smith and Hume can be interpreted as offering a novel, elegant explanation of what it means to acquire, develop, and embed this skill in preferences, and the endogenous process through which it may be done. Their analyses, which so far have not been incorporated into current formal research in the literature, may be interpreted as a glimpse of the promise of interdisciplinary inquiry for current research in economics and all other social sciences. Formal research, in turn, can provide an evaluation of their arguments as well as an analytical interpretation of the determinants of the nature and formation of human preferences, whose crucial importance for economics, law, public policy, political economy, and other areas of inquiry was emphasized by classical economists.

Appendix

Let $\{c^*_{t+i} : i = 0, \ldots, T\}$ denote the optimal consumption sequence from the perspective of an individual at time t and $\{c^{(h)*}_{t+i} : i = h, \ldots, T\}$ denote the subsequence that starts at time $t + h$; $h \in (0, T]$; that is, for given $h : c^*_{t+m} = c^{(h)*}_{t+m}, \forall m \in [h, T]$. Formally, an optimal consumption sequence from the perspective of time t is a *time-consistent equilibrium sequence* if it is optimal from the perspective of all future $t + h$ selves, $0 < h \leq T$; that is, if the *MRS*'s between consumption at any two periods do not change over time and, hence, if all of their subsequences coincide with the optimal consumption sequences from the perspectives

of all future $t + h$ selves, then

$$\left\{ c_{t+i}^{(h)*} : i = h, \ldots, T \right\}$$
$$= \left\{ c_{t+h+j}^{*} : j = 0, \ldots, T - h \right\}, \quad \forall h \in (0, T]. \quad \text{(A1)}$$

Three cases are analyzed next.

A. HABITS, MEMORIES, AND OTHER EFFECTS OF PAST EVENTS. If past experiences have an effect on current utility, utility at time t can be expressed without loss of generality as $u_{(t)} = u_{(t)}(c_t, u_{(t-1)})$ with $u'_{(t)1} = \partial u_{(t)}/\partial c_t > 0$, $u''_{(t)1} = \partial^2 u_{(t)}/\partial c_t^2 < 0$, $u'_{(t)2} = \partial u_{(t)}/\partial u_{(t-1)} > 0$. Then, a hyperbolic discounter who at time t maximizes $u_{(t)} + \delta \sum_{i=1}^{T} \beta^i u_{(t+i)}$ will equate the *MRS*'s of the different consumption goods to $1 + r$ to obtain his optimal consumption sequence. From the perspective of time t, the MRS_{t+2}^{t+1} between c_{t+1} and c_{t+2} is

$$\frac{\delta\beta u'_{(t+1)1} + \delta \sum_{i=1}^{T-1} \left[\beta^{i+1} \left(\prod_{\tau=2}^{i+1} u'_{(t+\tau)2} \right) \cdot u'_{(t+1)1} \right]}{\delta\beta^2 u'_{(t+2)1} + \delta \sum_{i=2}^{T-1} \left[\beta^{i+1} \left(\prod_{\tau=3}^{i+1} u'_{(t+\tau)2} \right) \cdot u'_{(t+2)1} \right]} = 1+r, \quad \text{(A2)}$$

whereas from the perspective of time $t + 1$ it becomes

$$\frac{\beta u'_{(t+1)1} + \delta \sum_{i=1}^{T-1} \left[\beta^{i+1} \left(\prod_{\tau=2}^{i+1} u'_{(t+\tau)2} \right) \cdot u'_{(t+1)1} \right]}{\delta\beta^2 u'_{(t+2)1} + \delta \sum_{i=2}^{T-1} \left[\beta^{i+1} \left(\prod_{\tau=3}^{i+1} u'_{(t+\tau)2} \right) \cdot u'_{(t+2)1} \right]} = 1+r. \quad \text{(A3)}$$

Note that regardless of the functional form of the utility function, the MRS_{t+2}^{t+1} as seen from the perspectives of time t and time $t + 1$ can *never* coincide. The numerator of the MRS_{t+2}^{t+1} from time t's perspective is always smaller than the one from $t + 1$'s perspective by the amount $(1-\delta)\beta u'_{(t+1)1}$. Therefore, a person will never behave as an exponential discounter. Note, however, that the effects of past experiences on current utility *mitigate* or *reduce in part* the size of the change in optimal choices (with respect to the case in which the past has no effects on current utility). This is because the difference between the MRS_{t+2}^{t+1} as seen from t and from $t + 1$ becomes smaller (and therefore so does the difference between the optimal consumption choices) the larger the numerator and denominator become. In other words, the term $(1 - \delta)\beta u'_{(t+1)1}$, which drives the difference, becomes relatively less important the greater the memory or custom effects that current consumption will have in the future.

B. Anticipation of Future Experiences through Imagination.
Without loss of generality, utility at time t can be represented as $u_{(t)} = u_{(t)}(c_t, u_{(t+1)})$ with $u'_{(t)1} = \partial u_{(t)}/\partial c_t > 0$, $u''_{(t)1} = \partial^2 u_{(t)}/\partial c_t^2 < 0$, $u'_{(t)2} = \partial u_{(t)}/\partial u_{(t+1)} > 0$. From the perspective of time t, the MRS_{t+2}^{t+1} between c_{t+1} and c_{t+2} is

$$\frac{u'_{(t)2} \cdot u'_{(t+1)1} + \delta\beta u'_{(t+1)1}}{u'_{(t)2} \cdot u'_{(t+1)2} \cdot u'_{(t+2)1} + \delta\beta u'_{(t+1)2} \cdot u'_{(t+2)1} + \delta\beta^2 u'_{(t+2)1}} = 1+r, \quad \text{(A4)}$$

whereas when the individual gets to time $t + 1$ it becomes

$$\frac{u'_{(t+1)1}}{u'_{(t+1)2} \cdot u'_{(t+2)1} + \delta\beta u'_{(t+2)1}} = 1 + r. \quad \text{(A5)}$$

Simple algebra shows that these two MRS's are exactly identical when $u'_{(t)2} = \partial u_{(t)}/\partial u_{(t+1)} = (1 - \delta)\beta$. Similar reasonings apply to all future periods. This effect generates a greater valuation or appreciation of the future, in this case *within* the utility itself, which overcomes the future relative "overvaluation" of the corresponding present so as to make choices intertemporally consistent.

C. Memory and Imagination. If both past and future experiences have an impact on today's well-being, utility can be defined as

$$u_{(t)} = u_{(t)}\left(u_{(t-1)}, c_t, u_{(t+1)}\right). \quad \text{(A6)}$$

Then, following reasoning similar to that employed above, a hyperbolic discounter will experience no changes in optimal choices if and only if his MRS's do not change over time. This in turn implies that preferences are intertemporally consistent if and only if

$$u'_{(t)c_{t+1}} = (1 - \delta)\beta u'_{(t+1)c_{t+1}} \quad \text{(A7)}$$
$$u'_{(t)c_{t+2}} = (1 - \delta)\beta u'_{(t+1)c_{t+2}}$$

for all t.

All the basic results obtained in these three cases are maintained if we consider instead more general recursive preferences such as $u_{(t)} = u_{(t)}(c_t, u_{(t-1)}, u_{(t-2)}, \dots)$ in the case of past experiences, and $u_{(t)} = u_{(t)}(c_t, u_{(t+1)}, u_{(t+2)}, \dots)$ in the case of future experiences. These specifications are necessary when the discounting structure is purely hyperbolic, $1/(k_1 + k_2 \cdot t)$, as opposed to the quasi-hyperbolic specification typically used in the literature. See Koopmans 1960 and Koopmans, Diamond, and Williamson 1964 for different postulates concerning the

underlying preference orderings when preferences are recursive in future utility, and Blackorby, Nissen, Primont, and Russell 1973 for an early paper—in the context of social choice with succeeding generations planning for the future—on how recursivity induces intertemporal consistency.

References

Ainslie, G. 1975. Specious Reward: A Behavioral Theory of Impulsiveness and Impulse Control. *Psychological Bulletin* 82:463–96.

——— . 1986. Beyond Microeconomics: Conflict among Interests in a Multiple Self as a Determinant of Value. In *The Multiple Self*, edited by Jon Elster 133–75. Cambridge: Cambridge University Press.

——— . 1992. *Picoeconomics.* Cambridge: Cambridge University Press.

Ainslie, G., and N. Haslam. 1992. Hyperbolic Discounting. In *Choice over Time*, edited by G. Loewenstein and J. Elster, 57–92. New York: Russell Sage.

Becker, G. S. 1992. Habits, Addictions, and Traditions. *Kyklos* 45:327–45.

——— . 1995. *The Essence of Becker.* Stanford, Calif.: Hoover Institution Press.

——— . 1996. *Accounting for Tastes.* Cambridge: Harvard University Press.

Becker, G. S., and C. B. Mulligan. 1994. On the Endogenous Determination of Time Preference. Working paper series no. 98. Chicago: Center for the Study of the Economy and the State, University of Chicago.

——— . 1997. The Endogenous Determination of Time Preference. *Quarterly Journal of Economics* 112:729–58.

Benabou, R., and J. Tirole. 2000. Willpower and Personal Rules. Mimeo, Princeton University.

Bentham, J. 1789. *Principles of Morals and Legislation*. Oxford: Clarendon.

Blackorby, C., D. Nissen, D. Primont, and R. R. Russell. 1973. Consistent Intertemporal Decision Making. *Review of Economic Studies* 40:239–48.

Böhm-Bawerk, E. V. [1891] 1971. *The Positive Theory of Capital.* Freeport, N.Y.: Books for Libraries Press.

Brocas, I., and J. Carrillo. 2000. The Value of Information When Preferences Are Dynamically Inconsistent. *European Economic Review* 44:1104–15.

Caillaud, B., and B. Jullien. 2000. Modelling Time-Inconsistent Preferences. *European Economic Review* 44:1116–24.

Campbell, J. Y., and J. H. Cochrane. 1999. By Force of Habit: A Consumption Based Explanation of Aggregate Stock Market Returns. *Journal of Political Economy* 107.2:205–51.

Carrillo, J., and T. Mariotti. 2000. Strategic Ignorance as a Self-Disciplining Device. *Review of Economic Studies* 67.3:529–44.

Chalmers, T. 1830. *The Works of Thomas Chalmers*. Philadelphia: J. Towar & D. M. Hogan.

———. 1832. *On Political Economy, in Connection with the Moral State and Moral Prospects of Society*. New York: D. Appleton.

De Villiers, P. A., and R. J. Herrnstein. 1976. Toward a Law of Response Strength. *Psychological Bulletin* 83:1131–53.

Elster, J. 1984. *Ulysses and the Sirens: Studies in Rationality and Irrationality*. Cambridge: Cambridge University Press.

Elster, J., and G. Loewenstein. 1992. Utility from Memory and Anticipation. In *Choice over Time*, edited by George Loewenstein and Jon Elster, 213–34. New York: Russell Sage.

Epstein, L. G., and J. A. Hynes. 1983. The Rate of Time Preference and Dynamic Economic Analysis. *Journal of Political Economy* 91:611–35.

Frank, R. H. 1988. *Passions within Reason: The Strategic Role of the Emotions*. New York: Norton.

Goldman, S. M. 1980. Consistent Plans. *Review of Economic Studies* 47:533–37.

Gul, F., and W. Pesendorfer. 2001. Temptation and Self-Control. *Econometrica* 69 (November): 1403–35.

Haakonssen, K. 1981. *The Science of a Legislator*. Cambridge: Cambridge University Press.

Herrnstein, R. J. 1961. Relative and Absolute Strengths of Responses as a Function of Frequency of Reinforcement. *Journal of Experimental Analysis of Animal Behavior* 4:267–72.

———. 1997. *The Matching Law: Papers in Psychology and Economics*. Edited by David I. Laibson. Cambridge: Harvard University Press.

Hirschman, A. O. 1977. *The Passions and the Interests*. Princeton, N.J.: Princeton University Press.

Hori, H. 1992. Utility Functionals with Nonpaternalistic Intergenerational Altruism: The Case Where Altruism Extends to Many Generations. *Journal of Economic Theory* 56:451–67.

Hume, D. [1739] 1978. *A Treatise of Human Nature*. First edition. New York: Oxford University Press.

Jevons, W. S. 1871. *Theory of Political Economy*. London: Macmillan.

Kimball, M. S. 1987. Making Sense of Two-Sided Altruism. *Journal of Monetary Economics* 20:301–26.

Koopmans, T. C. 1960. Stationary Ordinal Utility and Impatience. *Econometrica* 28:287–309.

Koopmans, T. C., P. A. Diamond, and R. E. Williamson. 1964. Stationary Utility and Time Perspective. *Econometrica* 32:82–100.

Laibson, D. I. 1997. Golden Eggs and Hyperbolic Discounting. *Quarterly Journal of Economics* 112:443–77.

Locke, J. 2000. *John Locke as Translator: Three of the Essais of Pierre Nicole in French and English*. Edited by J. S. Yolton. Oxford: Voltaire Foundation.

Loewenstein, G. 1987. Anticipation and the Valuation of Delayed Consumption. *Economic Journal* 97:666–84.

———. 1992. The Fall and Rise of Psychological Explanations in the Economics of Intertemporal Choice. In *Choice over Time*, edited by George Loewenstein and Jon Elster. New York: Russell Sage.

Loewenstein, G., and D. Prelec. 1992. Anomalies in Intertemporal Choice: Evidence and an Interpretation. *Quarterly Journal of Economics* 57:573–98.

Malthus, T. R. [1826] 1986. *An Essay on the Principle of Population*. In *The Works of Thomas Robert Malthus*, vols. 2 and 3. Edited by E. A. Wrigley and David Souden. London: W. Pickering.

Marshall, A. [1890] 1920. *Principles of Economics.* 8th ed. London: Macmillan.

Meghir, C., and G. Weber. 1996. Intertemporal Nonseparability or Borrowing Restrictions? A Disaggregate Analysis Using a U.S. Consumption Panel. *Econometrica* 64:1151–81.

Pareto, V. 1909. *Manuel d'économie politique.* Paris: Girard and Brieve.

Peleg, B., and M. E. Yaari. 1973. On the Existence of a Consistent Course of Action When Tastes Are Changing. *Review of Economic Studies* 40:391–401.

Pollak, R. A. 1968. Consistent Planning. *Review of Economic Studies* 35:201–8.

Posner, R. A. 1997. Are We One Self or Multiple Selves? Implications for Law and Public Policy. *Legal Theory* 3:23–35.

Prelec, D. 1991. Values and Principles: Some Limitations of Traditional Economic Analysis. In *Socioeconomics: Toward a New Synthesis*, edited by A. Etzioni and P. Lawrence, 131–45. New York: M. E. Sharpe.

Rabin, M. 1995. Moral Preferences, Moral Constraints, and Self-Serving Biases. Working paper no. 95-241, University of California, Berkeley.

———. 1999. Psychology and Economics. *Journal of Economic Literature* 36:11–46.

Schelling, T. C. 1984. Self-Command in Practice, in Policy, and in a Theory of Rational Choice. *American Economic Review* 74:1–11.

Skog, O.-J. 1997. The Strength of Weak Will. *Rationality and Society* 9:245–71.

Smith, A. [1759] 1976. *The Theory of Moral Sentiments.* First edition. London: Oxford University Press.

———. [1776] 1981. *An Inquiry into the Nature and Causes of the Wealth of Nations.* First edition. Indianapolis: Liberty Classics.

Strotz, R. H. 1956. Myopia and Inconsistency in Dynamic Utility Maximization. *Review of Economic Studies* 23:165–80.

Thaler, R. H., and H. M. Shefrin. 1981. An Economic Theory of Self-Control. *Journal of Political Economy* 89:392–406.

Weinstein, N. D. 1980. Unrealistic Optimism about Future Life Events. *Journal of Personality and Social Psychology* 39:806–20.

———. 1984. Why It Won't Happen to Me: Perceptions of Risk Factors and Susceptibility. *Health Psychology* 3:431–57.