On Intertemporal Choice, Rationality, and Time Consistency*

Tobias J. Klein†

May 2003

Abstract

This paper is inspired by the literature on procrastination in behavioral economics. I develop a general framework in order to define time consistent choice formally. The distinction between static and dynamic rationality is made and conditions are derived under which choice is time consistent.

Journal of Economic Literature Classification Number: D90.

Key Words: Time consistency, preferences, static and dynamic rationality.

*I would like to thank Martin Hellwig for an insightful discussion in which I got the motivation to write this paper.

†University of Mannheim, Department of Economics, 68131 Mannheim, Germany. *E-Mail:* klein@econ.uni-mannheim.de.

> "Consumers face two challenges: making good decisions and sticking to them. Economists have adopted optimistic assumptions on both counts. The consumers in mainstream economic models are assumed both to be exceptionally good decisionmakers and to be able to carry out their plans." – *Laibson, Repetto, and Tobacman (1998, p. 92).*

1 Introduction

Recently, there have been many contributions on the field of behavioral economics dealing with intertemporal choice. Catchwords are, for example, "procrastination", "present bias" as well as "sophistication and naivite". Rabin (1998) contains an excellent overview. This paper is an attempt to establish the link between this fascinating field and the standard approach to decision making in microeconomic theory where preference based approaches and approaches that build on choice rules play an important role. In behavioral economics, one's starting point in modeling is usually a utility function and it is kept secret what the properties of the preference relation that is represented by this utility function actually are. This includes the question what implicit assumptions on rationality we make once we start papers with the assertion that there exists such a utility function. We would like to fill this gap by first developing a formal framework and then adressing those questions directly.

2 Formal Framework

The formal framework I propose in this section is a slight generalization of the approach to decision making in standard microeconomic theory, in particular the first chapters of Mas-Colell, Whinston, and Green (1995). This proceeding has two advantages. On the one hand, this way to think about choice has proven to be very constructive because of its flexibility and on the other hand, it allows me to express the ideas of this paper in a fashion to which a lot of economists are used to.

Since our aim is to model intertemporal choice, we start with a set X of alternatives and discrete time $t = 1, \dots, T$. We will think of X as being given by

$$X = \mathbb{R}^{LT}_{+} = \left\{ x \in \mathbb{R}^{LT} : x_{lt} \geq 0 \text{ for } l = 1, \dots, L \text{ and } t = 1, \dots, T \right\}$$

so that, for example, consumption of some good l occurs in period 1 through T. In every period, the decision maker possesses a preference relation $\succeq_t$ on the set of alternatives. Throughout the paper, we want to make the following assumption which guarantees that in every t, the preference relation can be represented by a utility function.

Assumption 1: *The preference relation $\succeq_t$ possesses the following three properties:*

1. Completeness: *for all $x, y \in X$, we have that $x \succeq_t y$ or $y \succeq_t x$ (or both),*
2. Transitivity: *for all $x, y, z \in X$, if $x \succeq_t y$ and $y \succeq_t z$, then $x \succeq_t z$,*
3. Continuity: *for any sequence of pairs $\{x^n, y^n\}_{n=1}^{\infty}$ with $x^n \succeq_t y^n$ for all n, $x = \lim_{n\to\infty} x^n$, and $y = \lim_{n\to\infty} y^n$, we have $x \succeq_t y$.*

We imagine the decision process to be a combination of some irreversible action which occurs in this period, t, and some planned action in all future periods $t+1, \ldots, T$. An example for such an action is the process of buying some good. In other words, in every period t, there is a dichotomy between now and the future. Therefore, we can decompose X into a set N_t of alternatives among which is chosen now, a set F_t of alternatives among which will be chosen in the future, and a set P_t of alternatives among which has been chosen in the past. That is, in every t, we have $X = P_t \times N_t \times F_t$. Let P_t, N_t, and F_t have typical elements p_t, n_t, and f_t, respectively. Then, in period t, p_t is known to the decision maker, she fixes n_t according to her preference relation $\succeq_t$ by undertaking such an irreversible action, and she makes plans about choice actions in the future, f_t. To illustrate this, consider the following example.

Example 1 (One Good, Three Periods): Let x be a consumption path over time and $t = 1, 2, 3$ as well as $L = 1$. This is, consumption of one single good occurs in $t = 1, 2, 3$. Now, it might be that in $t = 1$, even though the decision maker has a preference relation on X, she can only choose period 1's consumption level, i.e. $P_1 = \emptyset$, $N_1 = \mathbb{R}_+$, and $F_1 = \mathbb{R}^2_+$. As a by-product, she makes plans f_1 for the future. In $t = 2$, she chooses n_2 and plans to choose f_2 in the future, and so on. Another situation arises when the decision maker chooses a complete consumption path x already in, $t = 1$ and uses a commitment device in order to commit herself to this choice so that she *has* to stick to it. Then, $N_1 = \mathbb{R}^3_+$ and $N_2 = N_3 = \emptyset$. Since her choice is irreversible, we have that $F_1 = F_2 = F_3 = \emptyset$.

To complete the description of the decision problem, we have to introduce a set of constraints. These constraints establish the link between what is chosen today, n_t, is planned to be choose in the future, f_t, and what has been chosen in the past, p_t. Recall that "choice" is used in the sense that n_t is implemented immediately, and some plan f_t for the future is made. Note that what the decision maker can commit to is part of n_t and not f_t. To avoid confusion, we assume throughout the paper that there is no such commitment device. Then, we have that, for $t = 1, \ldots, T$, $P_t = \mathbb{R}^{L(t-1)}_+$, $N_t = \mathbb{R}^L_+$, and $F_t = \mathbb{R}^{L(T-t)}_+$. For an example of such as set of constraints, let the interest rate between period 1 and t be denoted by r_t and let the price vector in t be denoted by $q_t \in \mathbb{R}^L$. Furthermore, let ω be the vector of endowments and let x_t be consumption which occurs in t. Then, the budget set is given by

$$B = \left\{ x \in \mathbb{R}^{LT}_+ : \sum_{t=1}^{T} \frac{1}{1+r_t} \cdot q_t \cdot x_t \leq \sum_{t=1}^{T} \frac{1}{1+r_t} \cdot q_t \cdot \omega \right\}.$$

In the remainder, we would like to be more general and will just say that there a family $\mathscr{B}$ of nonempty subsets of X. Then, for a set of intertemportal constraints, possibly empty, we have a subset $B \in \mathscr{B}$ of X which satisfies these constraints. That is, every element $B \in \mathscr{B}$ is a nonempty subset of X. This completes the description of the formal framework.

3 Time Consistency and Rationality

If the decision maker is preference maximizing in t, he chooses any of the elements in the set

$$C^*_t = \{x \in B : x \succeq_t y \text{ for every } y \in B\}.$$

So far, we did not make the underlying assumptions on rationality explicit. In fact, the assumptions of completeness and transitivity guarantee that, technically, $\succeq_t$ is what is called a weak order so that given some B, there exists a nonempty set C^*_t of best alternatives. This is where we assert that the decision maker is able to identify these most preferred alternatives, knows his preferences, and does not make mistakes. Without a fair amount of rationality, this would not be possible.

The ingrediants of our model, call it $\mathcal{M}$, are by now, for every t, a set $X = P_t \times N_t \times F_t$ of alternatives, a preference relation $\succeq_t$ on X, and a family $\mathscr{B}$ of nonempty subsets of X. Let us now define time consistency formally.

Definition 1 (Time Consistency): *We say that choice is* time consistent *if, for all $t, \tilde{t} = 1, \ldots, T$ and all $B \in \mathscr{B}$, we have that $x \in C_t^*$ implies $x \in C_{\tilde{t}}^*$*

Our first result follows almost immediately from the definition of time consistency.

Lemma 1: *Assume $\mathcal{M}$. If choice is time consistent, we have that $x \succeq_t y$ for some t implies $x \succeq_{\tilde{t}} y$ for all $\tilde{t} = 1, \ldots, T$.*

This is, as long as choice is time consistent, the rationality assumptions we made in Assumption 1 yield clear cut results. But once we ask ourselves under which conditions time consistency is achieved, we have to think about rationality a little more profoundly. This is because so far, we have only said something about rationality within each period t, namely that there exists, for every period, a weak order that is defined on the set of alternatives. We will refer to this as *static rationality.* Technically speaking, the preorder being the same in each period implies time consistency. We will see, however, that the converse need not hold. A concept which we will call *dynamic rationality* will be important in this context. Before proceeding with the discussion we show by counterexample that time consistency is not implied by Assumption 1.

Example 2 (Counterexample): Pick two periods s and t such that $0 < s < t \leq T$ and let $X = \mathbb{R}_+^2$. Assumption 1 implies that the preference relation in s and t can be represented by a utility function, respectively. Choose

$$U_s(x_s, x_t) = u(x_s) + \beta u(x_t)$$
$$U_t(x_s, x_t) = \beta u(x_s) + u(x_t)$$

where $\beta \notin \{0, 1\}$ and where the felicity function $u(\cdot)$ exhibits a positive first derivative, $u'(\cdot)$, and a negative second derivative, $u''(\cdot)$. For $\beta \in (0, 1)$, we get a untility function representing a preference relation that puts special weight on this period's pleasure. Suppose there is only an inner solution to the maximization problem. Then, if we use a standard utility maximization approach, we have that

$$\left.\frac{\mathrm{d}x_t}{\mathrm{d}x_s}\right|_{\text{in } s} = \frac{\frac{\partial U_s(x_s,x_t)}{\partial x_s}}{\frac{\partial U_s(x_s,x_t)}{\partial x_t}} = \frac{u'(x_s)}{\beta u'(x_t)} \neq \frac{\beta u'(x_s)}{u'(x_t)} = \frac{\frac{\partial U_t(x_s,x_t)}{\partial x_s}}{\frac{\partial U_t(x_s,x_t)}{\partial x_t}} = \left.\frac{\mathrm{d}x_t}{\mathrm{d}x_s}\right|_{\text{in } t}.$$

That is, the intertemporal rate of substitution between s and t depends on the time of the evaluation so that if the decision maker could, she would choose differently in s and t, i.e. $x \in C_s^*$ implies $x \notin C_t^*$. Consequently, time consisitency does not hold.

In the previous example, time consistency is still achievable if we require the decision maker to use all the information that is available to her. This additional information is given by her knowledge of the utility functions in *all future* periods that she can use to predict her future behavior. Intuitively, this imposes a set of additional constraints on her plans for the future. In terms of game theory, we could think of the decision maker in period s as being a different person as in period t, both playing a game with each other and taking the behavior of the other player into account. Then, if the equilibrium concept is Nash, it follows by construction that it must be subgame perfect and choice is time consistent. We can summarize those informal statements by first making an assumption and then showing that this assumption implies time consistency.

ASSUMPTION 2: *The decision maker is* dynamically rational. *That is, in $t < T$, he takes into account how he will behave in any period $\tilde{t} > t$.*

This is referred to as sophistication by O'Donoghue and Rabin (forthcoming) in contrast to naivite which is defined as the complete absence of sophistication. Assumption 2 allows us to state Lemma 2.

LEMMA 2: *Assume $\mathcal{M}$ and let Assumption 2 hold. Then, choice is time consistent.*

Proof. By backward induction. Given B, p_T and $\succeq_T$, the set C_T^* is known; therefore, given B, p_{T-1} and $\succeq_{T-1}$, the set $\{f_{T-1} : f_{T-1} \in C_T^*\}$ is known, and so on. That is, in t, given B and p_t, f_t depends on $\succeq_{t+1}, \ldots, \succeq_T$ and n_t and in a slight abuse of notation, we can write $f_t = f_t(n_t)$ so that the set of preference maximizing alternatives is given by

$$C_t^* = \left\{ n_t \in B : \begin{pmatrix} p_t \\ n_t \\ f_t(n_t) \end{pmatrix} \succeq_t y \text{ for every } y \in B \right\}$$

and the assertion follows. □

By Assumption 1, for every t, there exists a utility function $U_t : X \to \mathbb{R}$ representing $\succeq_t$. As an alternative to Assumption 2, we could make an assumptions on the utility function itself. We will refer to it as the "linearity assumption" because it allows us to express utility as a linear combination of utility contributions $u_t(x_t)$ that are independent of each other.

Assumption 3 (Linearity): *$U_t(\cdot)$ is so that, for every $0 \leq s < t \leq T$ and every $\Delta_1, \Delta_2 \in \mathbb{N} \cup \{0\}$, $\Delta_1, \Delta_2 \leq T - t$,*

$$\frac{\frac{\partial U_s}{\partial x_{t+\Delta_1}}}{\frac{\partial U_s}{\partial x_{t+\Delta_2}}} = \frac{\frac{\partial U_t}{\partial x_{t+\Delta_1}}}{\frac{\partial U_t}{\partial x_{t+\Delta_2}}}.$$

Note that Assumption 3 does not hold in Example 2. Furthermore, note that Assumption 3 holds for utility function that exhibit exponential discounting, i.e. utility functions that are of the form

$$U_t = \sum_{\tau=1}^{T} \delta^\tau u_\tau(x_\tau), \ \delta > 0,$$

or any positive monotone transformation of them. Finally, we should mention that Pollak (1968) has shown that Assumption 3 also holds for utility functions that are of the form

$$U_t = \sum_{\tau=1}^{T} \delta(\tau) \ln(x_\tau)$$

or any positive monotone transformation of them. One particular monotone transformation yields the well known Cobb-Douglas utility function.

Let us summarize the main results of this paper in a theorem.

Theorem 1: *Assume $\mathcal{M}$ and let Assumption 1 hold. Then, choice is time consisitent if Assumption 2 or Assumption 3 holds.*

Proof. The first half follows from Lemma 2. For the second half, note that B does not change over time so that together with Assumption 3, the assertion follows. □

Otherwise put, once we know that Assumption 1 and either Assumption 2 or Assumption 3 hold, we only have to know $\succeq_t$ for some t in order to predict the decision maker's choice behavior.

4 Concluding Remarks

We have seen that even under certainty, only very strong assumptions, either asserting that decision makers are marvelously rational or that their utility function has a particular functional form that is beyond belief, can guarantee that choice is time consisitent. The lack of grip on reality of these assumptions was the motivation for a variety of models to be constructed in what has become the field of behavioral economics. But still, one should realize that once a utility function shows up in these models, even if it depends on time, one implicitly makes the assumption that the decision maker is completely rational within each period. I have not found an answer to the question why, then, he should not be so over time and choose in a time consistent way. Maybe, we shall attempt to weaken the assumption of rationality within each period as a consequence.

References

Laibson, D. I., A. Repetto, and J. Tobacman (1998): "Self-Control and Saving for Retirement," *Brookings Papers on Economic Activity*, 1, 91–196.

Mas-Colell, A., M. D. Whinston, and J. R. Green (1995): *Microeconomic Theory.* Oxford University Press, New York, Oxford.

O'Donoghue, T., and M. Rabin (forthcoming): "Self Awareness and Self Control," in *Now of Later: Economic and Psychological Perspectives on Intertemporal Choice*, ed. by R. Baumeister, G. Loewenstein, and D. Read. Russell Sage Foundation Press, New York.

Pollak, R. A. (1968): "Consistent Planning," *Review of Economic Studies*, 35(2), 201–208.

Rabin, M. (1998): "Psychology and Economics," *Journal of Economic Literature*, 36, 11–46.