

# The Psychology of Intertemporal Tradeoffs

Marc Scholten  
ISPA University Institute

Daniel Read  
Durham Business School and Yale School of Management

It is commonly assumed that people make intertemporal choices by “discounting” the value of delayed outcomes, assigning discounted values independently to all options, and comparing the discounted values. We identify a class of anomalies to this assumption of alternative-based discounting, which collectively shows that options are not treated independently but rather comparatively: The time difference, or interval, between the options sometimes counts more and sometimes counts less if it is taken as a whole than if it is divided into shorter subintervals (*superadditivity* and *subadditivity*, respectively), and whether the interval counts more or less depends on the money difference, or compensation, involved (*inseparability*). We develop a model that replaces alternative-based discounting with attribute-based tradeoffs. In our model, people make intertemporal choices by weighing how much more they will receive or pay if they wait longer against how much longer the wait will be, or, conversely, how much less they will receive or pay if they do not wait longer against how much shorter the wait will be. This model, called the *tradeoff model*, accommodates, in a psychologically plausible way, all anomalies that the discounting approach can and cannot address.

*Keywords:* intertemporal choice, delay discounting, alternative-based choice, tradeoffs, attribute-based choice

Intertemporal choices involve tradeoffs between costs and benefits that occur at different points in time (Loewenstein & Elster, 1992; Loewenstein, Read, & Baumeister, 2003). These include choices such as taking a job now or getting an education and having a chance at a better job later, and spending money now or saving it and having more to spend later. Most intensively investigated, however, are much more elementary choices between smaller-sooner and larger-later amounts of money, such as receiving \$100 now or \$150 in 3 months. It is commonly assumed that people make intertemporal choices by “discounting” the value of delayed outcomes, assigning discounted values to the options, and

then comparing these discounted values. For this choice, they would compare the value of \$100 now with the discounted value of \$150 in 3 months. We argue that this does not accurately describe the *psychology* of intertemporal tradeoffs.

Discounting models belong to the broad class of *alternative-based* choice models, in which the options are independently assigned an overall value, these values are compared, and the option with the highest value is chosen. Alternative-based choice models can be contrasted with *attribute-based* ones (Payne, Bettman, & Johnson, 1988), in which the options are directly compared along their attributes, and the option favored by these comparisons is chosen.<sup>1</sup> Alternative-based discounting models can accommodate much of what we know about intertemporal choice, but there remains evidence, reviewed and strengthened in this article, that can only be addressed by an attribute-based choice model. In this article, we develop such a model, called the *tradeoff model*. Our model accommodates, in a psychologically plausible way, all anomalies that discounting models can and cannot address.

We begin with a review of discounting models and the evidence they can explain. We then identify three anomalies to discounting theory. Given these anomalies, we develop the tradeoff model and show how it accommodates *all* the evidence, whether consistent or inconsistent with discounting theory. After this largely qualitative development, we provide a fully parametric specification of the tradeoff model. We conclude by discussing how the tradeoff model deals with additional evidence and how it improves on related theories of choice.

---

Marc Scholten, Department of Social and Organizational Psychology, ISPA University Institute, Lisboa, Portugal; Daniel Read, Durham Business School, Durham University, Durham, England, and Yale School of Management, Yale University.

We acknowledge financial support from (1) the Fundação para a Ciência e Tecnologia (FCT), program POCI 2010 and projects POCTI/PSI/42275/2001, POCI/PSI/56030/2004, PPCDT/PSI/56030/2004, and PTDC/PSI-PCO/101447/2008; and (2) the Economic and Social Research Council (ESRC), grant RES-000-22-0201. We are eternally indebted to a number of people for the amazing generosity with which they helped us solve mathematical and technical problems. In order of appearance: António St. Aubyn (Universidade Lusíada de Lisboa, Lisboa, Portugal), Joaquim Judice (Universidade de Coimbra, Coimbra, Portugal), and João Patrício (Instituto Politécnico de Tomar, Tomar, Portugal), the contributions of whom will also be part of a separate article; Ana Rita Pires and Craig Desjardins (MIT, Department of Mathematics); and Peter Wakker (Erasmus University, Econometric Institute, Rotterdam, the Netherlands). Finally, we are immensely grateful to Andrew Meyer (Yale University) for his untiring assistance in data collection.

Correspondence concerning this article should be addressed to Marc Scholten, ISPA University Institute, Rua Jardim do Tabaco 34, 1149-041 Lisboa, Portugal. E-Mail: scholten@ispa.pt

---

<sup>1</sup> Analogous distinctions have been made between *independent* and *component-wise* evaluation (Tversky, 1969), *interdimensional* and *intradimensional* processing (Payne, 1976), *holistic* and *dimensional* choice strategies (Russo & Doshier, 1983), and *separate* and *joint* evaluation (Hsee, Loewenstein, Blount, & Bazerman, 1999).

**Discounting Theory and Anomalies in Intertemporal Choice**

We focus on intertemporal choices between pairs of single dated outcomes, one smaller-but-sooner (*SS*), the other larger-but-later (*LL*). An example is the choice between \$100 in 1 month and \$150 in 4 months. The outcomes are designated as  $x_S$  and  $x_L$  (\$100 and \$150), and their respective delays as  $t_S$  and  $t_L$  (1 and 4 months). We first describe how an elementary choice such as this is addressed by the discounted utility model, the normative standard for intertemporal choice, and then we discuss how successive reformulations of discounting theory have addressed preference patterns anomalous to this normative standard.

**Delay Discounting Model 1: Samuelson’s (1937) Discounted Utility Model**

In this model, intertemporal preferences are governed by the discounted utilities of the options. Discounted utilities are determined by first integrating the outcomes of choice with the baseline consumption level, assigning utilities to the consumption levels resulting from this asset integration, and then exponentially discounting these utilities as a function of the delays to the outcomes. Formally, the discounted utilities of *SS* and *LL* are given as

$$U_S = \delta^{t_S} u(c + x_S) + \delta^{t_L} u(c),$$

$$U_L = \delta^{t_S} u(c) + \delta^{t_L} u(c + x_L),$$

where  $0 < \delta < 1$  is an exponential discounting parameter,  $c$  is a constant baseline consumption level, and  $u$  is a concave utility function over consumption levels. The decision maker will be indifferent between *SS* and *LL* when

$$\delta^{t_L - t_S} = \frac{u(c + x_S) - u(c)}{u(c + x_L) - u(c)} \tag{1}$$

For elementary choices between *SS* and *LL*, the normative status of the discounted utility model derives largely from two of its features (for a more complete discussion, see Frederick, Loewenstein, & O’Donoghue, 2002): exponential discounting and asset integration. Exponential discounting ensures consistency in the treatment of time so that intertemporal preferences will not change merely because time has passed (Strotz, 1955–1956). The preference between two outcomes separated by a 3-month interval, for instance, should be the same regardless of whether the interval begins now, in 1 month, or in 9 months. Asset integration ensures consistency in the treatment of money. For instance, someone who prefers \$150 in 4 months to \$100 in 1 month should, in two separate choices between \$75 in 4 months and \$50 in 1 month, consistently prefer the former over the latter (see Prelec & Loewenstein, 1991).

**Delay Discounting Model 2: Laibson’s (1997) Quasi-Hyperbolic Discounting Model**

In this model, indifference between *SS* and *LL* is given by Equation 1 when both outcomes are delayed but not when *SS* is a smaller-*immediate* outcome. This is because a single “jolt” of discounting is added to the exponential discounting of delayed utilities:

$$U_S = u(c + x_S) + \beta \delta^{t_L} u(c),$$

$$U_L = u(c) + \beta \delta^{t_L} u(c + x_L),$$

where  $0 < \beta < 1$  is a quasi-hyperbolic discounting parameter. Thus, the decision maker will be indifferent between *SS* and *LL* when

$$\beta \delta^{t_L} = \frac{u(c + x_S) - u(c)}{u(c + x_L) - u(c)} \tag{2}$$

The implication of Equations 1 and 2 is the *immediacy effect* (Keren & Roelofsma, 1995; Read, Loewenstein, & Kalyanaraman, 1999; B. J. Weber & Chapman, 2005), also called *present-biased preference* (O’Donoghue & Rabin, 1999): There is more discounting over an interval that begins now than one that begins later. For instance, someone who is indifferent between \$150 in 3 months and \$100 now will prefer \$150 in 4 months to \$100 in 1 month.

**Delay Discounting Model 3: Loewenstein and Prelec’s (1992) Hyperbolic Discounting Model**

This model drops the assumptions of exponential discounting and asset integration. In this model, preferences are governed by the discounted *values* of the options. Discounted values are determined by first assigning values to *changes* in consumption level and then by hyperbolically discounting these values as a function of the delay to the outcomes. Formally, the discounted values of *SS* and *LL* are then given as

$$V_S = \prod_{i=1}^{t_S} \delta_i v(x_S),$$

$$V_L = \prod_{i=1}^{t_L} \delta_i v(x_L),$$

where  $v$  is a value function over changes in consumption level. For each time unit,  $i$ , there is a hyperbolic discounting parameter,  $0 < \delta_i < 1$ , which increases with  $i$ . At the indifference point,

$$\prod_{i=t_S+1}^{t_L} \delta_i = \frac{v(x_S)}{v(x_L)}$$

Hyperbolic discounting is motivated by the *common difference effect* (Loewenstein & Prelec, 1992): There is more discounting over an interval that begins earlier than one that begins later.<sup>2</sup> For instance, someone who is indifferent between \$150 in 4 months and \$100 in 1 month will prefer \$150 in 12 months to \$100 in 9

<sup>2</sup> Not counting the studies in which the effect of the delay to the later outcome is confounded with that of the interval *between* the outcomes (see Read, 2001), the common difference effect is seen in some (Green, Fristoe, & Myerson, 1994; Green, Myerson, & Macaux, 2005; Holt, Green, Myerson, & Estle, 2008; Keren & Roelofsma, 1995; Kirby & Herrnstein, 1995; Scholten & Read, 2006) but not, or not reliably, in others (Ahlbrecht & Weber, 1997; Baron, 2000; Holcomb & Nelson, 1992; Read, 2001; Read & Roelofsma, 2003). The common difference effect has been confirmed mostly in studies in which time periods were presented as delays proper (e.g., “in 1 year”) and disconfirmed mostly in studies in which time periods were presented as dates (e.g., “on December 1, 2007”). Recent studies show that using dates instead of delays indeed attenuates or eliminates the common difference effect (LeBoeuf, 2006; Read, Frederick, Orsel, & Rahman, 2005).

months. The common difference effect reduces to the immediacy effect when the earlier interval begins *now*.

Loewenstein and Prelec (1992) proposed a specific discount function to describe how  $\delta_i$  increases with  $i$ , and they also specified the properties of the value function  $v$ . Three of these value function properties are taken from prospect theory (Kahneman & Tversky, 1979). First, *reference dependence*, which is that outcomes are evaluated as gains and losses relative to a neutral reference point. These gains and losses are the changes in consumption level mentioned earlier. The neutral reference point usually corresponds to the baseline consumption level, in which case values are assigned to the outcomes as such. Second, *diminishing sensitivity*, which is that the marginal impact of an outcome decreases with the magnitude of the outcome (i.e.,  $v$  is concave over gains and convex over losses). Third, *loss aversion*, which is that losses loom larger than gains (i.e.,  $v$  is steeper for losses than for gains).<sup>3</sup>

Loewenstein and Prelec (1992) invoke both reference dependence and loss aversion to explain the *delay-speedup asymmetry*, which is that a positive outcome is discounted more, and a negative outcome discounted less, when it is delayed than when it is sped up over the same interval (Benzion, Rapoport, & Yagil, 1989; Shelley, 1993; see also Loewenstein, 1988, Experiment 3; Malkoc & Zauberman, 2006; E. U. Weber et al., 2007).<sup>4</sup> The intuition underlying this asymmetry can be conveyed by a thought experiment. Imagine you expect a parcel to be delivered today, but you learn it will be delayed for 1 week. How much would you need to be paid to compensate you for the pain of a late delivery? Call that amount  $C$ . Now imagine you expect a parcel to be delivered in 1 week, but you learn it can be sped up for delivery today. How much would you be willing to pay for the pleasure of an early delivery? Call that amount  $c$ . The delay-speedup asymmetry is that  $C$  is greater than  $c$ . Over the same interval, you will need more compensation for the pain of a late delivery than you will be willing to pay for the pleasure of an early one.

Loewenstein and Prelec's (1992) explanation of the delay-speedup asymmetry is as follows. People adapt to expected outcomes, so that the reference point for a given moment includes everything they expect to gain or lose at that moment. Receiving an amount when expected has, therefore, neutral value, because it has already been adapted to. For the same reason, however, *not* receiving the amount is a loss, which must be compensated by a gain at another moment. Specifically, when delaying a receipt, the earlier loss has to be compensated by a later gain. Aversion to the earlier loss drives up the magnitude of the compensating gain, leading to more discounting. Conversely, when speeding up a receipt over the same interval, the later loss has to be compensated by an earlier gain. Aversion to the later loss drives up the magnitude of the compensating gain, leading to less discounting.

Loewenstein and Prelec (1992) proposed that the value function has two properties in addition to those adopted from prospect theory. The first is that increasing the magnitude of both outcomes by the same multiplicative constant decreases the ratio between the values of  $x_S$  and  $x_L$ :

$$\frac{v(mx_S)}{v(mx_L)} < \frac{v(x_S)}{v(x_L)} \text{ iff } m > 1. \tag{3}$$

This property is motivated by the *absolute magnitude effect* (Loewenstein & Prelec, 1992; for a first demonstration, see Thaler,

1981): The discounting over an interval is inversely related to outcome magnitude. For instance, someone who is indifferent between \$150 in 4 months and \$100 in 1 month will prefer \$1,500 in 4 months to \$1,000 in 1 month. This is probably the most robust anomaly in intertemporal choice.<sup>6</sup>

The final property of Loewenstein and Prelec's (1992) value function is that reversing the sign of the outcomes from positive to negative decreases the ratio between the values of  $x_S$  and  $x_L$ :

$$\frac{v(-x_S)}{v(-x_L)} < \frac{v(x_S)}{v(x_L)} \text{ iff } x_S, x_L > 0. \tag{4}$$

This property is motivated by the *gain-loss asymmetry* (Loewenstein & Prelec, 1992): There is less discounting over a given interval when the outcomes are losses than when they are gains.<sup>7</sup> For instance, someone who is indifferent between receiving \$100 in 1 month and receiving \$150 in 4 months will prefer to pay \$100 in 1 month rather than pay \$150 in 4 months.

In sum, Loewenstein and Prelec's (1992) hyperbolic discounting model can accommodate several anomalies to the discounted utility model. There are, however, anomalies to discounting theory itself. These are identified in the next section.

### Why Discounting Does Not Work

In this section, we identify three anomalies to discounting theory. The first two rule out the alternative-based delay discounting models discussed in the previous section but not necessarily the "hybrid" (partly-alternative-partly-attribute-based) interval discounting model that we proposed recently (Scholten & Read, 2006) and that we discuss first. The third anomaly, however, rules out even this hybrid interval discounting model, which is our motivation for developing a fully attribute-based tradeoff model of intertemporal choice.

<sup>3</sup> Hyperbolic discounting captures diminishing sensitivity to *delays* (see Scholten & Read, 2006, Footnote 4; see also Takanishi, 2005).

<sup>4</sup> Given that this anomaly, like others below, is ascribed to the value function and not to the discount function, the term *discounting* loses its strict association with the discount function: It merely serves to describe an empirical regularity.

<sup>5</sup> The article by Loewenstein and Prelec (1992) contains a typographical error, in that their Inequality 18 reads "<" instead of ">".

<sup>6</sup> The studies by Kirby and Maraković (1995) are perhaps the only studies with humans in which the absolute magnitude effect was not corroborated. In animal studies, it has not been corroborated either (e.g., Grace, 1999; Green, Myerson, Holt, Slevin, & Estle, 2004; Ong & White, 2004; Richards, Mitchell, de Wit, & Seiden, 1997).

<sup>7</sup> Not counting the studies in which the effect of the *sign* of an outcome is confounded with that of *delaying or speeding up* an outcome (see Shelley, 1993), the gain-loss asymmetry is seen in some (Murphy, Vuchinich, & Simpson, 2001; Yates & Watts, 1975) but not, or not reliably, in others (Ahlbrecht & Weber, 1997; Benzion et al., 1989; Loewenstein, 1988; Shelley, 1993). The delay-speedup asymmetry is essentially a gain-loss asymmetry for compensations rather than outcomes: There is less discounting over a given interval when a compensation is to be paid (for speeding up a gain or delaying a loss) than when a compensation is to be received (for delaying a gain or speeding up a loss).

**Scholten and Read’s (2006) Interval Discounting Model**

Delay discounting models hold that the utility or value of an outcome is discounted as a function of its delay only. This is incompatible with evidence of *nonadditive discounting*, which implies that utilities or values are discounted not only as a function of the delays to the outcomes but also as a function of the interval between them. To illustrate, suppose that two outcomes are separated by a single day, and we use two different procedures for assessing the discounting over that day. In one procedure, we obtain a single measure of the discounting that takes place over the whole day. In the other, we obtain separate measures of the discounting that takes place over the four day segments (morning, afternoon, evening, and night), and then combine them into a single measure. Discounting is *additive* when the two procedures yield the same result. This is predicted by delay discounting models. On the other hand, discounting is *subadditive* when there is less discounting over the whole day than over its four segments, and it is *superadditive* when there is more discounting over the whole day than over the segments. There is evidence of both subadditive and superadditive discounting, as discussed below.

In the interval discounting model, discounting is governed by both intervals and delays. Let  $t_L - t_S$  be a unit interval and  $n$  be the number of equal-length intervals into which this unit interval is divided. According to the model, the decision maker will be indifferent between *SS* and *LL* when

$$\delta^{(n)} = \frac{v(x_S)}{v(x_L)}$$

where  $v$  is the same value function as in the hyperbolic discounting model, and

$$\delta^{(n)} = \sqrt[n]{\prod_{i=1}^n \delta_i}$$

which is the average discounting over the subintervals. Under exponential discounting,  $\delta_i = \delta$ , under quasi-hyperbolic discounting,  $\delta_1 = \beta\delta$  and  $\delta_{i > 1} = \delta$ , and under hyperbolic discounting,  $\delta_i$  increases with  $i$ , but, in each case,  $\delta^{(n)}$  is independent of  $n$ , which means that discounting is additive over partitions of the interval  $t_L - t_S$ . Discounting is subadditive, however, when  $\delta^{(n)}$  decreases with  $n$  (more discounting over partitions) and superadditive when  $\delta^{(n)}$  increases with  $n$  (less discounting over partitions).

The assumption of additive discounting has been undermined by a series of recent studies (Kinari, Ohtake, & Tsutsui, 2009; Read, 2001; Read & Roelofsma, 2003; Scholten & Read, 2006; Zauber- man, Kim, Malkoc, & Bettman, 2009). The pattern that emerges from these studies is that  $\delta^{(n)}$  decreases with  $n$  (subadditivity) up to a point, beyond which it increases with  $n$  (superadditivity). We proposed a specific discount function to describe how  $\delta_i$  increases with  $i$  (hyperbolic discounting) and how  $\delta^{(n)}$  first decreases and then increases with  $n$  (Scholten & Read, 2006). That discount function introduces a big change to discounting theory: It is not defined over delays,  $t$ , but rather over *differences* between weighted delays, that is,  $w(t_L) - w(t_S)$ , where  $w$  is a time-weighting function. The discount function thus incorporates direct comparisons along the time attribute.<sup>8</sup>

As a preliminary, we next discuss three anomalies to discounting theory. The first two show that subadditivity or superadditivity

can outweigh hyperbolic or exponential discounting and diminishing or constant sensitivity to outcomes. This rules out delay discounting models but not necessarily the interval discounting model, because it incorporates direct comparisons between delays into the discount function. The third anomaly, however, shows that, over a given range of intervals, superadditivity in intervals can change into superadditivity in compensations when the compensations change from large to small. This rules out even the interval discounting model, because it implies that an accurate model of intertemporal choice must incorporate direct comparisons between outcomes as well. Collectively, the anomalies demand a fully attribute-based tradeoff model of intertemporal choice. We now discuss the anomalies one by one.

**Anomaly 1: Strong Subadditivity**

We verified the first anomaly to discounting theory in a web-administered survey administered to 196 participants. All participants made eight choices, including the following:

$A = \$100$  in 19 months or  $B = \$118$  in 22 months;

$C = \$100$  in 16 months or  $D = \$136$  in 22 months.

Each choice was between  $(a, t - k)$  and  $(a + kb, t)$ , where the first element of each option is money and the second element is time. The options were obtained by setting  $a = \$100$ ,  $b = \$18$ ,  $t = 22$ , and  $k = 1$  for *A–B*, and  $k = 2$  for *C–D*. This is illustrated in Figure 1.

Table 1 gives the results. In the choice between *A* and *B*, a majority (60%) preferred *SS* to *LL*. Among the 34% who had a different preference between *C* and *D*, a majority (82%) shifted from *SS* to *LL*, so that a majority (62%) now preferred *LL* to *SS*. (A similar result was obtained by Rubinstein, 2003, Experiment 1.) This choice pattern is incompatible with hyperbolic discounting and diminishing sensitivity to outcomes, as we show next. (More generally, as we prove in Appendix A, it is incompatible with hyperbolic or exponential discounting and diminishing or constant sensitivity to outcomes.)

Consider a person who prefers *A* to *B*, or  $(a, t - 1)$  to  $(a + b, t)$ . For this person,  $(a, t - 2)$  should be preferred to  $(a + b, t - 1)$ , because, by hyperbolic discounting,  $\delta_{t-1} < \delta_t$ . In addition,  $(a + b, t - 1)$  should be preferred to  $(a + 2b, t)$ , because, by diminishing sensitivity to outcomes,  $v(a + 2b) - v(a + b) < v(a + b) - v(a)$ . By transitivity, therefore,  $(a, t - 2)$  should be preferred to  $(a + 2b, t)$ , or *C* should be preferred to *D*, contrary to our observation. We call the observed choice pattern *strong subadditivity*, because 1 interval of  $k$  periods counts less than  $k$  intervals of 1 period, and this interval effect is strong enough to outweigh the countervailing effects of hyperbolic discounting and diminishing sensitivity to outcomes.

Table 2 shows strong subadditivity for four pairs of choices. The pair on the left is between the shorter interval options ( $k = 1$ ), and the one on the right is between the longer interval options ( $k = 2$ ).

<sup>8</sup> Killeen (2009) suggested that the interval discounting model needs to adjust the parameters of the discount function to accommodate the absolute magnitude effect. However, the interval discounting model adopts the value function from Loewenstein and Prelec (1992), which captures the absolute magnitude effect with the property in Inequality 3.

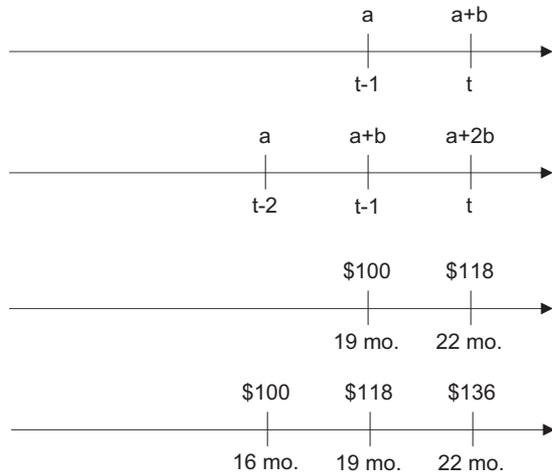


Figure 1. Strong subadditivity: The countervailing effects of hyperbolic discounting and diminishing sensitivity to outcomes ( $a$  and  $b$  are units of money,  $t$  is time).

In each case, there is a significant shift in preference from *SS* to *LL* as  $k$  increases.

Table 2 also shows the absolute magnitude effect: The proportion of *SS* choices increases as we move down the rows, from larger to smaller outcomes. As mentioned earlier, the absolute magnitude effect is probably the most robust anomaly in intertemporal choice. It may therefore serve as a benchmark against which to evaluate the size of the interval effect (strong subadditivity). Averaged across all 12 possible comparisons between intervals of equal length, the proportion of preference shifts due to outcome magnitude was .25, and the difference between the proportion of *SS* choices for smaller amounts and larger ones was .19. (Most of the absolute magnitude effect was located in the contrast between \$250, on the one hand, and \$100, \$50, and \$25, on the other; see Table 2.) In contrast, averaged across all four comparisons between sooner outcomes of equal magnitude, the proportion of preference shifts due to interval length was .28, and the difference between the proportion of *SS* choices for shorter intervals and longer ones was .19. The interval effect, therefore, was as large as the absolute magnitude effect. Moreover, the interval effect outweighed the countervailing effects of hyperbolic or exponential discounting and diminishing or constant sensitivity to outcomes. No such burden was placed on the absolute magnitude effect.

**Anomaly 2: Strong Superadditivity**

We verified the second anomaly to discounting theory in another web-administered survey administered to a further 233 participants. They made each of the following choices:

- A = \$8,250 in 12 months or B = \$10,250 in 24 months;
- C = \$6,250 in 12 months or D = \$10,250 in 36 months;
- E = \$4,250 in 12 months or F = \$10,250 in 48 months.

Each choice was between  $(a - kb, t)$  and  $(a, t + k)$ . The specific options were obtained by setting  $a = \$10,250$ ,  $b = \$2,000$ ,  $t = 12$ ,

and  $k = 1$  for A–B,  $k = 2$  for C–D, and  $k = 3$  for E–F. This is illustrated in Figure 2.

Table 3 gives the results. In the choice between A and B, a majority (60%) preferred *LL* to *SS*. Among the 38% who had a different preference between C and D, a majority (66%) shifted from *LL* to *SS*, so that a majority (52%) now preferred *SS* to *LL*. Similarly, among the 23% who had a different preference between E and F, a majority (76%) shifted from *LL* to *SS*, so that a majority (52%) now preferred *SS* to *LL*. This choice pattern is incompatible with hyperbolic discounting and diminishing sensitivity to outcomes, as we show next.

Consider a person who prefers B to A, or  $(a, t + 1)$  to  $(a - b, t)$ . For this person,  $(a, t + 2)$  should be preferred to  $(a - b, t + 1)$ , because, by hyperbolic discounting,  $\delta_{t+2} > \delta_{t+1}$ . In addition,  $(a - b, t + 1)$  should be preferred to  $(a - 2b, t)$ , because, by diminishing sensitivity to outcomes,  $v(a - b) - v(a - 2b) > v(a) - v(a - b)$ . By transitivity, therefore,  $(a, t + 2)$  should be preferred to  $(a - 2b, t)$ , or D should be preferred to C, contrary to our observation. We call this anomaly to discounting theory *strong superadditivity*, because 1 interval of  $k$  periods counts more than  $k$  intervals of 1 period, and this interval effect is strong enough to outweigh the countervailing effects of hyperbolic discounting and diminishing sensitivity to outcomes.

**Anomaly 3: Inseparability**

The third anomaly is that people do not treat time independently from money: A given interval may be treated as long or short, depending on whether the compensation is small or large. Indirect evidence of this interdependence effect comes from two studies of intransitive intertemporal choice.

In one study (Roelofsma & Read, 2000), small amounts of Dfl 7, 8, 9, and 10 (Dutch Guilders, before the Euro) could be earned by waiting 1, 2, 4, and 7 weeks, respectively. In the other study (Scholten & Read, 2006, Experiment 1), large amounts of £500, 525, 550, and 575 could be earned by waiting 1, 2, 3, and 4 weeks, respectively. Thus, across the studies, the waiting periods were similar but the amounts to be earned were not. The results are as follows: In the small-amounts study, the most common intransitivity was subadditivity (choice of *SS* from adjacent pairs but choice of *LL* from more distant pairs), whereas, in the large-amount study, the most common intransitivity was superadditivity

Table 1  
Strong Subadditivity

|                               |    | Choice over long interval CD  |           |     |
|-------------------------------|----|---|-----------|-----|
|                               |    | SS  | LL        |     |
| Choice over short interval AB | SS | 63  | <b>55</b> | 118 |
|                               | LL | <b>12</b>   | 67        | 79  |
|                               |    | 75  | 122       | 197 |
|                               |    | $\chi^2(1) = 7.72, p = .01^a$<br>$\chi^2(1) = 27.60, p = .00^b$<br>$\chi^2(1) = 11.21, p = .00^c$ |           |     |

<sup>a</sup> Chi-square test on of the right-hand margin. <sup>b</sup> McNemar’s chi-square test on the off-diagonals. <sup>c</sup> Chi-square test on the lower margin.

Table 2  
Strong Subadditivity and the Absolute Magnitude Effect

| Option pair A-B                               | $P_{SS}$ | Option pair C-D                           | $P_{SS}$ | $P_{SS \rightarrow LL}$ | $P_{LL \rightarrow SS}$ | $\chi^2(1)^a$ | $p$ |
|---|----------|---|----------|-------------------------|-------------------------|---------------|-----|
| \$250 in 19 months or \$295 in 22 months.     | .29      | \$250 in 16 months or \$340 in 22 months. | .13      | .19                     | .04                     | 20.45         | .00 |
| \$100 in 19 months or \$118 in 22 months.     | .60      | \$100 in 16 months or \$136 in 22 months. | .38      | .28                     | .06                     | 27.60         | .00 |
| \$50 in 19 months or \$59 in 22 months.       | .62      | \$50 in 16 months or \$68 in 22 months.   | .44      | .22                     | .04                     | 25.92         | .00 |
| \$25.00 in 19 months or \$29.50 in 22 months. | .68      | \$25 in 16 months or \$34 in 22 months.   | .47      | .25                     | .04                     | 31.50         | .00 |

Note.  $N = 196$ .

<sup>a</sup> McNemar's chi-square test of correlated proportions.

(choice of *LL* from adjacent pairs but choice of *SS* from more distant pairs).

Our interpretation of the above results is that people underweigh small differences between adjacent options in comparison with larger differences between more distant options and that what counts as a small difference along one attribute depends on the difference along the other attribute. In the large-amount study, participants treated time differences between adjacent options as small, yielding superadditivity *in intervals*. In the small-amount study, however, participants treated money differences between adjacent options as small, yielding superadditivity *in compensations*, or *subadditivity in intervals*. We next report direct evidence of this interdependence effect.

We conducted a web-administered survey study with 205 participants. We created two experimental conditions, shown in Table 4. In both conditions, different amounts could be earned by waiting 1, 2, 3,

or 4 weeks (corresponding to options *A, B, C,* and *D,* respectively). In one condition, the amounts were large, and in the other they were small, but in constant proportion. As shown in Table 4, each condition included all six possible pairwise choices between the four options.

Given six pairwise choices between four options, there are 64 possible choice patterns. These can be classified by two attributes: Presence or absence of superadditivity, and presence or absence of subadditivity. (For ease of exposition, these terms refer to superadditivity and subadditivity *in intervals*.) The combination of these attributes yields four classes of choice patterns: superadditive (16), subadditive (16), both superadditive and subadditive (8), and neither superadditive nor subadditive (24; these are all the *transitive* choice patterns). An example from each class is shown in Table 5.

Superadditivity occurs when *LL* is chosen over adjacent intervals but *SS* is chosen over the whole interval. In the first example, choice is superadditive (*AB* and *BC* vs. *AC*; *AB* and *BD* vs. *AD*; *AB, BC,* and *CD* vs. *AD*). Subadditivity occurs when *SS* is chosen over adjacent intervals but *LL* is chosen over the whole interval. In the second example, choice is subadditive (*AC* and *CD* vs. *AD*; *BC* and *CD* vs. *BD*; *AB, BC,* and *CD* vs. *AD*). The third example shows a choice pattern that is both superadditive (*BC* and *CD* vs. *BD*) and subadditive (*AB* and *BD* vs. *AD*). The final example shows a transitive choice pattern. Choice patterns were classified by superadditivity and subadditivity within each experimental condition.

We want to show that intervals are treated as short with large amounts, thus increasing the likelihood of superadditivity, and

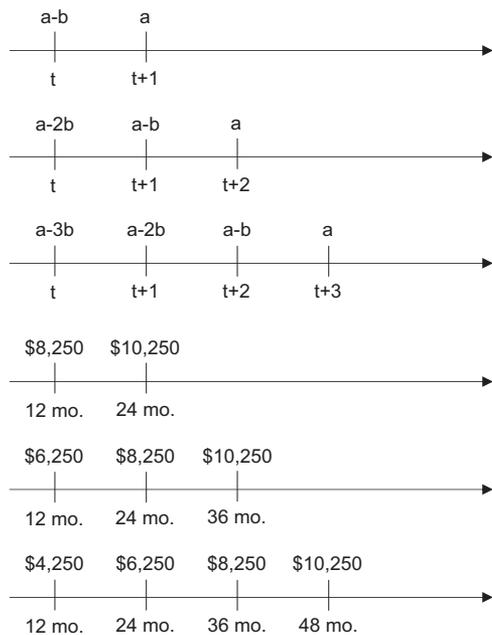


Figure 2. Strong superadditivity: The countervailing effects of hyperbolic discounting and diminishing sensitivity to outcomes (*a* and *b* are units of money, *t* is time).

Table 3  
Strong Superadditivity

|                                      | Choice over long interval <i>CD</i> |                               | Choice over long interval <i>EF</i> |                                |           |           |           |     |
|--------------------------------------|-------------------------------------|-------------------------------|-------------------------------------|--------------------------------|-----------|-----------|-----------|-----|
|                                      | <i>SS</i>                           | <i>LL</i>                     | <i>SS</i>                           | <i>LL</i>                      |           |           |           |     |
| Choice over short interval <i>AB</i> | <i>SS</i>                           | 63                            | <b>30</b>                           | 93                             | <i>SS</i> | 80        | <b>13</b> | 93  |
|                                      | <i>LL</i>                           | <b>59</b>                     | 81                                  | 140                            | <i>LL</i> | <b>41</b> | 99        | 140 |
|                                      |                                     | 122                           | 111                                 | 233                            |           | 121       | 112       | 233 |
|                                      |                                     | $\chi^2(1) = 9.48, p = .00^a$ |                                     |                                |           |           |           |     |
|                                      |                                     | $\chi^2(1) = 9.45, p = .00^b$ |                                     | $\chi^2(1) = 14.52, p = .00^b$ |           |           |           |     |
|                                      |                                     | $\chi^2(1) = 0.52, p = .47^c$ |                                     | $\chi^2(1) = 0.35, p = .56^c$  |           |           |           |     |

<sup>a</sup> Chi-square test on of the right-hand margin. <sup>b</sup> McNemar's chi-square test on the off-diagonals. <sup>c</sup> Chi-square test on the lower margin.

Table 4  
Six Option Pairs With Small Outcomes and Six Option Pairs With Large Outcomes

| Small outcomes         |  |  |                        | Large outcomes          |  |  |                         |
|------------------------|--|--|------------------------|-------------------------|--|--|-------------------------|
| SS                     |  |  | LL                     | SS                      |  |  | LL                      |
| A = \$26.40 in 1 week  |  |  | B = \$27.70 in 2 weeks | A = \$510.00 in 1 week  |  |  | B = \$535.50 in 2 weeks |
| B = \$27.70 in 2 weeks |  |  | C = \$29.10 in 3 weeks | B = \$535.50 in 2 weeks |  |  | C = \$562.30 in 3 weeks |
| C = \$29.10 in 3 weeks |  |  | D = \$30.60 in 4 weeks | C = \$562.30 in 3 weeks |  |  | D = \$590.40 in 4 weeks |
| A = \$26.40 in 1 week  |  |  | C = \$29.10 in 3 weeks | A = \$510.00 in 1 week  |  |  | C = \$562.30 in 3 weeks |
| B = \$27.70 in 2 weeks |  |  | D = \$30.60 in 4 weeks | B = \$535.50 in 2 weeks |  |  | D = \$590.40 in 4 weeks |
| A = \$26.40 in 1 week  |  |  | D = \$30.60 in 4 weeks | A = \$510.00 in 1 week  |  |  | D = \$590.40 in 4 weeks |

Note. To obtain the small outcomes, we divided the large outcomes by 19.30 and rounded to the nearest \$0.10.

treated as long with small amounts, thus increasing the likelihood of subadditivity. One way to test this interdependence effect is by comparing the two experimental conditions (small amounts and large amounts) with respect to each variant of intransitivity (superadditivity and subadditivity). This is done in the top left and bottom right cross-tabulations of Table 6. As expected, superadditivity was more likely with large amounts, whereas subadditivity was more likely with small amounts. Another way to test the interdependence effect is by comparing the two variants of intransitivity within each experimental condition. This is done in the top right and bottom left cross-tabulations of Table 6. As expected, superadditivity was more likely with large amounts, whereas subadditivity was more likely with small amounts. In sum, we obtain evidence of a new anomaly to discounting theory, which we call *inseparability*.

We also obtained evidence for the absolute magnitude effect. Averaged across all six comparisons between intervals of equal length, the proportion of preference shifts due to outcome magnitude was .43, and the difference between the proportions of SS choices for small amounts and large ones was .38. In contrast, averaged across all four comparisons testing the interdependence effect (inseparability), the proportion of preference shifts was .23, and the difference between the proportions in the “Yes” categories of Table 6 was .08. Thus, the interdependence effect was modest in comparison with the absolute magnitude effect. This does not imply, however, that it is less important for uncovering the psychological process driving these effects.

**Conclusion**

Both past evidence and Anomalies 1 and 2 show that people make direct comparisons along the time attribute. The solution of the interval discounting model is to introduce these comparisons into the discount function. However, the discount function still weighs the values of the outcomes independently of one another, as in any other discounting model. Therefore, the options are directly compared along the time attribute but not along the money attribute. Although the interval discounting model is not unique in combining alternative-wise valuation with attribute-wise comparisons (e.g., Mellers & Biagini, 1994; Shafir, Osherson, & Smith, 1993), it is unique in proposing that direct comparisons are made along one attribute (time) but not along the other (money). This seems psychologically implausible: If people make direct comparisons along the time attribute, why would they not make direct comparisons along the money attribute as well? Anomaly 3 shows that they do. This motivates a drastic departure from discounting and its alternative-based approach to intertemporal choice. We take an attribute-based approach, in which differences along the time attribute are directly weighted against differences along the money attribute. This enables us to accommodate *all* anomalies, those that discounting theory can address and those it cannot.

**Tradeoffs Without Discounting**

We present two versions of the tradeoff model. In the *single reference point model*, both outcomes are evaluated relative to the

Table 5  
Classification of Choice Patterns: Four Examples

| Superadditivity |    |        |    | Subadditivity |    |        |    | Both   |    |        |    | Neither |    |        |    |
|-----------------|----|--------|----|---------------|----|--------|----|--------|----|--------|----|---------|----|--------|----|
|                 |    | Option |    |               |    | Option |    |        |    | Option |    |         |    | Option |    |
| Option          | B  | C      | D  | Option        | B  | C      | D  | Option | B  | C      | D  | Option  | B  | C      | D  |
| A               | LL | SS     | SS | A             | SS | SS     | LL | A      | SS | SS     | LL | A       | SS | SS     | LL |
| B               |    | LL     | LL | B             |    | SS     | LL | B      |    | LL     | SS | B       |    | SS     | LL |
| C               |    |        | LL | C             |    |        | SS | C      |    |        | LL | C       |    |        | LL |

Note. Looking at option pairs spanning adjacent intervals (e.g., AB and BD) and the option pair spanning the whole interval (e.g., AD), intransitivity occurs when the choice for the whole interval differs from the choices for all adjacent intervals. Specifically, superadditivity occurs when LL is chosen over all adjacent intervals but SS is chosen over the whole interval; subadditivity occurs when the reverse is true.

Table 6  
Inseparability

|                                  |     |                                       |     |     |                                       |                                     |     |    |     |                                   |
|----------------------------------|-----|---------------------------------------|-----|-----|---------------------------------------|-------------------------------------|-----|----|-----|-----------------------------------|
|                                  | 205 | 180                                   | 25  |     |                                       |                                     | 188 | 17 | 205 |                                   |
| $\chi^2(1) = 6.81,$<br>$p = .01$ | 161 | 144                                   | 17  | No  | Superadditivity with<br>large amounts | No                                  | 150 | 11 | 161 | $\chi^2(1) = 14.88,$<br>$p = .00$ |
|                                  | 44  | 36                                    | 8   | Yes |                                       | Yes                                 | 38  | 6  | 44  |                                   |
|                                  |     | No                                    | Yes |     |                                       | No                                  | Yes |    |     |                                   |
|                                  |     | Superadditivity with<br>small amounts |     |     |                                       | Subadditivity with large<br>amounts |     |    |     |                                   |
|                                  |     | No                                    | Yes |     |                                       | No                                  | Yes |    |     |                                   |
| $\chi^2(1) = 0.35,$<br>$p = .56$ | 176 | 155                                   | 21  | No  | Subadditivity with<br>small amounts   | No                                  | 161 | 15 | 176 | $\chi^2(1) = 3.43,$<br>$p = .06$  |
|                                  | 29  | 25                                    | 4   | Yes |                                       | Yes                                 | 27  | 2  | 29  |                                   |
|                                  | 205 | 180                                   | 25  |     |                                       |                                     | 188 | 17 | 205 |                                   |

Note. McNemar’s chi-square test on the off-diagonals.

same reference point, the baseline consumption level. This model accommodates all preference patterns but the delay-speedup asymmetry. In the *multiple reference point model*, the outcome expected at a given time is evaluated relative to a reference point that has moved away from the baseline consumption level. This model accommodates the delay-speedup asymmetry as well. The single reference point model is a special case of the multiple reference point model, but, for simplicity of exposition, we present them separately.

**The Single Reference Point Model**

In the tradeoff model, intertemporal choices are made by directly weighing time differences against money differences. For a choice between a smaller-sooner outcome and a larger-later one,  $f > 0$  denotes the “advantage” (Tversky, 1969, p. 41) that the smaller-sooner gain or the larger-later loss has along the time attribute, and  $g > 0$  denotes the advantage that the larger-later gain or the smaller-sooner loss has along the money attribute. The decision maker will prefer the larger-later gain or the smaller-sooner loss if  $f < g$ , will prefer the smaller-sooner gain or the larger-later loss if  $f > g$ , and will be indifferent otherwise.

The advantages  $f$  and  $g$  are weighted time differences and money differences. The weighing is done by two intra- and two inter-attribute weighing functions. The two intra-attribute weighing functions are a time-weighing function  $w$ , which weighs delays against one another, and a value function  $v$ , which weighs outcomes against one another. The two inter-attribute weighing functions are tradeoff functions  $Q_{T|X}$  and  $Q_{X|T}$ , which weigh the absolute difference between weighted delays, called the *effective interval* and denoted  $T$ , against the absolute difference between valued outcomes, called the *effective compensation* and denoted  $X$ . Effective differences are neither subjective nor objective: They are weights put on time and money differences in the decision making process. Described in terms of  $v$ ,  $w$ ,  $Q_{T|X}$ , and  $Q_{X|T}$ , the decision maker will be indifferent between *SS* and *LL* when

$$Q_{T|X}(w(t_L) - w(t_S)) = \begin{cases} Q_{X|T}(v(x_L) - v(x_S)) & \text{if } x_L > x_S > 0, \\ Q_{X|T}(v(x_S) - v(x_L)) & \text{if } 0 > x_S > x_L. \end{cases} \tag{5}$$

The two intra-attribute weighing functions are increasing; the two inter-attribute weighing functions are nondecreasing. We next discuss additional properties of these functions.

**Reference dependence.** Both intra-attribute weighing functions,  $v$  and  $w$ , are reference dependent, which places the tradeoff model in the broader class of reference-dependent choice models (e.g., Köszegi & Rabin, 2006; Tversky & Kahneman, 1991). Delays are treated as positive deviations from the present, and, in the single reference point model, outcomes are treated as positive or negative deviations from the baseline consumption level. Reference dependence is the prerequisite for three additional properties of  $v$  and  $w$  that are discussed next: Decreasing absolute sensitivity, increasing proportional sensitivity, and constant loss aversion.

**Diminishing absolute sensitivity to delays and outcomes.** Diminishing absolute sensitivity has its roots in the Weber–Fechner law. For constant absolute increases of  $t$  and  $x > 0$ ,  $w(t)$  and  $v(x)$  increase by decreasing absolute amounts. For instance, adding \$1 to \$2 yields a greater absolute increase in value than adding \$1 to \$200. Similarly, for constant absolute decreases of  $x < 0$ ,  $v(x)$  decreases by decreasing absolute amounts. Diminishing absolute sensitivity to outcomes is a standard assumption in psychology and economics, and diminishing absolute sensitivity to delays is also captured by hyperbolic discounting (see Footnote 3).

Diminishing absolute sensitivity produces two preference patterns. The first, which we call the *proportional magnitude effect*, is shared with all discounting models: Increasing the magnitude of both outcomes by the same additive constant changes indifference into a preference for *SS* (gains) or for *LL* (losses).

The second pattern, the *common difference effect*, is that increasing both delays by the same additive constant changes indif-

ference into a preference for *LL* (gains) or for *SS* (losses). For gains, we have

$$Q_{T|x}(w(t_L) - w(t_S)) = Q_{x|T}(v(x_L) - v(x_S)) \text{ and}$$

$$Q_{T|x}(w(a + t_L) - w(a + t_S)) < Q_{x|T}(v(x_L) - v(x_S)), \text{ so that}$$

$$Q_{T|x}(w(a + t_L) - w(a + t_S)) < Q_{T|x}(w(t_L) - w(t_S)), \text{ or}$$

$$w(a + t_L) - w(a + t_S) < w(t_L) - w(t_S) \text{ iff } a > 0.$$

By diminishing absolute sensitivity, the additive constant decreases the absolute difference between weighted delays and, thus, decreases the weight of the time attribute.

**Augmenting proportional sensitivity to delays and outcomes.** For constant proportional increases of  $t$  and  $x > 0$ ,  $w(t)$  and  $v(x)$  increase by increasing absolute amounts. For instance, doubling \$100 yields a greater absolute increase in value than doubling \$1. Similarly, for constant proportional decreases of  $x < 0$ ,  $v(x)$  decreases by increasing absolute amounts.

As with diminishing absolute sensitivity, augmenting proportional sensitivity produces two preference patterns. The first, which we call the *common ratio effect*, is again shared with all discounting models: Increasing both delays by the same multiplicative constant changes indifference into a preference for *SS* (gains) or for *LL* (losses).<sup>9</sup>

The second pattern, the absolute magnitude effect, is that increasing the magnitude of both outcomes by the same multiplicative constant changes indifference into a preference for *LL* (gains) or for *SS* (losses).<sup>10</sup> For gains, we have

$$Q_{T|x}(w(t_L) - w(t_S)) = Q_{x|T}(v(x_L) - v(x_S)) \text{ and}$$

$$Q_{T|x}(w(t_L) - w(t_S)) < Q_{x|T}(v(mx_L) - v(mx_S)), \text{ so that}$$

$$Q_{x|T}(v(mx_L) - v(mx_S)) > Q_{x|T}(v(x_L) - v(x_S)), \text{ or}$$

$$v(mx_L) - v(mx_S) > v(x_L) - v(x_S) \text{ iff } m > 1.$$

By augmenting proportional sensitivity, the multiplicative constant increases the absolute difference between outcome values and, thus, increases the weight of the outcome attribute.

Although the tradeoff model ascribes the absolute magnitude effect to augmenting proportional sensitivity, the hyperbolic discounting model ascribes it to *increasing elasticity*: For constant proportional increases of  $x > 0$ ,  $v(x)$  increases by increasing proportional amounts, and, similarly, for constant proportional decreases of  $x < 0$ ,  $v(x)$  decreases by increasing proportional amounts (increasing elasticity is also captured by Inequality 3).<sup>11</sup> This difference between the models is important: In Appendix B, we prove that, when combined with diminishing absolute sensitivity, augmenting proportional sensitivity is satisfied by a broader class of value functions than increasing sensitivity. In particular, the power value function, since long a popular device in formal analyses of choice (Tversky, 1967b), satisfies the former but not the latter.

**Constant loss aversion.** The final property of the value function is constant loss aversion (Tversky & Kahneman, 1991): Reversing the sign of an outcome from positive to negative increases the magnitude of its value by a multiplicative constant, that is,

$v(-x) = -\Lambda v(x)$ , where  $x \geq 0$  and  $\Lambda > 1$ . In the single reference point model, constant loss aversion produces the gain-loss asymmetry: The sign reversal changes indifference into a preference for *SS*. Formally, we have

$$Q_{T|x}(w(t_L) - w(t_S)) = Q_{x|T}(v(x_L) - v(x_S)) \text{ and}$$

$$Q_{T|x}(w(t_L) - w(t_S)) < Q_{x|T}(v(-x_S) - v(-x_L)), \text{ so that}$$

$$Q_{x|T}(v(-x_S) - v(-x_L)) > Q_{x|T}(v(x_L) - v(x_S)), \text{ or}$$

$$v(-x_S) - v(-x_L) > v(x_L) - v(x_S).$$

By constant loss aversion, the multiplicative constant  $\Lambda$  increases the absolute difference between outcome values and, thus, increases the weight of the outcome attribute.

Constant loss aversion is the only variant of loss aversion that dissociates the asymmetric steepness of the value function (Kahneman & Tversky, 1979) from asymmetric elasticity (Loewenstein & Prelec, 1992; asymmetric elasticity is also captured by Inequality 4). Evidence on risky choice is equivocal on whether the asymmetry between gains and losses can be fully captured by constant loss aversion: Some studies suggest it can (e.g., Tversky & Kahneman, 1992), others that it cannot (e.g., Abdellaoui, Bleichrodt, & L'Haridon, 2008; Fennema & van Assen, 1999). In the tradeoff model, a qualitative account of the gain-loss asymmetry and, as discussed below, the delay-speedup asymmetry can entirely do without asymmetric elasticity. Drawing on considerations of parsimony, therefore, we only invoke constant loss aversion.

Constant loss aversion concludes our discussion of the intra-attribute weighing functions  $w$  and  $v$ , which transform objective intervals and compensations into effective ones. Thus, by applying the time-weighing function  $w$ , the difference between 4 months and 1 month may be *effectively* the same as that between 12 months and 6 months. Similarly, by applying the value function  $v$ , the difference between \$150 and \$100 may be *effectively* the same as that between \$1,500 and \$1,250, or that between -\$125 and -\$150. The effective intervals must now be weighted against effective compensations. This is done by the inter-attribute weighing functions discussed next.

<sup>9</sup> The common ratio effect cannot be expected to persist unconditionally: The difference between 40 and 20 years clearly matters, but the difference between 400 and 200 years clearly does not, because both delays are outside human experience. In risky choice, the common ratio effect pertains to the probability attribute and is anomalous to the expected utility model (Prelec & Loewenstein, 1991).

<sup>10</sup> Like the common ratio effect, the absolute magnitude effect cannot be expected to persist unconditionally, as Amos Tversky pointed out to Prelec and Loewenstein (1991): "Doubling an award of 10 billion dollars clearly matters less than doubling an award of \$500,000; likewise, the difference between losing one or two million dollars is less important than that between losing one or two thousand, to most people at least" (Prelec & Loewenstein, 1991, p. 778).

<sup>11</sup> Diminishing absolute sensitivity, augmenting proportional sensitivity, and increasing elasticity can be given a geometric interpretation. For  $x > 0$ , augmenting absolute sensitivity means that  $v(x)$  is a concave function of  $x$ . Augmenting proportional sensitivity means that  $v(x)$  is a convex function of  $\log(x)$ . Increasing elasticity means that  $\log(v(x))$  is a convex function of  $\log(x)$ .

**Inter-attribute sensitivity.** The tradeoff function  $Q_{T|X}$  can be given as

$$Q_{T|X}(T) = \begin{cases} \kappa_S T & \text{if } T < \varepsilon_{T|X} \\ \kappa_L T & \text{if } T \geq \varepsilon_{T|X}, \end{cases} \quad (6.1)$$

where  $0 < \kappa_S < \kappa_L$ , and  $\varepsilon_{T|X} > 0$  is a threshold. The step function considered by Tversky (1969) arises when  $\kappa_S = 0$ , producing a lexicographic semi-order. An example of  $Q_{T|X}$ , as given by Equation 6.1, is the bipartite function in Figure 3. Similarly, the tradeoff function  $Q_{X|T}$  can be given as

$$Q_{X|T}(X) = \begin{cases} \mu_S X & \text{if } X < \varepsilon_{X|T} \\ \mu_L X & \text{if } X \geq \varepsilon_{X|T}, \end{cases} \quad (6.2)$$

where  $0 < \mu_S < \mu_L$ , and  $\varepsilon_{X|T} > 0$  is a threshold. The scaling constants  $\kappa_S$ ,  $\kappa_L$ ,  $\mu_S$ , and  $\mu_L$  serve two functions: Scaling effective intervals and effective compensations to a common currency, and underweighing sub-threshold effective differences in comparison with supra-threshold effective differences. If only the relative magnitude of scaled effective differences is of interest, one of these scaling constants can be set to unity.

Like the inter-attribute threshold in González-Vallejo’s (2002) stochastic difference model, the intra-attribute thresholds  $\varepsilon_{T|X}$  and  $\varepsilon_{X|T}$  are affected by “choice context” (González-Vallejo, 2002, p. 137). Although choice context can include many things, we suggest that it includes the choice set itself, elsewhere called the “local context” (Tversky & Simonson, 1993, p. 1185). In line with Rubinstein (2003) and Leland (2002), who introduced the concept of *similarity* to the study of intertemporal choice, we propose that the intra-attribute thresholds  $\varepsilon_{T|X}$  and  $\varepsilon_{X|T}$  are affected by the relative similarity of the outcomes and the delays in the local context. If the outcomes are more similar to one another than the delays, there will be a bias in favor of *SS* (gains) or *LL* (losses), that is,  $X < \varepsilon_{X|T}$  but  $T \geq \varepsilon_{T|X}$ . Conversely, if the delays are more similar to one another than the outcomes, there will be a bias in favor of *LL* (gains) or *SS* (losses), that is,  $T < \varepsilon_{T|X}$

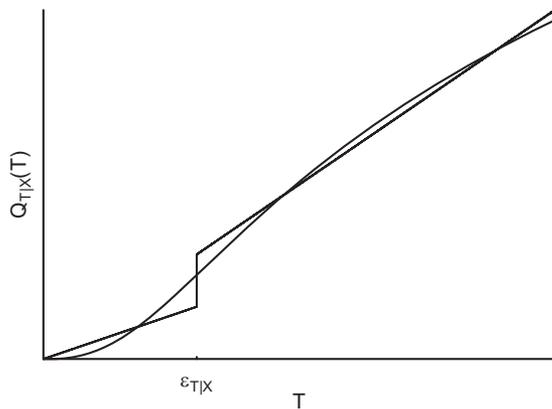


Figure 3. An example of a bipartite and a continuous tradeoff function  $Q_{T|X}$ . The bipartite tradeoff function is given as  $Q_{T|X}(T) = \kappa_S T$  if  $T < \varepsilon_{T|X}$  and  $Q_{T|X}(T) = \kappa_L T$  if  $T \geq \varepsilon_{T|X}$ , where  $0 < \kappa_S < \kappa_L$ , and  $\varepsilon_{T|X} > 0$  is a context-dependent threshold. The continuous tradeoff function is given as  $Q_{T|X}(T) = (\kappa/\alpha_{T|X})\log(1 + \alpha_{T|X}(T/\vartheta_{T|X})^{\vartheta_{T|X}})$ , where  $\alpha_{T|X} > 0$  is subadditivity and  $\vartheta_{T|X} > 1$  is superadditivity. The inflection occurs when  $T = (\vartheta_{T|X}((\vartheta_{T|X} - 1)/\alpha_{T|X})^{1/\vartheta_{T|X}})$ , so that  $\alpha_{T|X}$  and  $\vartheta_{T|X}$  take on the role of the context-dependent  $\varepsilon_{T|X}$  in locating the shallower and steeper regions of  $Q_{T|X}$ .

but  $X \geq \varepsilon_{X|T}$ . Similarity is *comparative* in nature. Therefore, its role in intertemporal choice underscores the basic proposition of the tradeoff model, which is that people make direct comparisons along both attributes (time and money).

We ascribe the three anomalies to discounting theory identified earlier (strong subadditivity, strong superadditivity, and inseparability) to biases induced by the effective similarity between the options along the attributes, as discussed next.

**Anomaly 1: Strong subadditivity.** With reference to Figure 1, participants preferred  $(a, t - 1)$  to  $(a + b, t)$ , so that, by diminishing absolute sensitivity to delays ( $w$ ), diminishing absolute sensitivity to outcomes ( $v$ ), and transitivity, they should also have preferred  $(a, t - 2)$  to  $(a + 2b, t)$ . However, they had the reverse preference. Our explanation is that, in the context of a 19- or 22-month wait, \$100 was likely to be similar to \$118, but, in the context of a 16- or 22-month wait, \$100 was less likely to be similar to \$136. The similarity of the outcomes produces a bias in favor of *SS*: An extra wait for a similar amount of money is unjustified. Without this bias, an extra wait for an extra amount of money may be justified, and the results show it was.

If there were *no* bias over the 1-period interval, that is,  $X \geq \varepsilon_{X|T}$ , then, by diminishing absolute sensitivity to delays, diminishing absolute sensitivity to outcomes, and transitivity, we would have

$$\begin{aligned} \mu_L(v(136) - v(100)) - \kappa_L(w(22) - w(16)) < \\ \mu_L(v(118) - v(100)) - \kappa_L(w(22) - w(19)) < 0. \end{aligned}$$

That is, if *SS* is preferred over the 1-period interval, then it should also be preferred over the 2-period interval. Instead, however, we had

$$\begin{aligned} \mu_S(v(118) - v(100)) - \kappa_L(w(22) - w(19)) < 0 < \\ \mu_L(v(136) - v(100)) - \kappa_L(w(22) - w(16)). \end{aligned}$$

That is, *LL* was preferred over the 2-period interval, but *SS* was preferred over the 1-period interval. Our interpretation is that, over the 1-period interval, the similarity of the outcomes produced a bias in favor of *SS*, that is,  $X < \varepsilon_{X|T}$ , so that the effective compensation was weighted by  $\mu_S < \mu_L$ , not by  $\mu_L$ .

**Anomaly 2: Strong superadditivity.** With reference to Figure 2, participants preferred  $(a, t + 1)$  to  $(a - b, t)$ , so that, by diminishing absolute sensitivity to delays ( $w$ ), diminishing absolute sensitivity to outcomes ( $v$ ), and transitivity, they should also have preferred  $(a, t + 2)$  to  $(a - 2b, t)$ . However, they had the reverse preference. Our explanation is that, the context of a \$8,250 or \$10,250 payoff, a 12-month wait was likely to be similar to a 24-month wait, but, in the context of a \$6,250 or \$10,250 payoff, a 12-month wait was likely to be similar to a 36-month wait. The similarity of the delays produces a bias in favor of *LL*: A similar wait for a much larger amount of money is justified. Without this bias, much longer waits for much larger amounts of money may not be justified, and the results show they were not.

If there were *no* bias over the 1-period interval, that is,  $T \geq \varepsilon_{T|X}$ , then, by diminishing absolute sensitivity to delays, diminishing absolute sensitivity to outcomes, and transitivity, we would have

$$\begin{aligned} \mu_L(v(10,250) - v(6,250)) - \kappa_L(w(36) - w(12)) > \\ \mu_L(v(10,250) - v(8,250)) - \kappa_L(w(24) - w(12)) > 0. \end{aligned}$$

That is, if *LL* is preferred over the 1-period interval, then it should also be preferred over the 2-period interval. Instead, however, we had

$$\begin{aligned} \mu_L(v(10,250) - v(8,250)) - \kappa_S(w(24) - w(12)) &> 0 > \\ \mu_L(v(10,250) - v(6,250)) - \kappa_L(w(36) - w(12)). \end{aligned}$$

That is, *SS* was preferred over the 2-period interval, but *LL* was preferred over the 1-period interval. Our interpretation is that, over the 1-period interval, the similarity of the delays produced a bias in favor of *LL*, that is,  $T < \epsilon_{T|X}$ , so that the effective interval was weighted by  $\kappa_S < \kappa_L$ , not by  $\kappa_L$ .

We suggested that a 12-months wait was similar to a 24-months wait in the context of significant large payoffs. These same delays would probably not be similar if the payoffs were reduced by a factor of 100, to \$82.50 and \$102.50. Such inter-attribute relativity is the third, and culminating, anomaly to discounting theory.

**Anomaly 3: Inseparability.** The tradeoff functions  $Q_{T|X}$  and  $Q_{X|T}$  can give rise to intransitive intertemporal choice. Consider three pairwise choices between three delayed gains. If a person is indifferent between *SS* and *MM* and indifferent between *MM* and *LL* but prefers *LL* to *SS* (subadditivity), we have

$$\begin{aligned} \kappa_L(w(t_M) - w(t_S)) &= \mu_S(v(x_M) - v(x_S)) \text{ and} \\ \kappa_L(w(t_L) - w(t_M)) &= \mu_S(v(x_L) - v(x_M)), \text{ so that} \\ \kappa_L(w(t_L) - w(t_S)) &= \mu_S(v(x_L) - v(x_S)), \text{ but, instead, we have} \\ \kappa_L(w(t_L) - w(t_S)) &< \mu_L(v(x_L) - v(x_S)), \text{ so that} \\ \mu_S &< \mu_L. \end{aligned}$$

Alternatively, if the person is indifferent between *SS* and *MM* and indifferent between *MM* and *LL* but prefers *SS* to *LL* (superadditivity), we have

$$\kappa_S < \kappa_L.$$

Our evidence of inseparability shows that which variant of intransitivity is most likely to occur (subadditivity or superadditivity) depends on the significance of the payoffs over a given range of intervals.

Superadditivity was more likely with large outcomes, whereas subadditivity was more likely with small ones. Our explanation is as follows. With large outcomes, the delays were more similar to one another than the outcomes. This induced a bias in favor of *LL*, that is,  $\kappa_S < \kappa_L$ , producing superadditivity. Conversely, with small outcomes, the outcomes were more similar to one another than the delays. This induced a bias in favor of *SS*, that is,  $\mu_S < \mu_L$ , producing subadditivity.

Inter-attribute sensitivity and the role of effective similarity in shaping inter-attribute sensitivity conclude the exposition of the single reference point model. We next generalize the tradeoff model to accommodate the delay-speedup asymmetry.

### The Multiple Reference Point Model

In the single reference point model, both outcomes are treated as deviations from the baseline consumption level, so that gains and losses coincide with the actual amounts to be received or paid. However, the reference point of the value function may not be located at the baseline consumption level. It can be

affected, for instance, by the formulation of the options (Kahneman & Tversky, 1979) and by the goals (Heath, Larrick, & Wu, 1999), aspirations (Payne, Laughhunn, & Crum, 1980, 1981), and expectations (Kahneman & Tversky, 1979) of the decision maker. We have already discussed how expectations play a role in Loewenstein and Prelec's (1992) analysis of the delay-speedup asymmetry. In their analysis, the reference point for a given moment includes everything they expect to gain or lose at that moment. Thus, receiving or paying an amount when expected is evaluated as a zero event, but *not* receiving or paying the amount (in the event of a delay or speedup) is evaluated as a loss or gain (see also Shelley, 1993).

In Loewenstein and Prelec's (1992) analysis, the reference point moves to the expected consumption level, meaning that people treat a compensated change from that level (compensated by a change from the baseline consumption level at another moment) exactly as they would treat an uncompensated change. It is unlikely, however, that someone who does not receive an expected \$100 today, knowing that he or she is to receive \$150 in 3 months, will treat the foregone \$100 exactly as he or she would if it was uncompensated. We suggest that people do not treat compensated and uncompensated changes from the expected consumption level in the same way, and that the presence of a compensation moderates the degree to which the reference point changes. That is, the effective reference point is likely to be intermediate between the expected consumption level (including the expected \$100) and the baseline consumption level. This proposal of incomplete adaptation is in agreement with Hoch and Loewenstein (1991); Strahilevitz and Loewenstein (1998); and Arkes, Hirshleifer, Jiang, and Lim (2008), who all discuss it in contexts other than the delay or speedup of consumption. We next develop a version of the tradeoff model with incomplete adaptation to expected outcomes.

In the single reference point model, the value of receiving  $x$  is  $v(x)$  and the value of not receiving it is  $v(0)$ , so that the value difference between receiving and not receiving  $x$  is  $v(x) - v(0) = v(x)$ . This is no longer the case if one adapts to receiving  $x$ . In the case of *complete* adaptation, the value of receiving  $x$  is  $v(0)$  and the value of not receiving it is  $v(-x)$ , so that the value difference between receiving and not receiving  $x$  is  $v(0) - v(-x) = -v(-x)$ . In general, however, the degree of adaptation, as indicated by reference-point shift  $r$ , will fall somewhere on the continuum between no adaptation, that is,  $r = 0$ , and complete adaptation, that is,  $r = x$ . The value of receiving  $x$  is then  $v(x - r)$  and the value of not receiving it is  $v(-r)$ , so that the value difference between receiving and foregoing  $x$  is  $v(x - r) - v(-r)$ . This expression, which also appears in Hoch and Loewenstein's (1991) analysis of impatience, serves to generalize the tradeoff model.

To accommodate the delay-speedup asymmetry, we propose that an expected outcome is evaluated relative to a reference point that has moved away from the baseline consumption level (adaptation), whereas the outcome that compensates for the delay or speedup of the expected outcome continues to be evaluated relative to the baseline consumption level. The decision maker will then be indifferent between *SS* and *LL* when

$$Q_{T|x}(T) = \begin{cases} Q_{x|T}(v(\hat{x}_L) - [v(x_S - r_S) - v(-r_S)]) & \text{if } r_S, x_S, \hat{x}_L > 0 \text{ (delaying a gain)} \\ Q_{x|T}(v(\hat{x}_S) - [v(x_L - r_L) - v(-r_L)]) & \text{if } r_L, \hat{x}_S, x_L < 0 \text{ (speeding up a loss)} \\ Q_{x|T}([v(x_L - r_L) - v(-r_L)] - v(\hat{x}_S)) & \text{if } r_L, \hat{x}_S, x_L > 0 \text{ (speeding up a gain)} \\ Q_{x|T}([v(x_S - r_S) - v(-r_S)] - v(\hat{x}_L)) & \text{if } r_S, x_S, \hat{x}_L < 0 \text{ (delaying a loss)}, \end{cases} \tag{7}$$

where  $\hat{x}_L$  and  $\hat{x}_S$  are the outcomes that compensate, respectively, for the delay and speedup of the expected outcome. If there is no adaptation to the expected outcome, that is,  $r_S = r_L = 0$ , the multiple reference point model in Equation 7 reduces to the single

reference point model in Equation 5. Under the assumption of constant sensitivity, as customarily invoked in formal analyses of the delay-speedup asymmetry (Loewenstein, 1988; Shelley, 1993), Equation 7 becomes

$$\kappa_L(t_L - t_S) = \begin{cases} \mu_L(\hat{x}_L - [(x_S - r_S) - \Lambda(-r_S)]) & \text{if } r_S, x_S, \hat{x}_L > 0 \text{ (delaying a gain)} \\ \mu_L(\Lambda\hat{x}_S - [\Lambda(x_L - r_L) - (-r_L)]) & \text{if } r_L, \hat{x}_S, x_L < 0 \text{ (speeding up a loss)} \\ \mu_L([(x_L - r_L) - \Lambda(-r_L)] - \hat{x}_S) & \text{if } r_L, \hat{x}_S, x_L > 0 \text{ (speeding up a gain)} \\ \mu_L([\Lambda(x_S - r_S) - (-r_S)] - \Lambda\hat{x}_L) & \text{if } r_S, x_S, \hat{x}_L < 0 \text{ (delaying a loss)}. \end{cases}$$

Rearranging,

$$\kappa(t_L - t_S) = \begin{cases} (\hat{x}_L - x_S) - (\Lambda - 1)r_S & \text{if } r_S, x_S, \hat{x}_L > 0 \text{ (delaying a gain)} & (8.1) \\ \Lambda(\hat{x}_S - x_L) + (\Lambda - 1)r_L & \text{if } r_L, \hat{x}_S, x_L < 0 \text{ (speeding up a loss)} & (8.2) \\ (x_L - \hat{x}_S) + (\Lambda - 1)r_L & \text{if } r_L, \hat{x}_S, x_L > 0 \text{ (speeding up a gain)} & (8.3) \\ \Lambda(x_S - \hat{x}_L) - (\Lambda - 1)r_S & \text{if } r_S, x_S, \hat{x}_L < 0 \text{ (delaying a loss)}, & (8.4) \end{cases}$$

where  $\kappa = \kappa_L/\mu_L$ . The compensation that one demands for delaying a gain or speeding up a loss, or that one offers for speeding up a gain or delaying a loss, can therefore be divided into a compensation that would be acceptable whether there were no adaptation and a compensation for adaptation. To derive the delay-speedup asymmetry for gains, we solve Equation 8.1 for the later outcome that compensates for the delay of the earlier one,

$$\hat{x}_L = \kappa(t_L - t_S) + x_S + (\Lambda - 1)r_S,$$

and we solve Equation 8.3 for the earlier outcome that compensates for the speedup of the later one,

$$\hat{x}_S = x_L + (\Lambda - 1)r_L - \kappa(t_L - t_S).$$

Setting the outcome to be sped up,  $x_L$ , equal to the outcome that compensates for the delay,  $\hat{x}_L$ , and simplifying, we get

$$\hat{x}_S = x_S + (\Lambda - 1)(r_S + r_L).^{12}$$

It is evident that  $\hat{x}_S > x_S$  as long as there is adaptation *and* loss aversion. In the same way, the delay-speedup asymmetry for losses can be derived from Equations 8.2 and 8.4.

The multiple reference point model introduces two improvements over current discounting models. The first, which can be introduced to discounting models as well, is that it allows for any degree of adaptation to an expected outcome. The second improvement, however, is a genuine strength of the tradeoff model. The hyperbolic discounting model ascribes the delay-speedup asymmetry to the value function being steeper in losses than in gains (loss aversion) but the gain-loss asymmetry to the value function being more elastic in losses than in gains: Two different devices for two similar phenomena (see Footnote 7). The tradeoff model ascribes both anomalies to loss aversion. The degree of adaptation to the expected outcome determines the degree to which the

delay-speedup asymmetry outweighs the gain-loss asymmetry. The tradeoff model thus offers a parsimonious and psychologically plausible explanation of these two similar phenomena.

The multiple reference point model concludes the qualitative development of the tradeoff model. In the next section, we develop a fully parametric specification of the model.

### Putting Tradeoffs to Work

In this section, we provide the parametric specification of the value function and the time-weighting function. The tradeoff functions are given by Equations 6.1 and 6.2, but, as we mention in the General Discussion section, they may, under certain conditions, be replaced by a single, continuous tradeoff function.

### The Value Function

The qualitative development of the tradeoff model is consistent with a broad range of value functions. However, the range of possible value functions can be narrowed by broadening the set of qualitative requirements that the value function must satisfy. Considering six qualitative requirements, we propose the following value function:

<sup>12</sup> Adaptation imposes a constraint on the magnitude of the loss-aversion coefficient,  $\Lambda$ . Because the equations have to satisfy  $\hat{x}_S < x_L$ , the multiple reference point model implies that, in the extreme case of complete adaptation to expected outcomes,  $x_S < x_L(2 - \Lambda)/\Lambda$ . Thus, we have  $1 < \Lambda < 2$ . In the case of incomplete adaptation, the constraint is relaxed, and in the case of no adaptation, the constraint is removed.

$$v(x) = \begin{cases} \frac{1}{\gamma} \log(1 + \gamma x) & \text{if } x \geq 0 \\ -\frac{\Lambda}{\gamma} \log(1 + \gamma(-x)) & \text{if } x < 0, \end{cases} \quad (9)$$

where  $\Lambda > 1$  is constant loss aversion and  $\gamma > 0$  is diminishing absolute sensitivity to outcomes. The six qualitative requirements that this value function satisfies are discussed below for the case of indifference between  $x_S$  at  $t_S$  and  $x_L$  at  $t_L$ . Unless stated otherwise, the exposition concerns the case of gains ( $x_S, x_L > 0$ ).

1. To maintain indifference, the absolute compensation  $x_L - x_S$  is larger for a longer  $t_L$ . This requires that  $v$  be an increasing function.

2. The absolute compensation  $x_L - x_S$  is larger for a larger  $x_S$ . This is the proportional magnitude effect and requires that  $v$  be a reference-dependent function satisfying diminishing absolute sensitivity.

3. The proportional compensation  $(x_L - x_S)/x_S$  is smaller for a larger  $x_S$ . This is the absolute magnitude effect, and requires that  $v$  be a reference-dependent function satisfying augmenting proportional sensitivity.<sup>13</sup>

4. The absolute compensation  $x_L - x_S$  for  $x_S, x_L > 0$  (gains) is larger than the absolute compensation  $x_S - x_L$  for  $x_S, x_L < 0$  (losses). This is the gain-loss asymmetry, and requires that  $v$  be a reference-dependent function satisfying constant loss aversion.

5. The compensations implied by the value function are bounded by two extremes, *constant compensation* and *proportional compensation*. Constant compensation occurs when the absolute compensation  $x_L - x_S$  and the magnitude of  $x_S$  are independent. For instance,  $x_L - x_S = \$10$  regardless of whether  $x_S$  is \$100 or \$900. Proportional compensation occurs when the proportional compensation  $(x_L - x_S)/x_S$  and the magnitude of  $x_S$  are independent. For instance,  $x_L - x_S = \$10$  when  $x_S$  is \$100, but  $x_L - x_S = \$90$  when  $x_S$  is \$900, so that  $(x_L - x_S)/x_S = .10$ . In the case of constant compensation, there would be no proportional magnitude effect and a maximum absolute magnitude effect; in the case of proportional compensation, there would be no absolute magnitude effect and a maximum proportional magnitude effect.

The value function in Equation 9 implies constant compensation as  $\gamma$  approaches 0 and proportional compensation as  $\gamma$  approaches infinity. When  $\gamma$  goes toward 0, the value function approaches *constant sensitivity*, in which case value increases linearly with outcome, that is,  $v(x) = x$  if  $x \geq 0$  and  $v(x) = \Lambda x$  if  $x < 0$ . When  $\gamma$  goes toward infinity, the value function approaches *insensitivity*, in which case any outcome is treated as if it had no value at all, that is,  $v(x) = 0$  for any  $x$ . These are *limiting* cases, however, meaning that constant and proportional compensation will never really exist.

6. The present value of  $x_L$  increases with the magnitude of  $x_L$ , but increases more steeply for a shorter  $t_L$  (Stevenson, 1986, 1992, 1993; see also Anderson & Shanteau, 1970; Shanteau, 1974; Tversky, 1967a, 1967b). That is, if we plot  $x_S$  as a function of  $x_L$  for different delays  $t_L$ , we get a diverging fan of lines, with a steeper slope corresponding to a shorter  $t_L$ . This is shown in Figure 4. These *bilinear interaction effects*, as they are called, impose severe constraints on the value function of the tradeoff model, because the present value is given as

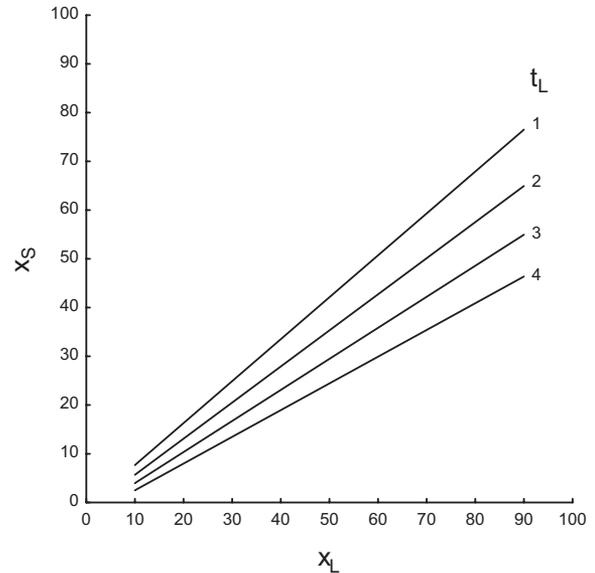


Figure 4. Bilinear interaction effects: The present value  $x_S$  of a delayed outcome  $x_L$  increases with the magnitude of the delayed outcome  $x_L$ , but it increases more steeply for a shorter delay  $t_L$ .

$$x_S = \begin{cases} v^{-1}[v(x_L) - Q_{X|T}^{-1}(Q_{T|X}(w(t_L)))] & \text{if } x_L > x_S > 0 \\ v^{-1}[v(x_L) + Q_{X|T}^{-1}(Q_{T|X}(w(t_L)))] & \text{if } 0 > x_S > x_L. \end{cases}$$

The subtraction or addition of  $v(x_S)$  and  $Q_{X|T}^{-1}(Q_{T|X}(w(t_L)))$  will not produce the diverging fan of lines in  $x_S - x_L$  space unless  $v^{-1}$  is an exponential function, which means that  $v$  itself must be a logarithmic function, as the one in Equation 9.<sup>14</sup> The diverging fan of lines in Figure 4 was drawn with  $\gamma = 0.15$  and all other functions ( $Q_{X|T}$ ,  $Q_{T|X}$ , and  $w$ ) set to identity.

The value function in Equation 9 also exhibits *decreasing* elasticity. In the tradeoff model, however, a qualitative account of the anomalies can entirely do without decreasing elasticity.

### The Time-Weighing Function

The time-weighing function is an adaptation of the value function to the weighing of delays:

<sup>13</sup> An example of a value function that does not satisfy this requirement is the double logarithmic value function

$$v(x) = \begin{cases} \frac{e}{\gamma} \log(\log(e + \gamma x)) & \text{if } x \geq 0 \\ -\Lambda \frac{e}{\gamma} \log(\log(e + \gamma(-x))) & \text{if } x < 0. \end{cases}$$

We owe this example to Craig Desjardins.

<sup>14</sup> An example of a value function that does not satisfy this requirement is the power value function

$$v(x) = \begin{cases} \frac{1}{1 + \gamma} x^{\frac{1}{1+\gamma}} & \text{if } x \geq 0 \\ -\frac{\Lambda}{1 + \gamma} (-x)^{\frac{1}{1+\gamma}} & \text{if } x < 0. \end{cases}$$

$$w(t) = \frac{1}{\tau} \log(1 + \tau t), \tag{10}$$

where  $\tau > 0$  is diminishing absolute sensitivity to delays. Diminishing absolute sensitivity is bounded by *constant sensitivity*, when  $\tau$  approaches 0, and *insensitivity*, as  $\tau$  approaches infinity. This time-weighting function is consistent with the functional form of discounting theory: Loewenstein and Prelec's (1992) hyperbolic discount function  $d(t) = (1 + \tau t)^{-\kappa/\tau}$  can be derived by combining  $w(t)$  with exponential discounting over weighted rather than stated delays, that is,  $d(t) = e^{-\kappa w(t)}$ .

Logarithmic specifications of  $w$  have recently been proposed by other authors as well. Takanishi (2005) proposes  $w(t) = \rho \log(1 + \tau t)$ , which reduces to Equation 10 when  $\rho = 1/\tau$ , and to  $w(t) = \log(1 + t)$ , proposed by Zauberman et al. (2009), when  $\rho = \tau = 1$ . These other authors, however, refer to  $w$  as a time-*perception* function rather than a time-*weighing* function. This is not just terminology: Time perceptions are subjective, whereas time weights, like probability weights in prospect theory, are neither subjective nor objective. They are not objective, because twice a delay will generally not receive twice the decision weight. Nor are they subjective, because twice a delay may be perceived as twice as long, especially when the delay is stated rather than experienced, and yet receive less than twice the decision weight.

### General Discussion

Discounting theory can accommodate many anomalies to Samuelson's (1937) discounted utility model, the normative standard for intertemporal choice. However, we identified three anomalies that it cannot accommodate. These are strong subadditivity, strong superadditivity, and inseparability. To account for these anomalies, we developed the tradeoff model of intertemporal choice. The tradeoff model is a drastic departure from discounting theory. It relinquishes the idea that people make intertemporal choices by comparing the discounted values of available options. According to the tradeoff model, people make intertemporal choices by weighing how much more they will receive or pay if they wait longer against how much longer the wait will be, or, conversely, how much less they will receive or pay if they do not wait longer against how much shorter the wait will be. The tradeoff model offers a parsimonious and psychologically plausible explanation of all anomalies that discounting theory can and cannot address. In this General Discussion section, we further strengthen our case for the tradeoff model by focusing on two of its main features: attribute-based choice and context dependence.

The tradeoff model has as its closest relatives the additive difference model proposed by Tversky (1969) and the stochastic difference model proposed by González-Vallejo and colleagues (González-Vallejo, 2002; González-Vallejo, Bonazzi, & Shapiro, 1996; González-Vallejo & Reid, 2006; González-Vallejo, Reid, & Schiltz, 2003). Both the tradeoff model and the stochastic difference model can be viewed as refinements of the additive difference model. However, they describe the inter-attribute weighing process in different ways, and this difference is important for the modeling of intertemporal choice.

In the tradeoff model, the *absolute* difference between valued outcomes is weighted against the *absolute* difference between weighted delays. In the stochastic difference model, when applied

to tradeoffs between time and money, the *proportional* difference between valued outcomes is weighted against the *proportional* difference between weighted delays. The weighing of proportional differences obviates the need for one tradeoff parameter (to scale effective intervals and effective compensations into a common currency). However, it does not eliminate the need for an increasingly elastic value function (to accommodate the absolute magnitude effect), and it *creates* the need for an increasingly elastic time-weighting function (to accommodate the common ratio effect). By weighing absolute differences, as in the tradeoff model, these problems are immediately resolved.

Whereas the stochastic difference model weighs proportional differences, and the tradeoff model weighs absolute differences, the interval discounting model—the only formulation of discounting theory that accommodates subadditivity and superadditivity—weighs absolute differences along the time attribute against proportional differences along the money attribute. If we neutralize all possible parameters of the stochastic difference model, the interval discounting model, and the tradeoff model, respectively, the decision maker will be indifferent between *SS* and *LL* when

$$\frac{t_L - t_S}{t_S} = \frac{x_L - x_S}{x_S}, \text{ or}$$

$$\frac{t_S}{t_L} = \frac{x_S}{x_L} \text{ (stochastic difference model),}$$

$$\frac{1}{1 + \kappa(t_L - t_S)} = \frac{x_S}{x_L}, \text{ or}$$

$$\kappa(t_L - t_S) = \frac{x_L - x_S}{x_S} \text{ (interval discounting model), and}$$

$$\kappa(t_L - t_S) = \begin{cases} x_L - x_S & \text{if } x_L > x_S > 0 \\ x_S - x_L & \text{if } 0 > x_S > x_L \end{cases} \text{ (tradeoff model).}$$

Thus, the successive formulations involve a progressive elimination of proportionality from the inter-attribute weighing process.

The tradeoff functions given by Equations 6.1 and 6.2 are bipartite functions, linear below the threshold and linear above it. The implication is additivity in the absence of threshold effects. This may be an oversimplification: There is evidence of subadditivity in the absence of any apparent threshold effects (Read, 2001; Read & Roelofsma, 2003; Scholten & Read, 2006; Zauberman et al., 2009). This may be an indication that the tradeoff functions are S-shaped—convex over small effective differences (superadditivity) but concave over larger ones (subadditivity)—and that the S-shaped curve over longer effective intervals is shallower than the S-shaped curve over larger effective compensations (subadditivity *in intervals*). If, in addition, the S-shaped curve over short effective intervals is steeper than the S-shaped curve over small effective compensations (*superadditivity* in intervals), the tradeoff process can be described by a single S-shaped tradeoff function  $Q_{TIX}$  (for an example of such a function, see Figure 3):

$$Q_{TIX}(w(t_L) - w(t_S)) = \begin{cases} v(x_L) - v(x_S) & \text{if } x_L > x_S > 0, \\ v(x_S) - v(x_L) & \text{if } 0 > x_S > x_L. \end{cases}$$

This is the attribute-based analogue of the partly-alternative-partly-attribute-based interval discounting model. Although both

models capture subadditivity, they do not really explain it: The origin of subadditivity deserves closer examination.

An S-shaped tradeoff function  $Q_{T|x}$  can also capture what Ebert and Prelec (2007) called *time sensitivity*. Greater time sensitivity consists in less discounting of near-future outcomes and more discounting of far-future ones. Ebert and Prelec’s studies, in which the effect of the delay to the later outcome was confounded with that of the interval between the outcomes (see Footnote 2), show that time sensitivity can be affected by the experimental context. For instance, unlimited deliberation time leads to greater time sensitivity than limited deliberation time, and a within-subject design leads to greater time sensitivity than a between-subjects design. If we invoke the S-shaped tradeoff function  $Q_{T|x}$  to deal with this evidence, greater time sensitivity corresponds to a more pronounced superadditivity and/or a less pronounced subadditivity (see the caption to Figure 3). This interpretation ties in with the evidence from one of our own studies (Scholten & Read, 2006, Experiment 2): Using a within-subject design, we observed a succession of superadditivity and subadditivity over intervals of increasing length. The nature and origin of time sensitivity, and its context dependence, deserves closer examination as well.

The experimental context also influences the overall level of discounting (for an earlier discussion, see Frederick et al., 2002). Consider, for instance, a choice-based matching procedure, in which one aspect (usually the outcome) of one option is repeatedly adjusted, and the point of indifference between *SS* and *LL* is inferred from the participant’s choices. (At the indifference point, the options are “matched” on preference.) A series of recent studies show that adjusting  $x_S$  in ascending order yields a higher level of discounting than adjusting it in descending order (Robles & Vargas, 2007, 2008; Robles, Vargas, & Bejarano, 2009).<sup>15</sup> A datum like this may prompt two legitimate, and perhaps not mutually exclusive, reactions.

One is to consider the direction in which  $x$  is adjusted as “theoretically irrelevant” (Frederick et al., 2002, p. 384) and to work out methodological solutions. For instance, one might take the average of the two indifference points as a best guess of the true indifference point (e.g., Green, Fry, & Myerson, 1994). Also, one might use other matching procedures, for example, one in which  $x$  is adjusted in a random order (e.g., Robles & Vargas, 2007) or one in which  $x$  is adjusted up and down as a function of the participant’s choices (a *titration* procedure; e.g., Read, 2001), and see whether these procedures yield similar levels of discounting.

The second reaction is to take the effect of the direction in which  $x$  is adjusted as something that requires a theoretical explanation. In Robles and Vargas’s (2007, 2008) experiments, the starting point of the adjustment process was different. On the first trial of the descending procedure,  $x_S$  was equal to  $x_L$ , and, on subsequent trials, began deteriorating; on the first trial of the ascending procedure,  $x_S$  was close to 0, and, on subsequent trials, began improving. The descending procedure may have emphasized that one is foregoing  $x_L$  in exchange for  $x_S$ , that is, that one is speeding up a gain. In the terminology of the tradeoff model, the amount of  $x_S$  at the indifference point, denoted  $\hat{x}_S$ , would be given by

$$Q_{T|x}(T) = Q_{x|T}([v(x_L - r_L) - v(-r_L)] - v(\hat{x}_S)).$$

For the ascending procedure, on the other hand, the indifference point would be given by

$$Q_{T|x}(T) = Q_{x|T}(v(x_L) - v(\hat{x}_S)).$$

Comparing the two procedures, adaptation and loss aversion would drive up  $\hat{x}_S$  from the descending procedure, and thus reduce discounting. In this way, Roble and Vargas’s finding fits neatly into the tradeoff model, without the need to invoke biases such as anchoring or indeed any errors on the part of the decision makers.

Attribute-based models are a departure from alternative-based models and models that combine alternative-wise valuation with attribute-wise comparisons. A prominent instance of the latter approach is Mellers and Biagini’s (1994) contrast-weighting theory. When applying this theory to tradeoffs between time and money, a discounted value  $V$  is assigned to each option but, in computing the discounted value, the outcome values receive a contrast weight that is inversely related to the absolute difference between the delays to the outcomes, and the discount factors receive a contrast weight that is inversely related to the absolute difference between the outcomes. Letting  $w_t(t_L - t_S)$  and  $w_x(x_L - x_S)$  be the contrast weights for outcome values and discount factors, respectively, the decision maker will be indifferent between *SS* and *LL* when

$$\left( \prod_{i=t_S+1}^{t_L} \delta_i \right)^{w_x(x_L - x_S)} = \left( \frac{v(x_S)}{v(x_L)} \right)^{w_t(t_L - t_S)}.$$

Contrast-weighting theory does not need increasing elasticity to accommodate the absolute magnitude effect: Increasing both outcomes by the same multiplicative constant increases the absolute difference between the outcomes and therefore decreases the contrast weight for the discount factors, so that

$$\left( \prod_{i=t_S+1}^{t_L} \delta_i \right)^{w_x(mx_L - mx_S)} > \left( \frac{v(mx_S)}{v(mx_L)} \right)^{w_t(t_L - t_S)} \text{ iff } m > 1.$$

However, contrast-weighting theory cannot, without revision, avoid invoking asymmetric elasticity to accommodate the gain-loss asymmetry. Specifically, it must assume that the contrast weights are inversely related to *effective*, rather than objective, differences along the attributes: Loss aversion increases the effective difference along the outcome attribute and therefore decreases the contrast weight for the discount factors. This revision, however, severely complicates the analysis of the delay-speedup asymmetry, in that reference-point shifts have an impact not only on the outcome values but also on the contrast weight for the discount factors. In sum, for the modeling of intertemporal choice, contrast-weighting theory raises as many problems as it resolves.

Even so, contrast-weighting theory is able to predict a classic preference pattern, and the tradeoff model must be able to predict it as well. This is the *comparability effect* (see Busemeyer, Jessup, & Dimperio, 2007), an idealized version of which is given in Figure 5. The effect occurs when choosing between one *standard* option with a constant outcome and delay and another *variable* option with a variable outcome and delay. The curve relating the outcome of the variable option to the probability of choosing it is

<sup>15</sup> In Robles and Vargas’s (2008) study, the difference between the two matching procedures persisted when they did not terminate at the indifference point but continued until a preestablished number of adjustments was reached.

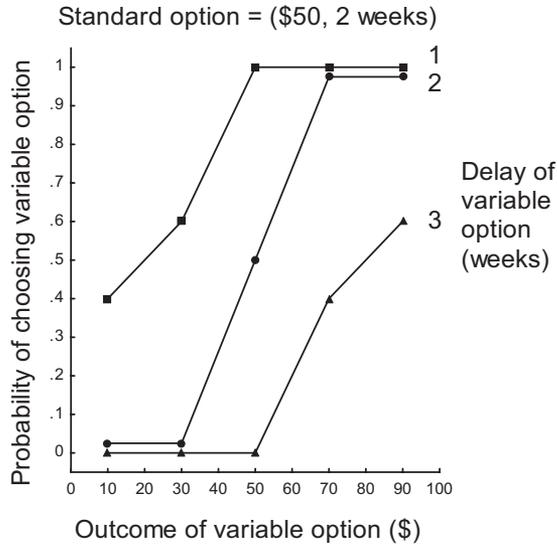


Figure 5. The comparability effect (cf. Mellers & Biagini, 1994, Figure 5, Panel B): The curve relating the outcome of the variable option to the probability of choosing it is steeper when the delays of both options are the same than when they are different.

steeper when the delays of both options are the same than when they are different. Mellers and Biagini (1994) proposed that the similarity or comparability of the options along one attribute enhances the difference between them along the other. This is, in their theory, captured by the contrast weights.

To show how the tradeoff model accommodates the comparability effect, we use the ratio rule to obtain the probability of choosing *LL*, *SS*, and, in case the options are identical, *EE*, where *E* denotes equal outcomes and equal delays, respectively:

$$P_{LL} = \frac{g}{g + f}, P_{SS} = \frac{f}{g + f}, \text{ and } P_{EE} = 1/2.^{16}$$

Similarly, choice probabilities can be obtained in case one option dominates the other. For instance, the probabilities of choosing *EL* and *ES*, that is, a later outcome and a sooner outcome of equal magnitude, are

$$P_{EL} = \frac{0}{0 + f} = 0 \text{ and } P_{ES} = \frac{f}{0 + f} = 1.$$

Using this specification of the choice rule, the comparability effect can be reconstructed as in Figure 6. The probabilities of choosing the variable option over the standard option are either moderate or extreme. Moderate probabilities occur when the options are identical or when they imply a tradeoff, whereas extreme probabilities occur when the variable option dominates, or is dominated by, the standard option. Thus, the tradeoff model explains the comparability effect by correctly distinguishing between dominance and nondominance situations.

In the above reconstruction of the comparability effect, we used a ratio rule to obtain choice probabilities. Although we do not commit ourselves to the ratio rule, we *do* commit ourselves to the other components of the tradeoff model that are needed in practice to obtain choice probabilities: The tradeoff functions in Equations

6.1 and 6.2 (but see the foregoing discussion of continuous tradeoff functions), the value function in Equation 9, and the time-weighting function in Equation 10. Our specification of the value function is a novelty in the formal analysis of choice. We must emphasize, however, that the value function emerged from six qualitative requirements *within the theoretical structure of the tradeoff model*. What is best for the tradeoff model may not be as good for choice models with a different structure. Even so, our value function deserves careful consideration. As mentioned earlier, our value function exhibits decreasing elasticity. Elsewhere (Scholten & Read, 2010), we show that this property can help to explain a well-known anomaly to the expected utility model: A change from risk seeking in small gains to risk aversion in larger gains, and, conversely, a change from risk aversion in small losses to risk seeking in larger losses (Markowitz, 1952). We show that prospect theory can only account for this anomaly if it adopts a decreasingly elastic value function.

One of the six qualitative requirements was that the tradeoff model would accommodate bilinear interaction effects: The present value of a delayed outcome increases with the magnitude of the outcome but increases more steeply the shorter the delay to the outcome (see Figure 4). Bilinear interaction effects are often taken as evidence for multiplicative, alternative-based choice models, but, as our development shows, it can just as well be taken as evidence for an attribute-based choice model in which diminishing sensitivity to attribute amounts is captured by logarithmic functions.

### Concluding Comments

Even though we propose the tradeoff model as an alternative to discounting theory, we have no disagreement with the computation of net present values in finance and economics or with any other normative application of discounting theory. Discounting theory is normatively compelling and, when equipped with a table of interest rates and a spreadsheet, easy to apply. The problems arise, however, with descriptive or even prescriptive applications.

Suppose you, a consultant, want to take a client’s sensitivity to ratios between delays into consideration (the *common difference effect*). This can easily be incorporated into your toolkit: Just replace exponential discounting by hyperbolic discounting. However, suppose you *also* want to take your client’s sensitivity to differences between outcomes into consideration (the *absolute magnitude effect*). Incorporating *this* into your toolkit will, because of the burden it places on the value function of a discounting model (see Appendix B), immediately create a convoluted and clumsy device, which may be impossible to apply.

Our recommendation: “Drop your tools” (Weick, 1996). View your client as someone who is trading off differences between outcomes against differences between delays. While you conduct your tradeoff analysis, remember the Weber–Fechner law, so as to take your client’s sense of proportionality into consideration. You will find that this device is easy to apply and that it is an accurate reflection of your client’s preferences.

<sup>16</sup> The advantages could be exponentiated in order to capture the error in choice behavior.

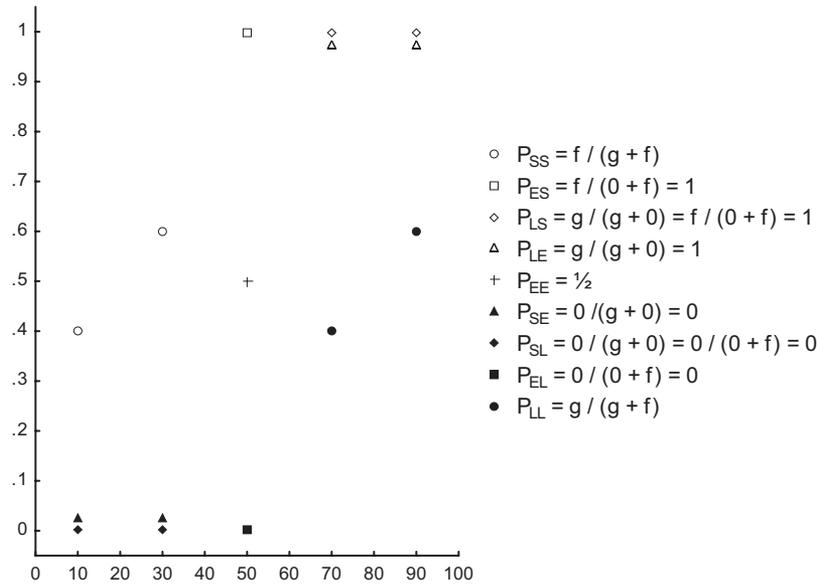


Figure 6. The comparability effect (see Figure 5) according to the tradeoff model. The probabilities of choosing the variable option over the standard option are either moderate or extreme. Moderate probabilities occur when the options are identical ( $P_{EE}$ , where  $E$  denotes equal outcomes and equal delays, respectively) or when they imply a tradeoff ( $P_{SS}$  and  $P_{LL}$ ). Extreme probabilities occur when the variable option dominates the standard option ( $P_{ES}$ ,  $P_{LS}$ , and  $P_{LE}$ ) or when it is dominated by the standard option ( $P_{SE}$ ,  $P_{SL}$ , and  $P_{EL}$ ). Symbols with the same shape but different colors denote complementary probabilities.

References

Abdellaoui, M., Bleichrodt, H., & L'Haridon, O. (2008). A tractable method to measure utility and loss aversion under prospect theory. *Journal of Risk and Uncertainty*, 36, 245–266.

Ahlbrecht, M., & Weber, M. (1997). An empirical study on intertemporal decision making under risk. *Management Science*, 43, 813–826.

Anderson, N. H., & Shanteau, J. C. (1970). Information integration in risky decision making. *Journal of Experimental Psychology*, 84, 441–451.

Arkes, H. R., Hirschleifer, D., Jiang, D., & Lim, S. (2008). Reference point adaptation: Tests in the domain of security trading. *Organizational Behavior and Human Decision Processes*, 105, 67–81.

Baron, J. (2000). Can we use human judgments to determine the discount rate? *Risk Analysis*, 20, 861–868.

Benzion, U., Rapoport, A., & Yagil, J. (1989). Discount rates inferred from decisions: An experimental study. *Management Science*, 35, 270–284.

Busemeyer, J. R., Jessup, R. K., & Dimperio, E. (2007). *Integrating sophisticated choice models with basic learning processes to more fully account for complex choice behavior*. Unpublished manuscript. Available from <http://indiana.academia.edu/JeromeRBusemeyer/Papers/36632/Busemeyer>

Ebert, J. E. J., & Prelec, D. (2007). The fragility of time: Time-insensitivity and valuation of the near and far future. *Management Science*, 53, 1423–1438.

Fennema, H., & van Assen, M. (1999). Measuring the utility of losses by means of the tradeoff method. *Journal of Risk and Uncertainty*, 17, 277–295.

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, 40, 351–401.

González-Vallejo, C. (2002). Making trade-offs: A probabilistic and context-sensitive model of choice behavior. *Psychological Review*, 109, 137–155.

González-Vallejo, C., Bonazzi, A., & Shapiro, A. J. (1996). Effects of

vague probabilities and of vague payoffs on preference: A model comparison analysis. *Journal of Mathematical Psychology*, 40, 130–140.

González-Vallejo, C., & Reid, A. A. (2006). Quantifying persuasion effects on choice with the decision threshold of the stochastic choice model. *Organizational Behavior and Human Decision Processes*, 100, 250–267.

González-Vallejo, C., Reid, A. A., & Schiltz, J. (2003). Context effects: The proportional difference model and the reflection of preference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 942–953.

Grace, R. C. (1999). The matching law and amount-dependent exponential discounting as accounts of self-control choice. *Journal of the Experimental Analysis of Behavior*, 71, 27–44.

Green, L., Fristoe, N., & Myerson, J. (1994). Temporal discounting and preference reversals in choice between delayed outcomes. *Psychonomic Bulletin & Review*, 1, 383–389.

Green, L., Fry, A. F., & Myerson, J. (1994). Discounting of delayed rewards: A life-span comparison. *Psychological Science*, 5, 33–36.

Green, L., Myerson, J., Holt, D. D., Slevin, J. R., & Estle, S. J. (2004). Discounting of delayed food rewards in pigeons and rats: Is there a magnitude effect? *Journal of the Experimental Analysis of Behavior*, 81, 39–50.

Green, L., Myerson, J., & Macaux, E. W. (2005). Temporal discounting when the choice is between two delayed rewards. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 1121–1133.

Heath, C., Larrick, R. P., & Wu, G. (1999). Goals as reference points. *Cognitive Psychology*, 38, 79–109.

Hoch, S., & Loewenstein, G. (1991). Time-inconsistent preferences and consumer self-control. *Journal of Consumer Research*, 17, 492–507.

Holcomb, J. H., & Nelson, P. S. (1992). Another experimental look at individual time preference. *Rationality and Society*, 4, 199–220.

Holt, D. D., Green, L., Myerson, J., & Estle, S. J. (2008). Preference reversals with losses. *Psychonomic Bulletin & Review*, 15, 89–95.

Hsee, C. K., Loewenstein, G. F., Blount, S., & Bazerman, M. H. (1999).

- Preference reversals between joint and separate evaluation of options: A review and theoretical analysis. *Psychological Bulletin*, 125, 576–590.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 363–391.
- Keren, G., & Roelofsma, P. (1995). Immediacy and certainty in intertemporal choice. *Organizational Behavior and Human Decision Processes*, 63, 287–297.
- Killeen, P. R. (2009). An additive-utility model of delay discounting. *Psychological Review*, 116, 602–619.
- Kinari, Y., Ohtake, F., & Tsutsui, Y. (2009). Time discounting: Declining impatience and interval effect. *Journal of Risk and Uncertainty*, 39, 87–112.
- Kirby, K. N., & Herrnstein, R. J. (1995). Preference reversals due to myopic discounting of delayed reward. *Psychological Science*, 6, 83–89.
- Kirby, K. N., & Maraković, N. N. (1995). Modeling myopic decisions: Evidence for hyperbolic delay-discounting within subjects and amounts. *Organizational Behavior and Human Decision Processes*, 64, 22–30.
- Kőszegi, B., & Rabin, M. (2006). A model of reference-dependent preferences. *Quarterly Journal of Economics*, 71, 1133–1165.
- Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, 62, 443–477.
- LeBoeuf, R. A. (2006). Discount rates for time versus dates: The sensitivity of discounting to time-interval description. *Journal of Marketing Research*, 43, 59–72.
- Leland, J. W. (2002). Similarity judgments and anomalies in intertemporal choice. *Economic Inquiry*, 40, 574–581.
- Loewenstein, G. (1988). Frames of mind in intertemporal choice. *Management Science*, 34, 200–214.
- Loewenstein, G., & Elster, J. (1992). *Choice over time*. New York, NY: Russell Sage Foundation.
- Loewenstein, G., & Prelec, D. (1992). Anomalies in intertemporal choice: Evidence and an interpretation. *Quarterly Journal of Economics*, 107, 573–597.
- Loewenstein, G., Read, D., & Baumeister, R. F. (2003). *Time and decision*. New York, NY: Russell Sage Foundation.
- Malkoc, S. A., & Zauberman, G. (2006). Deferring versus expediting consumption: The effect of outcome concreteness on sensitivity to time horizon. *Journal of Marketing Research*, 43, 618–627.
- Markowitz, H. (1952). The utility of wealth. *Journal of Political Economy*, 60, 151–158.
- Mellers, B. A., & Biagini, K. (1994). Similarity and choice. *Psychological Review*, 101, 505–518.
- Murphy, J. G., Vuchinich, R. E., & Simpson, C. A. (2001). Delayed reward and cost discounting. *The Psychological Record*, 51, 571–588.
- O'Donoghue, T., & Rabin, M. (1999). Doing it now or later. *American Economic Review*, 89, 103–124.
- Ong, E. L., & White, K. G. (2004). Amount-dependent temporal discounting? *Behavioural Processes*, 66, 201–212.
- Payne, J. W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance*, 16, 366–387.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 534–552.
- Payne, J. W., Laughhunn, D. J., & Crum, R. (1980). Translation of gambles and aspiration level effects in risky choice behavior. *Management Science*, 26, 1039–1060.
- Payne, J. W., Laughhunn, D. J., & Crum, R. (1981). Further tests of aspiration level effects in risky choice behavior. *Management Science*, 27, 953–958.
- Prelec, D., & Loewenstein, G. (1991). Decision making over time and under uncertainty: A common approach. *Management Science*, 37, 770–786.
- Read, D. (2001). Is time-discounting hyperbolic or subadditive? *Journal of Risk and Uncertainty*, 23, 5–32.
- Read, D., Frederick, S., Orsel, B., & Rahman, J. (2005). Four score and seven years from now: The date/delay effect in temporal discounting. *Management Science*, 51, 1326–1335.
- Read, D., Loewenstein, G., & Kalyanaraman, S. (1999). Mixing virtue and vice: Combining the immediacy effect and the diversification heuristic. *Journal of Behavioral Decision Making*, 12, 257–273.
- Read, D., & Roelofsma, P. H. M. P. (2003). Subadditive versus hyperbolic discounting: A comparison of choice and matching. *Organizational Behavior and Human Decision Processes*, 91, 140–153.
- Richards, J. B., Mitchell, S. H., de Wit, H., & Seiden, L. S. (1997). Determination of discount functions in rats with an adjusting-amount procedure. *Journal of the Experimental Analysis of Behavior*, 67, 353–366.
- Robles, E., & Vargas, P. A. (2007). Functional parameters of delay discounting assessment tasks: Order of presentation. *Behavioural Processes*, 75, 237–241.
- Robles, E., & Vargas, P. A. (2008). Parameters of delay discounting assessment: Number of trials, effort, and sequential effects. *Behavioural Processes*, 78, 285–290.
- Robles, E., Vargas, P. A., & Bejarano, R. (2009). Within-subject differences in degree of delay discounting as a function of order of presentation of hypothetical cash rewards. *Behavioural Processes*, 81, 260–263.
- Roelofsma, P. H. M. P., & Read, D. (2000). Intransitive intertemporal choice. *Journal of Behavioral Decision Making*, 13, 161–177.
- Rubinstein, A. (2003). “Economics and psychology”? The case of hyperbolic discounting. *International Economic Review*, 44, 1207–1216.
- Russo, J. E., & Doshier, B. A. (1983). Strategies for multiattribute binary choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 676–696.
- Samuelson, P. (1937). A note on measurement of utility. *Review of Economic Studies*, 4, 155–161.
- Scholten, M., & Read, D. (2006). Discounting by intervals: A generalized model of intertemporal choice. *Management Science*, 52, 1426–1438.
- Scholten, M., & Read, D. (2010). Back to Markowitz: How and why losses are not the mirror image of gains. Working Paper 1504630. Social Science Research Network (SSRN) eLibrary. Available from [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1504630](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1504630)
- Shafir, E. B., Osherson, D. N., & Smith, E. E. (1993). The advantage model: A comparative theory of evaluation and choice under risk. *Organizational Behavior and Human Decision Processes*, 55, 325–378.
- Shanteau, J. C. (1974). Component processes in risky decision making. *Journal of Experimental Psychology*, 103, 680–691.
- Shelley, M. K. (1993). Outcome signs, question frames, and discount rates. *Management Science*, 39, 806–815.
- Stevenson, M. K. (1986). A discounting model for decisions with delayed positive or negative outcomes. *Journal of Experimental Psychology: General*, 115, 131–154.
- Stevenson, M. K. (1992). The impact of temporal context and risk on the judged value of future outcomes. *Organizational Behavior and Human Decision Processes*, 52, 455–491.
- Stevenson, M. K. (1993). Decision making with long-term consequences: Temporal discounting for single and multiple outcomes in the future. *Journal of Experimental Psychology: General*, 122, 3–22.
- Strahilevitz, M. A., & Loewenstein, G. (1998). The effect of ownership history on the valuation of objects. *Journal of Consumer Research*, 25, 276–289.
- Strotz, R. H. (1955–1956). Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies*, 23, 165–180.
- Takanishi, T. (2005). Loss of self-control in intertemporal choice may be attributable to logarithmic time-perception. *Medical Hypotheses*, 65, 691–693.
- Thaler, R. (1981). Some empirical evidence on dynamic inconsistency. *Economics Letters*, 8, 201–207.

Tversky, A. (1967a). Additivity, utility, and subjective probability. *Journal of Mathematical Psychology*, 4, 175–201.

Tversky, A. (1967b). Utility theory and additivity analysis of risky choices. *Journal of Experimental Psychology*, 75, 27–36.

Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, 76, 31–48.

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *Quarterly Journal of Economics*, 107, 1039–1061.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5, 297–323.

Tversky, A., & Simonson, I. (1993). Context-dependent preferences. *Management Science*, 39, 1179–1189.

Weber, B. J., & Chapman, G. B. (2005). The combined effects of risk and time on choice: Does uncertainty eliminate the immediacy effect? Does delay eliminate the certainty effect? *Organizational Behavior and Human Decision Processes*, 96, 104–118.

Weber, E. U., Johnson, E. J., Milch, K. F., Chang, H., Brodscholl, J. C., & Goldstein, D. G. (2007). Asymmetric discounting in intertemporal choice: A query-theory account. *Psychological Science*, 18, 516–523.

Weick, K. E. (1996). Drop your tools: An allegory for organizational studies. *Administrative Science Quarterly*, 41, 301–313.

Yates, J. F., & Watts, R. A. (1975). Preferences for deferred losses. *Organizational Behavior and Human Decision Processes*, 13, 294–306.

Zauberman, G., Kim, B. K., Malkoc, S. A., & Bettman, J. R. (2009). Discounting time and time discounting: Subjective time perception and intertemporal preferences. *Journal of Marketing Research*, 46, 543–556.

## Appendix A

### Strong Subadditivity

Consider the case in which a person prefers  $(a, t - 1)$  to  $(a + b, t)$  but  $(a + kb, t)$  to  $(a, t - k)$ , where the first element of each option is money, and the second element is time. We will prove that this is incompatible with hyperbolic or exponential discounting and diminishing or constant sensitivity to outcomes: A person who prefers  $(a, t - 1)$  to  $(a + b, t)$  should also prefer  $(a, t - k)$  and  $(a + kb, t)$ . We owe this proof to Ana Rita Pires.

Given the following conditions:

$$a, b > 0, \tag{A1}$$

$$v(0) = 0, v(x) > 0 \text{ for } x > 0, v'(x) > 0, \text{ and } v''(x) \leq 0, \tag{A2}$$

$$0 < \delta_{t-k} \leq \delta_{t-k+1} \leq \dots \leq \delta_{t-1} \leq \delta_t < 1, \text{ where } k < t, \text{ and} \tag{A3}$$

$$\prod_{i=1}^t \delta_i v(a + b) - \prod_{i=1}^{t-1} \delta_i v(a) < 0, \text{ or} \tag{A4}$$

$$\delta_t < \frac{v(a)}{v(a + b)},$$

we want to prove that

$$\prod_{i=1}^t \delta_i v(a + kb) - \prod_{i=1}^{t-k} \delta_i v(a) < 0, \text{ or}$$

$$\prod_{i=t-k+1}^t \delta_i < \frac{v(a)}{v(a + kb)}.$$

Given Condition A3, it is sufficient to prove that

$$\delta_t^k < \frac{v(a)}{v(a + kb)}.$$

Let  $V(x) = \log v(x)$ . Condition A2 implies

$$V'(x) = \frac{v'(x)}{v(x)} > 0, \text{ and}$$

$$V''(x) = \frac{v''(x)v(x) - v'(x)^2}{v(x)^2} < 0. \tag{A5}$$

Because  $v(x) = e^{V(x)}$ , Condition A4 becomes

$$e^{V(a) - V(a+b)} > \delta_t, \text{ or}$$

$$V(a) - V(a + b) > \log(\delta_t),$$

and what we want to prove becomes

$$e^{V(a) - V(a+kb)} > \delta_t^k, \text{ or}$$

$$V(a) - V(a + kb) > k \log(\delta_t), \text{ or}$$

$$\frac{V(a) - V(a + kb)}{k} > \log(\delta_t).$$

Given Condition A5, we have

$$\frac{V(a) - V(a + kb)}{k} > V(a) - V(a + b),$$

which completes the proof.

(Appendices continue)

## Appendix B

### Augmenting Proportional Sensitivity Versus Increasing Elasticity

We will prove that, when combined with diminishing absolute sensitivity, augmenting proportional sensitivity is satisfied by a broader class of value functions than increasing sensitivity. We give the proof for the case of gains; a parallel proof for the case of losses is straightforward.

In the tradeoff model, the absolute magnitude effect occurs because

$$v(mbx) - v(mx) > v(bx) - v(x),$$

where  $m, b > 1$  and  $x > 0$ . Thus,  $H(x) = v(bx) - v(x)$  increases in  $x$ . That is,  $H'(x) > 0$ ,

$$v'(bx)b - v'(x) > 0,$$

$$bxv'(bx) > xv'(x).$$

Thus,  $h(x) = xv'(x)$  increases in  $x$  (augmenting proportional sensitivity), meaning that  $h'(x) > 0$ ,

$$v'(x) + xv''(x) > 0,$$

$$v''(x) > -\frac{v'(x)}{x}. \quad (\text{B1})$$

In a discounting model, the absolute magnitude effect occurs because

$$\frac{v(mbx)}{v(mx)} > \frac{v(bx)}{v(x)}.$$

Thus,  $H(x) = v(bx)/v(x)$  increases in  $x$ . That is,  $H'(x) > 0$ ,

$$\frac{v'(bx)bv(x) - v(bx)v'(x)}{v(x)^2} > 0,$$

$$v'(bx)bv(x) - v(bx)v'(x) > 0,$$

$$bx\frac{v'(bx)}{v(bx)} > x\frac{v'(x)}{v(x)}.$$

Thus,  $h(x) = xv'(x)/v(x)$  increases in  $x$  (increasing elasticity), meaning that  $h'(x) > 0$ ,

$$\frac{[xv'(x)]'v(x) - xv'(x)v'(x)}{v(x)^2} > 0,$$

$$[v'(x) + xv''(x)]v(x) - xv'(x)^2 > 0,$$

$$v''(x) > \left[ \frac{v'(x)}{x} - 1 \right] \frac{v'(x)}{x}. \quad (\text{B2})$$

Because  $x, v(x), v'(x) > 0$  but  $v''(x) < 0$  (diminishing absolute sensitivity), Inequality B1 is satisfied by a broader class of value functions than Inequality B2, which completes the proof.

Received January 24, 2007

Revision received November 10, 2009

Accepted March 4, 2010 ■