

Original citation:

Scholten, Marc, Read, Daniel and Sanborn, Adam N.. (2016) Cumulative weighing of time in intertemporal tradeoffs. *Journal of Experimental Psychology: General*, 145 (9). pp. 1177-1205.

Permanent WRAP URL:

<http://wrap.warwick.ac.uk/79211>

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions. Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:

© APA 2016

Link to the authoritative document: <http://dx.doi.org/10.1037/xge0000198>

This article may not exactly replicate the authoritative document published in the APA journal. It is not the copy of record.

A note on versions:

The version presented here may differ from the published version or, version of record, if you wish to cite this item you are advised to consult the publisher's version. Please see the 'permanent WRAP URL' above for details on accessing the published version and note that access may require a subscription.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

RUNNING HEAD: Cumulative Weighing of Time in Intertemporal Tradeoffs

Cumulative Weighing of Time in Intertemporal Tradeoffs

Marc Scholten

Universidade Europeia, Lisbon, Portugal

Email: marc.scholten@universidadeeuropeia.pt

Daniel Read

Warwick Business School, Coventry, United Kingdom

Email: daniel.read@wbs.ac.uk

Adam Sanborn

University of Warwick, Coventry, United Kingdom

Email: a.n.sanborn@warwick.ac.uk

WORD COUNT (ABSTRACT AND BODY): 18,540

Abstract

We examine preferences for sequences of delayed monetary gains. In the experimental literature, two prominent models have been advanced as psychological descriptions of preferences for sequences. In one model, the instantaneous utilities of the outcomes in a sequence are discounted as a function of their delays, and assembled into a discounted utility of the sequence. In the other model, the accumulated utility of the outcomes in a sequence is considered along with utility or disutility from improvement in outcome utilities and utility or disutility from the spreading of outcome utilities. Drawing on three threads of evidence concerning preferences for sequences of monetary gains, we propose that the accumulated utility of the outcomes in a sequence is traded off against the *duration* of utility accumulation. In our first experiment, aggregate choice behavior provides qualitative support for the tradeoff model. In three subsequent experiments, one of which incentivized, disaggregate choice behavior provides quantitative support for the tradeoff model in Bayesian model contests. The third experiment addresses one thread of evidence that motivated the tradeoff model: When, in the choice between two single dated outcomes, it is conveyed that receiving less sooner means receiving nothing later, preference for receiving more later increases, but when it is conveyed that receiving more later means receiving nothing sooner, preference is left unchanged. Our results show that this *asymmetric hidden-zero effect* is indeed driven by those supporting the tradeoff model. The tradeoff model also accommodates all remaining evidence on preferences for sequences of monetary gains.

Keywords

Intertemporal choice; sequences; discounting; utility; tradeoff.

Cumulative Weighing of Time in Intertemporal Tradeoffs

Intertemporal choices are those in which the outcomes of available options are distributed over time, and tradeoffs must be made between *what* will be experienced and *when*. Examples of choices with intertemporal features are those between professional careers, medical treatments, and home mortgages. In these situations, the options of choice are well characterized as streams, or *sequences*, of outcomes.

Most laboratory studies on the psychology of intertemporal choices restrict their scope to choices between two *single* dated outcomes, as that between \$100 in one year and \$150 in two years. Here, the outcomes are “distributed over time,” but only *between* the available options. Some studies broaden the scope to choices between *sequences* of dated outcomes, as that between a rising and a falling profile of monetary gains, e.g., {\$100 in one year, \$150 in two years} and {\$150 in one year, \$100 in two years}. Here, the outcomes are also “distributed over time,” but now *within*, as well as between, options. This comes closer to the real-life examples given above. The purpose of this paper is to investigate preferences for sequences of monetary gains. In doing so, we introduce a substantive innovation in the psychological modeling of individuals’ choices between sequences, and present both qualitative and quantitative support for this new approach.

We consider three candidate models for describing preferences for sequences: Two prominent models from the experimental literature, and a third model that we present as a challenger. In what we call the *discounted instantaneous utility model* (Abdellaoui, Attema, & Bleichrodt, 2010), instantaneous utilities are assigned to the dated outcomes in a sequence, each utility is discounted as a function of the delay to the outcome, the discounted utilities of the outcomes are assembled into the discounted utility of the sequence, and the sequence with the highest discounted utility is chosen. In what we call the *sequences model* (Loewenstein & Prelec, 1993), the accumulated utility of the outcomes in a sequence is considered along with utility or

disutility from improvement in outcome utilities and utility or disutility from the spreading of outcome utilities, and the sequence with the highest utility is chosen. As a challenger, we propose a modification of the *tradeoff model* (Read & Scholten, 2012; Scholten & Read, 2010), in which the accumulated utility of the outcomes in a sequence is traded off against the *duration* of utility accumulation, and the sequence favored by this tradeoff is chosen. Qualitative and quantitative evidence from four experiments provides credence to the challenger.

Our model development is motivated by three threads of evidence concerning preferences for sequences of *unlabeled* monetary gains, meaning money that is not attributed to a specific origin, such as wage payments or rental income (e.g., Loewenstein & Sicherman, 1991). That evidence shows deviations from the *Net Present Value* (NPV) model, the standard of economic rationality found in textbooks on finance and economics (e.g., Brealey et al., 2012; Hey, 2003), and due originally to Fisher (1930). In the NPV model, each outcome in a sequence is discounted as a function of the delay to the outcome, at an interest rate determined by the opportunity cost of money, and the discounted outcomes are assembled into the NPV of the sequence. Psychologists often cite a variant of the NPV model, referred to as the exponential-discounting model, as the normative standard. In their variant, r is any *personal* interest rate that may differ from the market interest rate or opportunity cost of money (Read, Frederick, & Scholten, 2013; Scholten & Read, 2013). To illustrate the NPV model, consider the following choice pair, in which outcomes being distributed over three consecutive years:

Option	Delay to outcome			Exponentially discounted outcome		
	1	2	3	$(x_1, 1)$	$(x_2, 2)$	$(x_3, 3)$
Low NPV		\$1,000			$\left(\frac{1}{1+r}\right)^2 1,000$	
High NPV	\$500		\$500	$\left(\frac{1}{1+r}\right) 500$		$\left(\frac{1}{1+r}\right)^3 500$

The term between brackets is the *per-period discount factor*, and, within it, r is the interest rate. By *any* nonzero interest rate, i.e., either positive ($r > 0$) or, which is nonstandard in economics, negative ($-1 < r < 0$), the “high NPV option” has a higher Net Present Value than the “low NPV option,” i.e., $500 / (1 + r) + 500 / (1 + r)^3 > 1,000 / (1 + r)^2$, which solves for $r^2 > 0$. In case of a zero interest rate, the NPV of each sequence is simply the sum of its outcomes, which, in this illustration, and in most research on preferences for sequences, is the same for both sequences. The three threads of evidence are as follows.

Thread 1: Undecided preference between decreasing and constant sequences. In incentivized studies, Manzini, Mariotti, and Mittone (2010) and Gigliotti and Sopher (1997) examined preferences between decreasing, constant, and increasing sequences. Sequences totaled the same amount of money, e.g., $\{\$32, \$16\}$ for a decreasing sequence, $\{\$24, \$24\}$ for a constant sequence, and $\{\$16, \$32\}$ for an increasing sequence. Majorities preferred decreasing and constant sequences to increasing ones (80% and 93%, respectively, in Manzini et al., 2010; 62% and 82%, respectively, in Gigliotti & Sopher, 1997). This is consistent with the NPV model, because a given amount of money sooner is always better than the same amount later. By the same logic, however, decreasing sequences are better than constant ones, and, yet, no consensus emerged, with only about half of the participants preferring decreasing sequences over constant

ones (66% in Manzini et al., 2010, and 40% in Gigliotti & Sopher, 1997). Our experiments confirm that preference between decreasing and constant sequences is basically undecided (i.e., close to chance level).

Thread 2: The asymmetric hidden-zero effect. The *hidden-zero effect* (Magen, Dweck, & Gross, 2008) is that people are less patient when choosing between single dated outcomes, e.g., “\$100 today” (*SS*, for Smaller-Sooner) and “\$150 in one year” (*LL*, for Larger-Later), than when choosing between the outcome sequences “\$100 today and \$0 in one year” (*SS*₀) and “\$0 today and \$150 in one year” (*LL*₀). That is, choosing *LL* rather than *SS* is less likely than choosing *LL*₀ rather than *SS*₀. This is incompatible with the NPV model, because zero outcomes have zero value, and should therefore not affect preference. Subsequent research (Read, Olivola, & Hardisty, in press; Wu & He, 2012) has shown, however, that the zeroes are not equal: Adding a later zero (changing *SS* into *SS*₀) increases patience, but adding a sooner zero (changing *LL* into *LL*₀) has no effect. We will show that this *asymmetric hidden-zero effect* is driven by those individuals whose preferences are best described by the tradeoff model.

Thread 3: Preference for faster accumulation. The third thread of evidence originally came from informal observations, which we then confirmed experimentally. It concerns the preference between the low and high NPV options discussed earlier. Given *any* time preference, the NPV model predicts a preference for the high NPV option over the low NPV option. We conducted informal surveys of colleagues and students, and, without any exception, large majorities preferred the *low* NPV option, arguing that, with the low NPV option, they would *receive the full amount sooner*. That is, they would receive the full \$1,000 by the second period, whereas, with the high NPV option, they would have to wait until the third period to get that full amount. Confirming these informal observations, 71% chose the low NPV option in an MTurk survey of 429 American residents, and 79% did so in a survey of 321 British residents.¹ We will

show that, by holding on to the logic of our colleagues and students, which is that the low NPV option *accumulates faster* to \$1,000 than the high NPV option, we arrive at a model of intertemporal choice that offers the best account of preferences for monetary sequences.

Models of Preferences for Sequences

Below, we describe the three candidate models of preferences for sequences, and discuss how each deals with the three threads of evidence discussed in the introduction. It turns out that the discounted instantaneous utility model and the sequences model fail to offer a coherent account of the evidence, whereas the tradeoff model does, and in a parsimonious way. This will set the stage for four experiments, in which we examine whether the tradeoff model outperforms the other candidate models in accounting for fresh evidence.

The discounted instantaneous utility model

We earlier described the NPV model, in which *outcomes* are discounted at a constant rate as a function of their delays. An alternative to the NPV model is the discounted instantaneous utility model, in which utilities are assigned to the outcomes (by means of a utility function that may not be linear), and the utilities are discounted at a personal rate (which may deviate from the rate deriving from the economically appropriate interest rate). Formally, the discounted utility of a sequence is given as:

$$U(x_1, t_1; \dots; x_n, t_n) = \sum_{i=1}^n \delta^{t_i} u(x_i), \quad (1)$$

where x_i is the i th outcome in the sequence, t_i is its delay, n is the number of outcomes in the sequence, and u is an instantaneous utility function satisfying $u(0) = 0$ (Abdellaoui et al., 2010, Wakker, 2010). Even though u might be any monotonically increasing function, it is commonly assumed to be concave over gains, meaning that it exhibits *diminishing sensitivity* (Kahneman & Tversky, 1979; Loewenstein & Prelec, 1992). Diminishing sensitivity is that the marginal impact

of an outcome decreases with its absolute magnitude; e.g., adding \$1 to \$100 has psychologically less impact than adding \$1 to \$10.

Table 1 summarizes how the discounted instantaneous utility model (and the models yet to be discussed) deals with the three threads of evidence. We begin with Thread 1, or the undecided preference between decreasing and constant sequences. Consider the decreasing sequence {\$400, \$200} and the constant sequence {\$300, \$300}. By constant discounting of outcomes, or computation of Net Present Value, we have $\delta 400 + \delta^2 200 > \delta 300 + \delta^2 300$, or $100 > \delta 100$, so that the decreasing sequence would be preferred to the constant one. The discounted instantaneous utility model, however, posits discounting of outcome *utilities*. Concave utility countervails discounting, in that $u(400) + u(200) < 2u(300)$, which favors the constant sequence. When concave utility *outweighs* discounting, we have $\delta u(400) + \delta^2 u(200) < \delta u(300) + \delta^2 u(300)$, or $u(400) - u(300) < \delta [u(300) - u(200)]$, so that the constant sequence would be preferred to the decreasing one. With discounting and concave utility working in opposite directions, there may be no consensus on the preference between decreasing and constant sequences, which is Thread 1. The discounted instantaneous utility model further predicts that the increasing sequence {\$200, \$400} will lose from the decreasing sequence {\$400, \$200}, by discounting, and from the constant sequence {\$300, \$300}, by both discounting and concave utility, consistent with the results from Manzini et al. (2010) and Gigliotti and Sopher (1997).

<Insert Table 1 about here>

Importantly, any preference for a constant sequence over a decreasing one is a violation of *dominance*, a cornerstone in the rational choice paradigm (Tversky & Kahneman, 1986), which, in the domain of intertemporal choice, affirms that, when the cumulative outcome of sequence X in the successive time periods is never worse than that of sequences Y , and better in at least one time

period, then X dominates Y , and *should* be chosen (Scholten & Read, 2014). This definition of dominance assumes that more money is better than less (*monotonicity*), and that money is better sooner than later (*positive time preference*; Fisher, 1930). Thus, holding the total amount earned constant, a decreasing sequence dominates a constant one, and a constant sequence dominates an increasing one. By dominance, people should indeed prefer decreasing and constant sequences to increasing ones, as observed, but they should *also* prefer decreasing sequences to constant ones, while no consensus emerged. The discounted instantaneous utility model offers an explanation for this, by suggesting that concave utility countervails discounting, and is therefore a *nonnormative* model.

The discounted instantaneous utility model can further accommodate the preference for faster accumulation (Thread 3) when assuming that discounting is outweighed by *convex* utility, or *augmenting sensitivity*, in which case we have $\delta u(500) + \delta^3 u(500) < \delta^2 u(1,000)$, or $\delta^{-1} u(500) + \delta u(500) < u(1,000)$. Concave utility over gains is a nonstandard assumption, and, in any case, the discounted instantaneous utility model needs *concave* utility to accommodate the undecided preference between decreasing and constant sequences (Thread 1). It thus cannot offer a coherent account of the prior evidence.

The discounted instantaneous utility model cannot account for the (asymmetric) hidden-zero effect (Thread 2) either, because zero outcomes have zero utility, and should therefore not affect preference.

The sequences model

Loewenstein and Prelec (1993) developed the sequences model. “The guiding idea behind the model,” according to the authors, “is that evaluation of sequences reflects the interaction between two motives: a basic preference for improvement tempered by a desire to spread the better outcomes more or less uniformly over the entire interval” (p. 92). In proposing a preference

for improvement, they were partly motivated by Loewenstein and Sicherman's (1991) finding that people preferred increasing over decreasing sequences of monetary gains. As mentioned earlier, Manzini et al. (2010) and Gigliotti and Sopher (1997) found a strong preference for decreasing over increasing sequences. While Loewenstein and Sicherman (1991) used large amounts of money, labeled the money, i.e., specified its origin (wage payments or rental income), and presented the sequences using column charts to museum visitors, whereas Manzini et al. (2010) and Gigliotti and Sopher (1997) used small amounts of money, did not label the money, and presented sequences using delay-by-payoff matrices to college students. It is not entirely clear which methodological differences are responsible for the opposite results, but preference for increasing sequences is more pronounced for larger amounts of money (Duffy & Smith, 2013), more pronounced for income from labor than for income from rent (Loewenstein & Sicherman, 1991) and income from lotteries (Duffy & Smith, 2013; Duffy, Smith, & Woods, 2015), perhaps because people *expect* wages to increase over time (Chapman, 1996), and more pronounced among naïve people than among (financially) sophisticated people (Guyse, Keller, & Eppel, 2002; Loewenstein & Sicherman, 1991; Matsumoto, Peecher, & Rich, 2000). In any case, the sequences model is very general, in that it allows individuals to either like or dislike improvement, and either like or dislike spreading, so we will give it maximum latitude.

As applied to sequences of future outcomes distributed over contiguous time periods (e.g., three years, as in the choice between the high and low NPV options), the sequences model has three components: One is the sum of instantaneous utilities, and the others are the “gestalt” properties of sequences” (Loewenstein & Prelec, 1993, p. 94): Improvement and spreading.

The sequences model draws on the concept of *cumulative* instantaneous utilities. Imagine that, in each period, some instantaneous utility is received, and a running total of that utility is kept. This is the cumulative instantaneous utility. This “real” running total is compared with an

“imaginary” running total from a constant utility stream across the periods. More precisely, in each period, the utility that has accumulated until that period is compared with the utility that would have accumulated until that period had the same sum of utility been distributed uniformly over time:

$$d_i = \frac{i}{n} \sum_{j=1}^n u(x_j) - \sum_{j=1}^i u(x_j). \quad (2a)$$

where the second term on the right is the utility that has accumulated until each period i , and the first term on the right is the utility that would have accumulated until that period had the same sum of utility been distributed uniformly over time. To see how this works, assume linear utility, i.e., $u(x) = x$, and consider the sequence $S = \{\$100, \$200, \$300\}$, the cumulative outcomes of which are $C_S = \{\$100, \$300, \$600\}$. Had the outcomes been distributed evenly over time, we would have had the (flat) sequence $F = \{\$200, \$200, \$200\}$, the cumulative outcomes of which are $C_F = \{\$200, \$400, \$600\}$. We see that C_F “runs ahead” of C_S by $\$200 - \$100 = \$100$ in period 1, and by $\$400 - \$300 = \$100$ in period 2. The *improvement score* of the sequence S is the sum of these differences, i.e., $\sum d_i$, or $\$100 + \$100 = \$200$, which is a positive number, and therefore an indication of improvement rather than *decline* (Loewenstein & Prelec, 1993). It is no coincidence that the improvement score equals the difference between the outcome afforded by S in period 3 and the outcome afforded in period 1, i.e., $\$300 - \$100 = \$200$, because, for any three-outcome sequence $\{x_1, x_2, x_3\}$, the improvement score is $u(x_3) - u(x_1)$, which is the difference between the utility of the last outcome and the utility of the first outcome in the sequence. For an increasing sequence, the improvement score is positive; for a constant sequence, it is zero; and, for a decreasing sequence, it is negative. To capture an individual’s attitude toward improvement, the improvement score is weighted by β , where $\beta > 0$ is a desire for improvement, $\beta = 0$ is neutrality toward improvement, and $\beta < 0$ is distaste for improvement.²

The *spreading score* of a sequence is the sum of the *absolute* differences, i.e., $\sum |d_i|$, which indicates how much the sequence deviates from a uniform distribution. For the sequence S , all differences are positive, so that its spreading score equals the improvement score. More generally, for a three-outcome sequence $\{x_1, x_2, x_3\}$, the spreading score is $\frac{1}{3}(|u(x_3) + u(x_2) - 2u(x_1)| + |2u(x_3) - u(x_2) - u(x_1)|)$, which contrasts the utilities in periods 3 and 2 with the utility in period 1, and contrasts the utility in period 3 with the utilities in periods 2 and 1. For a constant sequence, the spreading score is zero; for an increasing or a decreasing sequence, it is positive. To capture an individual's attitude toward spreading, the spreading score is weighted by $-\sigma$, where $\sigma > 0$ is a desire for spreading (meaning that the utility of the sequence *decreases* more with a larger deviation from a uniform distribution), $\sigma = 0$ is neutrality toward spreading, and $\sigma < 0$ is distaste for spreading (meaning that the utility of the sequence *increases* more with a larger deviation from a uniform distribution).³ Formally, the utility of a sequence is given as:

$$U(x_1, t_1; \dots; x_n, t_n) = \sum_{i=1}^n u(x_i) + \beta \sum_{i=1}^n d_i - \sigma \sum_{i=1}^n |d_i|. \quad (2b)$$

Loewenstein and Prelec (1993) applied the model in Equations (2a) and (2b) to decisions about consumption experiences, e.g., weekends differing in their pleasantness, and observed a preference for improvement tempered by a desire for spreading. As mentioned, however, the model is very general, and allows for any attitude toward improvement and spreading. Moreover, its flexibility can be enhanced even further by alternative treatments of concealed zero outcomes.

In one variant of the sequences model, which we call the *no-zero sequences model*, concealed zeroes are *not* inferred. The no-zero sequences model can accommodate the asymmetric hidden-zero effect (Thread 2) if it invokes a desire for improvement and a desire for spreading, consistent with Loewenstein and Prelec's (1993) findings. Either zero creates a sequence that deviates from a uniform distribution, which, under a desire for spreading, is evaluated negatively.

The later zero (changing SS into SS_0) creates a decreasing sequence, which, under a desire for improvement, is evaluated negatively as well. With the desire for improvement reinforcing the desire for spreading, the later zero will increase preference for LL . In contrast, the sooner zero (changing LL into LL_0) creates an increasing sequence, which, under a desire for improvement, is evaluated positively. With the desire for improvement countervailing the desire for spreading, the sooner zero may not affect preference between SS and LL_0 , and it will not do so if the desire for improvement precisely offsets the desire for spreading. Thus, to accommodate the asymmetric hidden-zero effect, the no-zero sequences model can preserve a desire for improvement and a desire for spreading. However, to accommodate preference for decreasing over increasing sequences, which accompanies the undecided preference between decreasing and constant sequences (Thread 1) and preference for constant over increasing sequences, the non-zero sequences model needs *distaste* for improvement. It therefore fails to offer a coherent account of the prior evidence.

In the other variant of the sequences model, which we call the *zero sequences model*, concealed zeroes *are* inferred. For this variant, the asymmetric hidden-zero effect (Thread 2) is insurmountable. When the later zero is revealed by the experimenter (changing SS into SS_0), the sooner zero is inferred by the decision maker (changing LL into LL_0), and, when the sooner zero is revealed by the experimenter (changing LL into LL_0), the later zero is inferred by the decision maker (changing SS into SS_0). Because either zero changes the choice into one between SS_0 and LL_0 , it cannot be that one zero has an effect, whereas the other has not. Thus, the zero sequences model also fails to accommodate the prior evidence. Table 1 provides additional details on how the sequences model can enforce its flexibility to address the isolated pieces of prior evidence.

The tradeoff model

Whereas, in the sequences model, the accumulated utility of the outcomes is combined with the utility deriving from a desire or distaste for improvement and the utility deriving from a desire or distaste for spreading, we propose an instantiation of intertemporal the tradeoff model in which the accumulated utility is traded off against the *duration* of utility accumulation. Generally, the duration of a sequence is some average of the constituent delays:

$$\hat{t} = \sum_{j=1}^n a_j t_j, \text{ where } \sum_{j=1}^n a_j = 1.$$

In the instantiation of the intertemporal tradeoff model proposed by Read and Scholten (2012), duration lies between two limits. One is the unweighted average of constituent delays, $a_j = 1 / n$, and the other is the weighted average of constituent delays, with each delay being weighted by the relative magnitude of the outcome received following that delay:

$$a_j = \frac{x_j}{\sum_{i=1}^n x_i}.$$

Whatever the computation, the duration of a sequence lies between the delay to the first outcome (t_1) and the delay to the last outcome (t_n), because money is received after t_1 , meaning that the duration will be longer than t_1 , and money is received before t_n , meaning that the duration will be shorter than t_n : The duration thus lies “somewhere in between.” In light of Thread 3 (preference for faster accumulation), we propose to preserve the concept of duration, but, in its computation, to weigh individual delays by *cumulative outcome utilities*, meaning that a_j is given as:

$$a_j = \frac{\sum_{k=1}^j u(x_k)}{\sum_{i=1}^n \left(\sum_{k=1}^i u(x_k) \right)}.$$

When a_j is specified in this way, \hat{t} becomes the *duration of utility accumulation*, obtained through a *cumulative weighing of time*, with each t_j being weighted by the utility accumulating until period j , normalized by the utility accumulating until all time periods:

$$\hat{t} = \sum_{j=1}^n \frac{\sum_{k=1}^j u(x_k)}{\sum_{i=1}^n \left(\sum_{k=1}^i u(x_k) \right)} t_j. \quad (3a)$$

Consider the high NPV option { \$500 in one year, \$500 in three years }, with which the utility accumulating to period 1 is $u(500)$, the utility accumulating until period 3 is $u(500) + u(500)$, and the utility accumulating until both time periods is $u(500) + (u(500) + u(500))$. Thus, the duration of the high NPV option is:

$$\hat{t}_H = \frac{u(500)}{u(500) + (u(500) + u(500))} \times 1 + \frac{u(500) + u(500)}{u(500) + (u(500) + u(500))} \times 3 = 2.33.$$

In contrast, the low NPV option (\$1,000 in two years) is a single dated outcome, which has no duration, and its delay is $t_L = 2$. Because the duration of the high NPV option is longer than the delay of the low NPV option, i.e., $\hat{t}_H = 2.33 > 2 = t_L$, the cumulative weighing of time contributes to a preference for the low NPV option.

To complete the analysis, we must specify the tradeoff model in its full form. The tradeoff model is an *attribute-based* model, in that outcome and time advantages of the options are traded off against one another, and the option favored by the tradeoff is chosen. In contrast, the previous models are *alternative-based* models, in that each option receives a utility, independently of the other option, and the option with the highest utility is chosen. According to the tradeoff model, the decision maker will be indifferent between the high NPV option (H) and the low NPV option (L) when:

$$\sum_{i=1}^n u(x_{H_i}) - \sum_{i=1}^n u(x_{L_i}) = \kappa(w(\hat{t}_H) - w(\hat{t}_L)), \quad (3b)$$

where, analogous to $\delta < 1$ in discounting models, $\kappa > 0$ captures impatience, with higher values benefiting the sequence with the shortest duration, and w is a (concave) time-weighing function (thus exhibiting diminishing sensitivity).⁵ Furthermore, the difference between the accumulated utilities is an advantage of H when positive and an advantage of L when negative, whereas the difference between weighted durations is an advantage of L when positive and an advantage of H when negative. In the above choice between the high NPV option and the low NPV option, the low NPV option is a single dated outcome, so that $\hat{t}_L = t_L$, and this delay is shorter than the duration of the high NPV option, i.e., $t_L < \hat{t}_H$, which contributes to preference for the low NPV option. By diminishing sensitivity to outcomes (concavity of the utility function u), however, the accumulated utility of the high NPV option will be higher than the utility of the low NPV option, i.e., $u(500) + u(500) > u(1,000)$, contributing to a preference for the high NPV option. Preference for the low NPV option (Thread 3) occurs when concave utility is outweighed by the cumulative weighing of time.

The tradeoff model in Equations (3a) and (3b) also accounts for the asymmetric hidden-zero effect (Thread 2). The smaller-sooner outcome (SS) is “\$100 today.” This option has no duration, and its delay is $t_{SS} = 0$. Revealing the later zero outcome, i.e., “\$100 today and \$0 in one year” (SS_0), creates a sequence, which *has* duration. Specifically, the duration of SS_0 is:

$$\hat{t}_{SS_0} = \frac{u(100)}{u(100) + u(100) + u(0)} \times 0 + \frac{u(100) + u(0)}{u(100) + u(100) + u(0)} \times 1 = \frac{1}{2}.$$

Because $\hat{t}_{SS_0} = 1/2 > 0 = t_{SS}$, revealing the later zero increases preference for LL . Conversely, the larger-later outcome (LL) is “\$150 in one year,” the delay of which is t_{LL} . Revealing the sooner

zero outcome, i.e., “\$0 today and \$150 in one year” (LL_0), creates a sequence, the duration of which is:

$$\hat{t}_{LL_0} = \frac{u(0)}{u(0) + u(0) + u(150)} \times 0 + \frac{u(0) + u(150)}{u(0) + u(0) + u(150)} \times 1 = 1.$$

Because $\hat{t}_{LL_0} = 1 = t_{LL}$, revealing the sooner zero does not affect preference between SS and LL .

Therefore, a cumulative weighing of time naturally produces the asymmetric hidden-zero effect.

Finally, the tradeoff model in Equations (3a) and (3b) agrees with the one proposed by Read and Scholten (2012) that duration is longer for increasing than for constant sequences, and longer for constant than for decreasing sequences. Increasing sequences lose from decreasing sequences because of their longer duration, and lose from constant sequences not only because of their longer duration, but also because of concave utility. The undecided preference between decreasing and constant sequences (Thread 1) occurs because duration (favoring the decreasing sequence) is countervailed by concave utility (favoring the constant sequence). Overall, and as summarized in Table 1, a tradeoff model in which the accumulated utility of a sequence is traded off against the duration of utility accumulation offers a coherent and parsimonious account of all prior evidence.

As mentioned in Footnote 5, Equation (3b) may be rearranged into an (arithmetically) “discounted utility” model, by which $\Sigma u_H - \kappa \hat{t}_H = \Sigma u_L - \kappa \hat{t}_L$. The current specification of the tradeoff model thus seems to blur the distinction between attribute-based choice and alternative-based choice. Our application of the tradeoff model would seem to do so as well, because choice will be assumed to follow a stochastic difference rule, so that the probability of choosing the high NPV option is $F[(\Sigma u_H - \Sigma u_L) - \kappa(\hat{t}_H - \hat{t}_L)]$, which is equivalent to $F[(\Sigma u_H - \kappa \hat{t}_H) - (\Sigma u_L - \kappa \hat{t}_L)]$. However, if the tradeoff model were an alternative-based choice model, a sequence of *positive*

outcomes might receive a *negative* “discounted utility,” which is formally unacceptable and psychologically implausible. This constitutes a compelling argument for maintaining that the tradeoff model *is* an attribute-based choice model.

Candidate models

We identified three candidate models for describing preferences for sequences: The Discounted Instantaneous Utility Model (DIUM), in Equation (1), the Sequences Model (SM), in Equations (2a) and (2b), as previously applied by Loewenstein and Prelec (1993) and Guyse et al. (2002), and the Tradeoff Model (TM), in Equations (3a) and (3b). Although Loewenstein and Prelec (1993) applied Equations (2a) and (2b) in their empirical test of SM, they also showed how discounting of individual outcome utilities could be incorporated into their model. As evident from their simulations (Loewenstein and Prelec, 1993, Figure 4), discounting interacts with the “gestalt” properties of sequence when a sequence starts in (or close to) the present (e.g., 0, 1, and 2) or when time periods are noncontiguous (e.g., 1, 2, and 10). Similar to Loewenstein and Prelec (1993), who used contiguous one-week intervals starting in one week, we use contiguous one-year intervals starting in one year (i.e., 1, 2, and 3). And, as Loewenstein and Prelec (1993) did, we compare the performance of SM with that of DIUM. The novelty is that we introduce TM, with its cumulative weighing of time, as a candidate model, and that we apply the candidate models to sequences of unlabeled monetary gains rather than consumption experiences.

In Experiment 1, we compare the *qualitative* predictions of the candidate models in an aggregate analysis, and, in Experiments 2 to 4, we compare their *quantitative* predictions in a Bayesian model contest in disaggregate analyses. We find that TM best describes both the “representative agent” and the majority of agents individually. In Experiment 3, we validate the Bayesian model contest against the asymmetric hidden-zero effect, which we show to be indeed

driven by those participants best described by TM. We conclude with an assessment of how TM deals with remaining evidence on preferences for sequences of unlabeled monetary gains.

Experiment 1:

Concealed Zeroes Revealed

In Experiment 1, we examine, within the “hidden-zero paradigm,” how the preference between low and high NPV sequences is affected when the zero outcomes are revealed. We introduce two changes to elicitation of preference between the low and high NPV options. First, we use multiples of £100 below the £1,000 mark. Indeed, a partial explanation of the preference for the low NPV option might be that the number “1,000” acquires special significance among numbers that are merely multiples of “100.” We therefore use tasks in which the numbers are “300” and “600,” or “400” and “800.” Second, we present the choice pairs in delay-by-payoff matrices, with the options designated as “plans,” as shown in Figure 1 (see also Manzini et al., 2010; Gigliotti & Sopher, 1997). In our informal inquiries and initial surveys, the choice pairs were presented as follows:

Which option do you prefer?

- Receive \$500 in one year and receive \$500 in three years.
- Receive \$1,000 in two years.

A partial explanation of the preference for the low NPV option might be that it “ended earlier” in a reading from left to right. Delay-by-payoff matrices resolve this problem, and ensure that, when zero outcomes are revealed, nothing changes in the presentation format, except for the appearance of the zeroes.

<Insert Figure 1 about here>

So the question is how preference changes when, for instance, the choice between £800 in two years and {£400 in one year, £400 in three years} is changed into a choice between {£0 in one

year, £800 in two years, £0 in three years} and {£400 in one year, £0 in two years, £400 in three years}. According to the Tradeoff Model (TM), concealed zeroes are not inferred. With zero outcomes concealed, TM produces a preference for the low NPV option (£800 in two years) when duration outweighs concave utility. When, however, the low NPV option is changed into {£0 in one year, £800 in two years, £0 in three years}, year one has no weight, because zero utility accumulates until year one, year two is weighted by the accumulated utility $u(0) + u(800) = u(800)$, and so is year three, i.e., $u(0) + u(800) + u(0) = u(800)$. Thus, compared with a situation in which zeroes are concealed, year three has weight, which increases the duration of the low NPV option. In addition, when the high NPV option is changed into {£400 in one year, £0 in two years, £400 in three years}, year one is weighted by the accumulated utility $u(400)$, year two as well, i.e., $u(400) + u(0) = u(400)$, and year three is weighted by the accumulated utility $u(400) + u(0) + u(400) = 2u(400)$. Thus, compared with a situation in which zeroes are concealed, year two has weight, which decreases the duration of the high NPV option. In sum, revealing zeroes increases the duration of the low NPV option, and decreases the duration of the high NPV option, and should thus increase preference for the high NPV option.

The Discounted Instantaneous Utility Model (DIUM) can produce a preference for the low NPV option if concave utility, or diminishing sensitivity, is replaced by convex utility, which would have to outweigh discounting. However, because zero outcomes have zero utility, and therefore do not affect the discounted utilities of the options, DIUM predicts that preference for the low NPV option should be unaffected by revealing zero outcomes.

The Sequences Model (SM) shares its prediction with either TM or DIUM, depending on the specific assumptions invoked. To reach the *TM prediction*, SM may assume that concealed zeroes are *not* inferred (the No-Zero Sequences Model, or SMNZ). It can produce a preference for the low NPV option if concave utility is replaced by convex utility (see Table 1). When the zeroes

are revealed, both options become sequences that deviate from a uniform distribution, but the high NPV option deviates less from a uniform distribution than the low NPV option. Thus, if SMNZ assumes a desire for spreading, it predicts that revealing zeroes increases preference for the high NPV option, as TM does on other grounds.

To reach the *DIUM prediction*, SM may assume that concealed zeroes *are* inferred (the Zero Sequences Model, or SMZ). Both options are sequences that deviate from a uniform distribution, but the high NPV option deviates less from a uniform distribution than the low NPV option. SMZ can then produce a preference for the low NPV option if convex utility reinforces distaste for spreading or outweighs a desire for spreading, or if distaste for spreading outweighs concave utility (see Table 1). However, because zero outcomes, revealed *or* concealed, are considered anyway, SMZ predicts that preference for the low NPV option should be unaffected by revealing zero outcomes, as DIUM does on other grounds.

It thus appears that SM can have it both ways. However, this is more apparent than real, because both SMNZ and SMZ fail to accommodate the prior evidence, as discussed earlier. On the one hand, SMNZ cannot reconcile the asymmetric hidden-zero effect with the preference for decreasing over increasing sequences; on the other hand, SMZ cannot produce the asymmetric hidden-zero effect whatsoever. Because there is no parameterization of SM that accommodates prior evidence, SM fails to produce an a priori prediction for the present experiment.

Method

Participants

Participants were 703 British residents, recruited through Maximiles (now rebranded to Bilendi), an Internet service that awards its members points, tradeable for goods and services, for completing surveys. The sample was 57% female, with an average age of 39. The great majority (86%) were employed (61% full time), and a majority (59%) had an academic degree (41%

bachelors, 15% masters, 3% PhD). For 50% of the participants, the total combined income of all members of the family was between £20,000 and £59,999, for 15% it was £60,000 or more, and for 24% it was less than £20,000, with 11% preferring not to report their income.

The samples of all our experiments, including this one, were recruited in “one go,” and no participants were ever removed from the samples (except for three participants in Experiment 4, who were dropped from further analysis due to a programming error). We recruited rather large samples, because, in the subsequent modeling experiments, participants would disperse across candidate models as “their” model, and, in the present experiment, they disperse across two between-participants conditions. The largest sample is collected in Experiment 3, which is a modeling experiment with three between-participants conditions.

Stimuli

There were four choice pairs, as displayed in Table 2. Two pairs were compromised of a “low NPV option” and a “high NPV option” as discussed throughout the theoretical introduction, or a single dated outcome in period 2, and a sequence of two equal outcomes in periods 1 and 3 affording the same total amount. In two other pairs, the sequence was changed by increasing the outcome in period 1 and correspondingly decreasing the outcome in period 3. According to each candidate model, this should increase the preference for the sequence (and so it did).

<Insert Table 2 about here>

Procedure

Participants were randomly assigned to either the zero condition or the no-zero condition. As mentioned earlier, choice pairs were presented in delay-by-payoff matrices, with the options designated as “plans.” The participant selected either the Plan A button or the Plan B button at the bottom of the screen, and could change his or her selection before confirming it by clicking on the

arrow in the bottom-left corner of the screen. The order of the choice pairs, and the left-right order of the options in each choice pair, was randomized across participants.

Results

The results are reported in Table 2. Majorities preferred the low NPV option to the high NPV option in the no-zero condition, and, in the zero condition, preference for the high NPV option increased, such that, now, majorities preferred *that* option. As suspected, the majorities preferring the low NPV option to the high NPV option in the no-zero condition were smaller than in our informal inquiries and initial surveys.

Discussion

The results favored the Tradeoff Model (TM) over the Discounted Instantaneous Utility Model (DIUM), because TM correctly predicted an increased preference for the high NPV option in the zero condition, whereas DIUM predicted that preference between the high and low NPV options would not change. The evidence also favors TM over the Sequences Model (SM), because TM provides an a priori prediction of the observed change in preference, whereas SM does not; a posteriori, the results would favor SMNZ over SMZ.

Experiment 2:

No Size Fits All

The purpose of Experiment 2 is to conduct a first Bayesian contest between our candidate models. In the previous experiment, we tested the models in aggregate choice behavior, basically assuming that the average participant is representative of all participants; in this experiment, we recognize that the “representative agent” may not exist, or that the average participant is perhaps representative of many, but not all.

The models differ in their policy on zero outcomes. The Discounted Instantaneous Utility Model (DIUM) holds that it does not matter whether revealed or concealed zeroes are ignored or

considered, because zero outcomes have zero utility, and therefore do not affect discounted utility. The Tradeoff Model (TM) holds that it *does* matter how zeroes are treated, and proposes that concealed zeroes are ignored, whereas revealed zeroes are considered. Finally, the Sequences Model (SM) also holds that it matters how zeroes are treated, but is silent on whether concealed zeroes are, as in the No-Zero Sequences Model (SMNZ), ignored or, as in the Zero Sequences Model (SMZ), considered. We therefore include both variants of SM as candidate models in our Bayesian contest.

Method

Participants

Participants were 520 British residents, recruited through Maximiles. The sample was 61% female, with an average age of 37. The great majority (83%) were employed (61% full time), and a large majority (94%) had an academic degree (63% bachelors, 27% masters, 5% PhD). For 52% of the participants, the total combined income of all members of the family was between £20,000 and £59,999, for 20% it was £60,000 or more, and for 16% it was less than £20,000, with 13% preferring not to report their income.

Stimuli

Stimuli were selected from a space spanned by two options and three time periods, making for six slots. At each stimulus construction trial, each slot was filled with an amount randomly selected from among £800, £600, £400, and £200, along with a “blank” (zero) outcome, so that the outcomes were equally spaced and below the £1,000 mark. Prior criteria were that the options should not be identical, that at least one option should be a sequence, that, across options, at least one nonzero outcome occurred in each time period, that, across options and time periods, at least one zero outcome occurred, that, in no time period, the options had identical outcomes (avoiding *common consequences*), and that the options afforded the same total amount.

Selection of choice tasks took place in two runs, one searching for nondominance relations (ND), the other searching for dominance relations (D), between options affording the same total amount. Choice tasks in which Net Present Value could, depending on the discount rate, yield either preference order of the options were discarded. Choice tasks in which Net Present Value yielded the same preference order of the options as duration (computed with outcomes, rather than unknown outcome utilities) were also discarded. Run ND identified 28 choice tasks, while run D identified 1 choice task. One choice task identified by run ND, and the choice task identified by run D, included options with equal duration.

From the 29 choice tasks identified, seven with nondominance relations were dropped to reduce repetitiveness in the stimulus set, and thus fatigue and boredom on the part of the participants. These were replaced by four tasks with dominance relations: Two included a decreasing sequence and an increasing sequence affording the same total amount (one with a linear progression, the other with a nonlinear progression); the remaining two included a constant sequence and either a linearly decreasing sequence or a linearly increasing sequence affording the same total amount. These four tasks served to relate our results back to those from Manzini et al. (2010) and Gigliotti and Sopher (1997). Table 3 shows the 26 choice tasks used in Experiment 2, with the first five exhibiting dominance relations.

<Insert Table 3 about here>

Procedure

Participants were informed they would make 26 choices, and that, for each choice, they should assume they were entitled to receive money, which they could receive according to two plans. They were told that all plans spanned a period of three consecutive years, and that they could choose the plan they liked best. Participants started with a rehearsal choice between two options exhibiting a transparent dominance relation, one yielding £200 in one year, £300 in two

years, and £400 in three years, and the other yielding £100 less in each consecutive year. After making their choice, we told participants they probably noticed that the choice was very easy: One option yielded more money in each year, so surely they wanted that. We added, however, that the remaining choices would not be so easy, because the options would yield the same total amount.

Model estimation and evaluation

We investigated the support for each model from each individual participant using a Bayesian analysis. A Bayesian analysis allows the user to make clear and coherent statements about the direction and magnitude of the evidence in favor of each model for each individual participant. It automatically adjusts for each model's complexity using the priors assigned to the parameters. The key measure of a Bayesian analysis is the result of the prior probability of the parameters multiplied by the likelihood of the data given those parameters, and the logic of this measure can be conveyed with a stock-market metaphor. Each model can be thought of as a stock portfolio, which starts with the same amount of money (the prior) that can be split amongst its companies (the parameter settings). The monetary gain from a company is the product of how much money is invested in that company and its performance, which is its share price at sale divided by the share price at purchase (the likelihood of the data given that parameter setting). The overall monetary gain from the portfolio is the sum of the results from each company in the portfolio. The narrower the range of companies in a portfolio, the more that can be invested in each company, and the larger the amount that will be gained if those companies do well. It thus becomes intuitive why, for a given model, a narrower range of prior parameter settings is better than a wider range if that those parameters perform equally well.

We present the evidence in the form of Bayes Factors (Kass & Raftery, 1995), which are statements of the relative likelihood of obtaining the data, D , from a pair of models, M_1 and M_2 ,

$$BF = \frac{p(D | M_1)}{p(D | M_2)},$$

where $p(D|M)$ is the probability of the data given model M . Using Bayes Factors as a measure of the evidence easily allows readers to integrate their personal opinions about the plausibility of each model with that evidence. What a reader should believe about each model after observing the data is just their relative prior beliefs in two models $p(M_1) / p(M_2)$ multiplied by the Bayes Factor,

$$\frac{p(M_1 | D)}{p(M_2 | D)} = BF \frac{p(M_1)}{p(M_2)},$$

and, when we report the posterior probabilities of models, we assume that the priors over models are always equal.

Bayes Factors are calculated by essentially finding the likelihood of the data for every possible setting of each model's parameters, with each set of parameters weighted by its prior belief in that set of parameters, which is known as the marginal likelihood

$$p(D | M) = \int p(D | M, \theta) p(\theta | M) d\theta,$$

where θ is a set of parameters for model M . We did not discover an analytic solution to this equation, but on visual inspection we noticed that the posterior distributions were not particularly narrow compared to the prior distributions for individual participants. This potentially allows us to sidestep complex simulation methods such as Markov chain Monte Carlo and instead use a simple Monte Carlo approximation for the marginal likelihood:

$$p(D | M) = \frac{1}{n} \sum_{j=1}^n p(D | M, \theta_j),$$

where the θ_j are samples from the prior over parameters $p(\theta | M)$. When the posterior is not too narrow compared to the prior distribution, samples drawn from the prior will often fall in regions of high posterior probability, yielding a stable estimate of the marginal likelihood. We assessed

the stability of our marginal likelihoods by running our main analysis twice. The same model was found to have the highest marginal likelihood for 99.2% of the participants, and, as an absolute measure of stability, the average deviation in model posterior probability between the two runs, assuming equal priors for all models, was 0.001 across all participants and models.

Unpacking the Bayesian analysis even further, the likelihood of the data was simply the likelihood of the model producing the same set of binary choices as a participant, which can be expressed as:

$$\log(p(D | M, \theta)) = \sum_{i=1}^m (x_i \log(\hat{P}_i) + (1 - x_i) \log(1 - \hat{P}_i)),$$

where x_i acquires the value of 1 if the high NPV option is chosen from choice pair i , and the value of 0 otherwise, \hat{P}_i is the predicted probability of choosing the high NPV option from choice pair i using model M , and m is the number of choice pairs. The predicted probability of choosing the high NPV option is given by the logistic distribution function:

$$\hat{P}_i = \frac{1}{1 + \exp(-\varepsilon(U_{HNPV_i} - U_{LNPV_i}))},$$

where ε regulates the degree to which choice probabilities are determined by the model, with a value of zero indicating that choice is totally random (only “noise”). For the choice pairs in which one option dominates the other (e.g., the choice between a decreasing sequence and an increasing one yielding the same payoff), we introduce an adjustment $0 \leq \vartheta \leq 1$ for dominance detection. Imagine an avid dominance detector. If a substantive model, by its proper psychology, already predicts high probabilities that dominant options will be chosen, then the adjustment is small (ϑ close to zero); conversely, if the model predicts low probabilities that dominant options will be chosen, then the adjustment is large (ϑ close to one). The predicted probability of choosing the dominant over the dominated option is given as:

$$\hat{P}_i = \vartheta + (1 - \vartheta) \frac{1}{1 + \exp(-\varepsilon(U_{HNPV_i} - U_{LNPV_i}))},$$

where \hat{P}_i is the probability of choosing the dominant option, as the combined result of the model and dominance detection.

The weighing parameters of TM and SM were transformed into relative weights, thus encouraging an independent role of ε in the distribution function and making values easier to interpret. For TM, the accumulated utilities were weighted by $1 / (1 + \kappa)$, and duration by $\kappa^* = \kappa / (1 + \kappa)$. Analogously, for SM, the accumulated utilities were weighted by $1 / (1 + |\beta| + |\sigma|)$, the improvement scores by $\beta^* = \beta / (1 + |\beta| + |\sigma|)$, and the spreading scores by $\sigma^* = \sigma / (1 + |\beta| + |\sigma|)$.

For all three models, we specified the utility function as the power function $u(x) = x^\gamma$ (e.g., Abdellaoui et al., 2010; Dai & Busemeyer, 2014) with $0 \leq \gamma \leq 5$, and assumed a uniform prior distribution over its entire range.⁶ For δ and ϑ , we assumed the same uniform prior between zero and one. For the regulation parameter ε , we assumed a uniform prior over the range of 0 to 100. For TM, we specified, as mentioned in Footnote 5, the time-weighing function as the identity function $w(\hat{t}) = \hat{t}$, and let κ have a uniform prior from 0 to 100. For SM, we allowed participants to either like ($\beta > 0$) or dislike ($\beta < 0$) improvement, and to either like ($\sigma > 0$) or dislike ($\sigma < 0$) spreading. Both parameters were given uniform prior distributions from -100 to 100 . The priors for κ , β , and σ cover the entire range from exclusive weight of accumulated utilities and (practically) exclusive weight of the model components associated with κ , β , and σ . Additional analyses with different sets of prior distributions produced qualitatively the same headline results.⁷

Model recovery

We will estimate models with four to five parameters from individual choice behavior in 26 binary choice tasks. Because information is relatively scarce, one might wonder whether the

information is *sufficient* at all to identify the “right” model. We therefore undertook a model-recovery exercise. At each step in the exercise, a parameter set was drawn from the prior of the *generating model*, and one individual’s worth of data were generated from those parameters. The marginal likelihoods were found for the data, and the winning model was selected, thus treated as the *recovered model*. This procedure was repeated 1,000 times for each of the models taking a turn as the generating model. The results are reported in Table 4. The rate of model recovery is high: On more than 87% of the 4,000 occasions, the generating model was successfully recovered. We therefore conclude that the information provided by our experimental paradigm is sufficient to identify the “right” model.

<Insert Table 4 about here>

Parameter estimation

While our model recovery exercise demonstrated that we could identify the best model using individual data, the best parameters for individuals were not as easily identifiable. As shown in the supplemental material, the posterior distributions over parameters for each individual were broad and sometimes multimodal, so any single point does not provide a good summary of these distributions.⁸ Instead, we ran a secondary analysis to identify the best parameters. Participants were split into groups according to which model produced the highest marginal likelihood for their individual data. The data of each group were aggregated, and the aggregate data was used to determine the posterior distribution over the parameters for the model associated with that group. Thus, the posterior distribution over parameters for each model was determined using a different group of participants.

We needed a new approximation to calculate the posterior distributions, because the simple approximation we used in the individual-analyses does not work well when the posterior distributions are narrow. Thus, in the group-level analyses, we used the Metropolis-Hastings

algorithm, which is a Markov chain Monte Carlo method that starts in a particular state, and produces a chain of states that are used as samples from the posterior distribution over parameters. We started each chain using a random draw from the prior distribution and collected one million states for each chain, discarding the first half so as to remove the influence of the arbitrary starting state. During this “burn-in period,” we continually recalibrated the proposal distribution, i.e., the distribution that determined what the next state could be, by setting the covariance of the Gaussian proposal distribution to a constant multiple of the sample covariance of a recent set of states, plus a small additive diagonal factor, in order to ensure the covariance matrix did not collapse to zero (Rosenthal, 2011). The samples were summarized by finding both the median of each parameter and the Highest Density Interval (HDI), which was the single shortest continuous interval that contained 95% of the posterior distribution. Across two runs of this analysis using random starting points for the sampler, we found consistent medians and HDIs.

Main contest and parallel contest

We run a “main contest” and a “parallel contest” between models. The *main contest* is between complete specifications of the four substantive psychological models, i.e., TM, DIUM, SMNZ, and SMZ. The *parallel contest* is between the same models, and reduced forms of these models, either by setting ϑ to zero, or by setting ε to zero, or by setting both ϑ and ε to zero. We next explain the logic behind this.

In the main contest, each candidate model includes the parameter ϑ , an adjustment for dominance detection. This parameter increases the probability of choosing the dominant option; when it is zero, an individual has no stronger proclivity toward the dominant option than is given by the model. The adjustment parameter, however, does make the models less parsimonious, so that, when dominance detection is modest, models without the adjustment parameter might, by their parsimony, fare better. In the Appendix, we specify all four models with and without the

adjustment parameter, and report the relative performance of the models accordingly, but we are generally interested in how the models perform regardless of whether increased parsimony pays off or not, just as we are interested, in the main contest, in how the models perform regardless of whether \mathcal{G} is high or low.

Another aspect of the main contest is that we only include substantive psychological models as candidates. Each model, however, includes, as in modeling exercises elsewhere, the parameter ε , which regulates the degree to which choice probabilities are determined by the model; when it is zero, choice is totally random. As choice becomes increasingly random, a *Random-choice model* (RM) might, by its parsimony, fare better than any substantive psychological model. In the Appendix, we include RM as a fifth model. When specified without the parameter \mathcal{G} , RM holds that *any* choice is random; when specified with this parameter, RM holds that choice is random *unless* one option dominates another, in which case choice of the dominant option may become more likely. In our discussion of the results from the main contest, we will refer to the results from the parallel contest.

Descriptive accuracy

While Bayes factors allow us to assess how well the models perform *relative to one another*, they do not provide an absolute standard against which to measure the performance of each model. So, in addition to the Bayes factors, we calculated *Bayesian p-values*, which measure descriptive accuracy by evaluating the posterior predictions of a model against the observed data (see Meng, 1994; Myung, Karabatsos, & Iverson, 2005). In particular, a Bayesian *p-value* is the probability that the discrepancy between the model and the posterior predicted data is greater than the discrepancy between the model and the observed data, $p\{T(D^{\text{rep}}) > T(D)|D, M\}$, with values

below the conventional threshold of .05 meaning that it is unlikely that the observed data would have arisen from a particular model, or that the model does not fit well.

A discrepancy function T must be chosen to evaluate the “closeness” of the data to the model. While the Pearson chi-square discrepancy function is a common choice (Myung et al., 2005), the predicted probabilities from our model were often within machine precision of zero or one, which makes that discrepancy function difficult to work with. Instead, we chose the sum-of-squares discrepancy function

$$T(D; \theta, M) = \sum_i (d_i - p(d_i | \theta, M))^2,$$

where d_i are the data from each individual choice i . Using the discrepancy function T , the formula for the Bayesian p -value is

$$\begin{aligned} p\{T(D^{\text{rep}}) > T(D) | D, M\} &= \\ &= \int \frac{1/n \sum_k^n [p(D_k^{\text{rep}} | \theta, M) I(T(D_k^{\text{rep}}; \theta, M) > T(D; \theta, M))] p(D | \theta, M) p(\theta | M)}{p(D|M)} d\theta, \end{aligned}$$

where I is an indicator function that is 1 if the argument is true and 0 if it is false.

As with the Bayes factors, the Bayesian p -value needs to be approximated. To do this, we follow a procedure similar to that for approximating the Bayes factors for individual participants. Specifically, we drew samples from $P(\theta|M)$, and, for each sample, we generated a single D^{rep} by sampling choices using the choice probabilities predicted by model M with parameters θ . We calculated the discrepancies $T(D)$ and $T(D^{\text{rep}})$ using the predictions generated by model M using parameters θ . If, while iterating over the samples, $T(D^{\text{rep}}) > T(D)$, we added $p(D|\theta, M) / p(D|M)$ to the Bayesian p -value, otherwise we added nothing.

Results

Table 3 shows the percentages of participants choosing the high NPV option among those supporting the respective models in the Bayesian analyses, to be reported shortly, and among all those participating in the experiment. The results for choice pairs #2 to #4 confirm the results from Manzini et al. (2010) and Gigliotti and Sopher (1997): Participants preferred a constant sequence to an increasing one, and they preferred a decreasing sequence to an increasing one, but their preference between a decreasing sequence and a constant one was basically undecided.

Figure 2 depicts the proportions of participants whose support for one model over *all* other candidate models in the main contest is, by informal rules of thumb (Jeffreys, 1961; Wetzels et al., 2011), anecdotal (both Bayes Factors greater than one), substantial (both BFs greater than three), strong (both BFs greater than 10), very strong (both BFs greater than 30), and decisive (both BFs greater than 100). TM receives the most support; e.g., it receives substantial support from over 40% of the participants, while the other candidate models together receive substantial support from under 20% of the participants.

<Insert Figure 2 about here>

For a strictly quantitative assessment of relative performance (without reliance on rather arbitrary rules of thumb), we averaged posterior probabilities (the ratios between which are, under the assumption of equal priors, the Bayes Factors) of the respective models across participants, yielding .54 for TM, .17 for DIUM, .16 for SMNZ, and .12 for SMZ. Figure A1 depicts the average posterior probabilities of the five candidate models (with or without adjustment for dominance detection) in the parallel contest. TM performs best, and RM performs second best, suggesting that participants engaged in a fair amount of guessing. Moreover, RM obtains more support *with* adjustment for dominance detection than *without*, suggesting that guessing was not just a matter of distraction or carelessness, but rather of not knowing what to do in the absence of

dominance relations. Table A1 shows how the supporters of the four candidate models in the main contest distribute across the five candidate models in the parallel contest. The participants either stay with their model or switch to RM. In raw numbers, most switches are from TM to RM, but that is because there are so many more supporters of TM; it is, actually, a *minority* that switches from TM to RM, just as we see minorities switching from DIUM to RM, and from SMZ to RM. Among supporters of SMNZ, on the other hand, we see a true exodus to RM, with over 70% of the supporters taking part in it, meaning that SMNZ is particularly prone to attracting individuals with “noisy” choice behavior.

Bayes Factors and posterior probabilities inform us about the relative performance of the models, not about how well each model describes the choices of those best described by it. Figure 3 depicts the observed proportions of participants choosing the high NPV option among those lending anecdotal support to the respective models in the main contest (horizontal), plotted against the average predicted probabilities of choosing the high NPV option generated by the models (vertical). Not surprisingly, each model describes the choices of *its* supporters best. Interesting is how TM and DIUM describe the supporters of one another. Most choice pairs, except for #1 to #4, were selected for generating an ordinal contrast between preference entailed by duration (TM, with linear utility) and preference entailed by Net Present Value (DIUM, with linear utility). Table 3 shows that supporters of TM indeed mostly prefer the low NPV option, whereas supporters of DIUM mostly prefer the high NPV option, and Figure 3 shows that TM tends to underpredict high NPV choices among supporters of DIUM, whereas DIUM tends to overpredict high NPV choices among supporters of TM (as the sequences model do as well).

<Insert Figure 3 about here>

For a quantitative assessment of descriptive accuracy, Table A2 shows the proportions of Bayesian p -values lower than .05 among participants lending anecdotal support to the respective

models in the parallel contest. Not counting RM without dominance detection, which holds that each choice is a coin toss, and therefore receives a Bayesian p -value of zero by definition, the models fit well overall, with TM, whether with or without adjustment for dominance detection, showing no p -values lower than .05. When specified with adjustment for dominance detection, SMZ also does not show p -values lower than .05, and nor does RM, which means that, also in terms of descriptive accuracy, a combination of guessing and dominance detection captures well what a portion of our participants is doing. Across specifications, DIUM is least accurate, and SMNZ second-least accurate.

Table 5 provides, for each model, the medians of the parameter posteriors, along with the lower and upper bounds of the 95% Highest Density Intervals (HDIs). In all four models, the utility function is concave (i.e., $\gamma < 1$), thus exhibiting diminishing sensitivity. This argues against all explanations of prior evidence that resorted to convex utility.

<Insert Table 5 about here>

Concavity of the utility function is least pronounced among supporters of TM, and *most* pronounced among supporters SMNZ and SMZ, which may seem counterintuitive, since the sequences models include a desire for spreading ($\sigma > 0$) as a separate component, which would seem to relieve the utility function from capturing any desire for spreading through concave utility ($\gamma < 1$). This would be the case, however, if the spreading component and concave utility would both favor an option for approximating a uniform distribution. But this is not what the spreading component does: It *disfavors* an option for *departing from* a uniform distribution. The result is a hydraulic relationship between the spreading component, which disfavors an option for departing from a uniform distribution, and concave utility, which *does* favor an option for approximating a

uniform distribution. Thus, not even in the sequences models, γ provides a “pure” assessment of diminishing sensitivity, independently from a desire for spreading.

While the HDIs for σ are wide, so is the HDI for κ in TM. This is because the utility function of TM exhibits only slight concavity (γ near to 1) over the outcomes of sequences that, in any given choice pair, total the same amount, so that differences between the accumulated utilities of the sequences change little across choice pairs. Therefore, TM explains choice behavior mostly on the basis of differences between the durations of the sequences, which creates redundancy between κ and ε , in that higher values of κ can more easily be compensated by lower values of ε . For instance, if concavity vanishes, and the differences between the accumulated utilities of the sequences are therefore always zero, an increase of κ from 18.73 (lower bound of the HDI) to 100.00 (upper bound of the HDI), which is an increase of κ^* (relative weight) from .9493 to .9901, can be compensated by a decrease of ε from 4.2350 to 4.0605 (taken at equal distance from the median). In comparison to the narrow range of κ^* , we obtain, for the sequences models, wider ranges of β^* (between $-.79$ and $-.46$ for SMNZ, between $-.51$ and $.06$ for SMZ) and σ^* (between $-.20$ and $.87$ for SMNZ, between $.48$ and $.93$ for SMZ); yet, it can be concluded that these models tend to converge on distaste for improvement and a desire for spreading.

Discussion

In a Bayesian contest among candidate models at the disaggregate level, the Tradeoff Model (TM) came out as best describing most participants. In a parallel contest, RM came out as the strongest competitor to TM, especially when, alongside the choice processes described by these models, dominance detection was taken into account. The overall picture is that *weighing cumulative utility against the duration of utility accumulation, guessing, and dominance detection go a long way in explaining the behavior of our participants.*

In Experiment 1, we compared the competing models with respect to their predictions about the effects of revealing zero outcomes on aggregate choice behavior, and, in Experiment 2, we compared the models with respect to their ability to describe individual choice behavior. In the next experiment, we validate our disaggregate analysis against the asymmetric hidden-zero effect, one thread of evidence from the theoretical introduction.

Experiment 3:

The Cumulative Weight of Nothing

We expect the asymmetric hidden-zero effect to be driven by those supporting the Tradeoff Model (TM), since the later zero added to the smaller-sooner outcome (changing SS into SS_0) has cumulative weight, whereas the sooner zero added to the larger-later outcome (changing LL into LL_0) has not. In contrast, we expect those supporting the Discounted Instantaneous Utility Model (DIUM) to be immune to zero outcomes, because zero outcomes have zero utilities, and thus leave the discounted utilities of sequences unchanged.

Drawing on the analysis in Table 1, the No-Zero Sequences Model (SMNZ) would expect the later zero to have an effect as far as a desire for improvement reinforces a desire for spreading, and would expect the sooner zero to have no effect at all if the desire for improvement precisely offsets the desire for spreading. SMNZ would expect a *symmetric* hidden-zero effect if the desire for improvement eclipses the desire for spreading. The Zero Sequences Model (SMZ), however, expects *any* hidden-zero effect to be symmetrical, because revealing one zero means that the other zero will be inferred.

It should be emphasized that Bayesian analysis does not include the hidden-zero tasks serving to evaluate the hidden-zero effect, so that the above derivations are truly *predictions*.

Method

Participants

Participants were 848 British residents, recruited through Maximiles. The sample was 54% female, with an average age of 42. The great majority (84%) were employed (61% full time), and a large majority (91%) had an academic degree (61% bachelors, 25% masters, 6% PhD). For 50% of the participants, the total combined income of all members of the family was between £20,000 and £59,999, for 23% it was £60,000 or more, and for 16% it was less than £20,000, with 11% preferring not to report their income. Overall, the sample characteristics were very similar to those in Experiment 2.

Stimuli

The stimuli were the same as in Experiment 2, except that one set of three choice tasks was added. These were choices between a smaller-sooner and a larger-later outcome, with the smaller-sooner outcome being £150, £300, or £450 today, and the larger-later outcome being £200, £400, and £600, respectively, in 1 year. Thus, the interest rate offered was held constant at 33.33%. Either no zero outcomes were stated, or only the sooner zero was stated, or only the later zero was stated.

Procedure

The three hidden-zero tasks were presented first. Participants were asked to imagine for these three decisions that they were entitled to receive money, that they would be given 2 options to choose from, and that, in each case, they would have to select the option they preferred. The order of the choice tasks was randomized across participants. Furthermore, the statement of zero outcomes was manipulated between participants. As in past research on the hidden-zero effect, the options were written in two consecutive lines, with *LL* (or *LL₀*) always following *SS* (or *SS₀*). The rest of the procedure was identical to that in Experiment 2.

Results

Bayesian model contests

Figure 2 depicts the proportions of participants lending progressively more convincing support for one model over all other candidate models in the main contest. By the informal rules of thumb, TM again receives the most support; e.g., it receives more than twice as much substantial support as the other candidate models together. On average, posterior probabilities were .53 for TM, .18 for DIUM, .17 for SMNZ, and .12 for SMZ, almost identical to those in Experiment 2. Figure A1 depicts the average posterior probabilities of the five candidate models in the parallel contest, also almost identical to those in Experiment 2. Table A1 shows how the supporters of the four candidate models in the main contest distribute across the five candidate models in the parallel contest, and they essentially do so in the same way as in Experiment 2. Figure 4 depicts the observed proportions of participants choosing the high NPV option among those lending anecdotal support to the respective models in the main contest, plotted against the average predicted probabilities of choosing the high NPV option generated by the models, and Table A2 shows the proportions of Bayesian p -values lower than .05 among those participants lending anecdotal support to the respective models in the parallel contest. It is safe to conclude that, both quantitatively and qualitatively, the results show a strong resemblance to those from Experiment 2.

<Insert Figure 4 about here>

External validation against the asymmetric hidden-zero effect

Concerning the three hidden-zero tasks, we computed, for each participant, the proportion of choices favoring the larger-later outcome (a measure of patience). The top panel of Figure 5 shows the results for those providing anecdotal support for the respective models. As expected, DIUM supporters are not affected by zero outcomes. In contrast, those supporting TM, SMNZ,

and SMZ *are* affected by them, but more so by the later zero than by the sooner one. Let p_L , p_S , and p_N be the average proportions of patient choices in the later-zero, sooner-zero, and no-zero conditions, respectively. A measure of the asymmetry of the hidden-zero effect is then $(p_L - p_N) - (p_S - p_N) = p_L - p_S$, where a *strictly* asymmetric hidden-zero effect is $p_L - p_N > 0$ and $p_S - p_N = 0$. The asymmetry is more pronounced among those supporting TM, $p_L - p_S = .18$, 95% CI [.08; .28], than among those supporting SMNZ, $p_L - p_S = .08$, 95% CI [-.13; .29], and SMZ, $p_L - p_S = .15$, 95% CI [-.13; .43], but those supporting TM do seem to be affected somewhat by the sooner zero, $p_S - p_N = .09$, 95% CI [-.02; .19]. We must realize, however, that we are including all participants who provide *anecdotal* support for a model, several of whom do not lend *convincing* support to it. The bottom panel of Figure 5 therefore shows the results among those lending progressively more convincing support to TM. We see the asymmetric hidden-zero effect that we expected: An effect of the later zero, but no effect of the sooner one, among those lending substantial support, $p_L - p_N = .27$, 95% CI [.16; .38], $p_S - p_N = .06$, 95% CI [-.06; .19], among those lending strong support, $p_L - p_N = .19$, 95% CI [.06; .32], $p_S - p_N = .02$, 95% CI [-.13; .17], among those lending very strong support, $p_L - p_N = .24$, 95% CI [.09; .39], $p_S - p_N = .01$, 95% CI [-.17; .18], and among those lending decisive support, $p_L - p_N = .24$, 95% CI [.05; .43], $p_S - p_N = -.05$, 95% CI [-.27; .16].

<Insert Figure 5 about here>

Discussion

The results from Experiment 3 confirm that the asymmetric hidden-zero effect is driven by those supporting the Tradeoff Model (TM): Among those individuals, the later zero has an effect but the sooner zero has not, while, among those supporting the Discounted Instantaneous Utility

Model (DIUM), neither zero has an effect. This constitutes an external validation of both models, and, for TM, it corroborates the cumulative weighing of time.

The Sequences Model (SM), however, fails the test of external validity. Among supporters of both variants, there were indications that *both* zeroes produced a rise in patience, but that the later zero produced a greater rise in patience than the sooner zero. For the No-Zero Sequences Model (SMNZ), this would mean that a desire for improvement outweighs a desire for spreading, whereas we see in Table 5 that *distaste* for improvement outweighed a desire for spreading. For the Zero Sequences Model (SMZ), any asymmetry of the hidden-zero effect is at odds with its assumption that, once a decision involves outcome sequences, concealed zeroes are inferred by the decision maker.

As to the Bayesian model contest itself, the overall picture that emerges from it is the same as in Experiment 2: *Weighing cumulative utility against the duration of utility accumulation, guessing, and dominance detection go a long way in explaining the behavior of our participants.*

Experiment 4:

Top It Up

In Experiment 4 we conducted an incentivized version of our previous experiments. This was done at the University of Warwick, with participants being entered into a lottery where they had a 1:40 chance that one of their choices would be paid out for real. Earnings were in the form of “Eating at Warwick” top ups (see [<http://www2.warwick.ac.uk/services/retail/eating>]), which can be added to the ID cards of every student and staff member at the University of Warwick. Primarily, as the name implies, the card is used for eating at university restaurants and coffee shops, but it can also be used to purchase groceries, theatre tickets (at the Warwick Arts Center), and books (at the university bookstore). These cards can be preloaded by anyone who has the

person's student or staff number. For instance, parents can prepay for their child's food by topping up their cards from home. These prepayments take effect immediately.

Method

Participants and payments

Participants were 239 students from the University of Warwick, Coventry, Great Britain. (Three participants were dropped from further analysis due to a programming error.) They were recruited on the University of Warwick's research participation system (SONA), and participated in exchange for a promise that six participants would receive one of their choices, worth up to £80, paid for real. On average, the sample was 19 years of age, with 59% being female.

Stimuli

The outcomes from Experiments 2 and 3 were divided by 20, yielding outcomes of £10, £20, £30, and £40. Real total payoffs therefore ranged from £20 (if choice pair #12 were selected for payout) to £80 (if choice pair #5 were selected for payout). Furthermore, the delays from the previous experiments were divided by six, yielding delays of two, four, and six months. All else about the stimuli remained the same.

Procedure

The experiment was run online, on November 5, 2014. In the introduction to the online questionnaire, participants were informed that they would make 26 decisions between payment plans, and that the payment plans were Eating at Warwick top ups on their student card. They were told that every plan would specify a top up amount to be paid in two months (January 5, 2015), in four months (March 5, 2015), and in six months (May, 5 2015), and that, for every choice they made, there would be a chance it would be paid for real. They were further told that six participants would be chosen at random, and that the chosen participants would actually be given one of their preferred plans. The instruction then read: "To enable us to enter you in the

draw please provide your email address.” They were informed that everybody would receive a number upon completion of the experiment, and that the experimenter would randomly select six of these numbers; if their number was selected, one of the choices they made would be picked at random and paid out for real. They were further informed that everyone would receive an email listing the randomly selected numbers, and that the participants with these numbers would be sent a separate email stating what they would receive and when.

The next page of the instruction presented a rehearsal task, analogous to the one used in Experiments 2 and 3. After the rehearsal trial, participants were reminded that each option they chose might be paid out for real, so that they would have to choose carefully.

Results

Figure 2 depicts the proportions of participants lending progressively more convincing support for one model over all other candidate models in the main contests. By the informal rules of thumb, TM is still the best performing model, but it faces fiercer competition from SMZ, and especially from DIUM; e.g., it receives 6.8% more substantial support than DIUM, its closest competitor, but it receives 5.1% less substantial support than DIUM and the other candidate models together. On average, the posterior probabilities were .36 for TM, .29 for DIUM, .13 for SMNZ, and .22 for SMZ, also attesting to the rise of SMZ, and especially of DIUM. Figure A1 shows the average posterior probabilities of the five candidate models in the parallel contest, and a stiff competition between the models is evident, with RM and DIUM now sharing the second place, and each coming right after TM, which is still in the first place. All four substantive psychological models obtained more support *with* adjustment for dominance detection than *without*. Table A1 shows how the supporters of the four candidate models in the main contest distribute across the five candidate models in the parallel contest, and they basically do so as in the same way as in Experiments 2 and 3, except that there is less migration from SMNZ to RM.

Figure 6 depicts the observed proportions of participants choosing the high NPV option among those lending anecdotal support to the respective models in the main contest, plotted against the average predicted probabilities of choosing the high NPV option generated by the models. There is more dispersion around the identity line than in Experiments 2 and 3, especially for TM and SMNZ. However, Table A2 shows that, when specified with adjustment for dominance detection, all models in the parallel contest, including TM and SMNZ, hit the 0% rate of Bayesian p -values lower than .05. TM remained at the 0% rate when specified without adjustment for dominance detection, as did DIUM, meaning that, also in terms of descriptive accuracy, DIUM performed better than in Experiments 2 and 3.

<Insert Figure 6 about here>

Table 5 provides, for each model, the medians of the parameter posteriors, along with the lower and upper bounds of the 95% Highest Density Intervals (HDIs). First, across the board, but especially among those supporting SMNZ and SMZ, choice behavior is determined more by the respective models (higher ε), meaning that the choice behavior is less “noisy.” Second, while adjustment for dominance detection (ϑ) was, among supporters of DIUM, already low in previous experiments, it dropped, among supporters of TM, SMNZ, and SMZ, to low levels as well, with the lowest level reached by TM. This is consistent with the fact that, in the parallel contest, the models benefited more than in Experiments 2 and 3 from an increased parsimony when the adjustment parameter was suppressed. The reason why ϑ goes to zero is a markedly reduced probability of choosing the decreasing sequence instead of the constant sequence it dominates (analogous to choice pair #3 in Table 3). Figure 6 plots the four choice pairs with dominance relations as solid circles. The choice pair involving the decreasing and constant sequence is the one far above the identity line, while the other three choice pairs involving dominance relations

are reasonably close to that line. Thus, one does not want adjustment for dominance detection to make matters worse.

Discussion

The results from Experiment 4, which introduced incentives, differ somewhat from those in Experiments 2 and 3. Most importantly, we saw an increase in support for the Zero Sequences Model (SMZ), which, as shown in Footnote 2, is the holistic analogue of atomistic discounting, and especially the Discounted Instantaneous Utility Model (DIUM). Why this change, and what to conclude from it?

The differences between Experiment 4 and the previous experiments include smaller outcome magnitudes, shorter delays, student population, less abstract outcomes, and the use of non-hypothetical payoffs. All of these changes, especially the use of less abstract outcomes, may have introduced motives in addition to those operating in the previous experiments. Importantly, payment in the “Eating at Warwick” currency may have promoted a discounting of individual outcome utilities, because later payoffs occurred close to the summer break, when many students might be going home or even leaving the university altogether, and would no longer be using their Eating at Warwick card. We thus ascribe the rise in support for DIUM to *a trimmed time horizon* in the mind of the decision makers. If correct, this explanation offers an important insight into the psychology of preferences for sequences, and into the variability of relative model performance across settings.

The main goal of all the changes was to provide non-hypothetical payoffs. This might bring decisions closer to the economic model and reduce errors in responding (e.g., Camerer & Hogarth, 1999; Hertwig & Ortmann, 2001). We indeed found choice behavior to be less “noisy.” We also found a surge in support for DIUM. Because DIUM is similar to the NPV model, which is the standard of economic rationality, one might infer that decisions were indeed brought “closer

to the economic model.” However, the NPV model operates on money, whereas DIUM operates on the *instantaneous utility* of money, as a result of which it actually moves *away from* economic rationality by allowing for violations of dominance. Thus, increasing support for DIUM cannot be automatically taken to be increasing support for “the economic model.” Moreover, *even if* DIUM were taken to be the standard of economic rationality, it would definitely be less plausible that the incentives “magically” led to a discounting of individual outcome utilities than that the incentives led to a trimmed time horizon, which, in turn, led to a discounting of individual outcome utilities. In any case, the Tradeoff Model (TM) still received more support than any other model in the contest.

Finally, there is yet another indication that incentivization did not bring decisions closer to the economic model: The choice between a decreasing and a constant sequence was no longer a coin toss, which by itself is nonnormative, but rather strongly disfavored the decreasing sequence in relation to the constant one. This is a gross violation of dominance, meaning that participants indeed moved away from the economic model. The violation of dominance was problematic for all models in the contest, in that they far overpredicted the probability of choosing the dominant option.

Old Evidence from a New Perspective

Given that the Tradeoff Model (TM), with its cumulative weighing of time, emerges as the winning model overall, the question becomes whether it can also accommodate previously reported evidence on preferences for sequences of unlabeled monetary gains, which, without exception, concerns *aggregate* choice behavior involving *two-outcome* sequences. We saw how it accommodates the asymmetric hidden-zero effect, and how, in disaggregate choice behavior, it is driven by those best described by TM, but what about the remaining evidence? Table 6 compiles that evidence, which, with the exception of the *asymmetric* hidden-zero effect, was addressed by

Read and Scholten (2012). In this section, we discuss how TM accommodates the listed anomalies to choice by the Net Present Value model.

<Insert Table 6 about here>

Front-end amount effect and its reversal

We begin with the single one anomaly that TM, as specified in Equations (3a) and (3b), does *not* explain, the one that Read and Scholten (2012) call the *front-end amount effect*, as first reported by Rao and Li (2011). Consider a choice between “\$10 today” (SS) and “\$30 in two months” (LL), and another between “\$510 today” (SS_F) and “\$500 today and \$30 in two months” (LL_F). The introduction of the common front-end amount of \$500 decreases patience, i.e., the probability of choosing LL_F is lower than the probability of choosing LL . This contradicts TM as specified in Equations (3a) and (3b). By a cumulative weighing of time, the delay of LL is two, while the duration of LL_F is shorter than two, which would benefit LL_F . Furthermore, by concave utility, the utility difference between the options in period 1 (“today”) is less strongly in favor of SS_F than it is in favor of SS , i.e., $u(510) - u(500) < u(10) - u(0)$, which would benefit LL_F as well. On both counts, the front-end amount should *increase* patience.

The front-end amount effect must be appreciated along with a *reversal* of the effect, as reported by Read and Scholten (2012). Consider a choice between “\$510 today” (SS_A) and “\$530 in two months” (LL_A), and another between “\$1,510 today” (SS_{FA}) and “\$1,000 today and \$530 in two months” (LL_{FA}). Now, the introduction of the common front-end amount of \$1,000 *increases* patience, i.e., the probability of choosing LL_{FA} is higher than the probability of choosing LL_A . This is, of course, compatible with a cumulative weighing of time and concave utility, but, then, *why* the front-end amount effect, and *why* its reversal?

At this point, it must be recognized that the prime motivation for developing the tradeoff model were *similarity effects* (Scholten & Read, 2010), showing that intertemporal choices are

influenced the similarity between the options along the time and outcome attributes. Accordingly, we attribute the reversal of the front-end amount effect to the similarity of the options along the outcome attribute upon introduction of the front-end amount.⁹ In the choice between *SS* and *LL*, a large majority preferred *LL*, and, indeed, the interest rate on offer was 73.2% per month (72,800% per year). Therefore, according to TM,

$$u(30) - u(10) > \kappa_{T|X} w(2),$$

where the left-hand side is the outcome advantage of *LL*, and the right-hand side is the time advantage of *SS*, with the former being larger than the latter, so that *LL* is preferred. In the subscript to κ , which measures impatience, X is the outcome advantage, and T is the time advantage, with $T|X$ meaning that κ also depends on the similarity of the outcomes, as explained below. Upon introduction of the front-end amount, which leaves the interest rate unchanged, the majority preference for *LL* shrunk toward indifference between *SS_F* and *LL_F*, which we write as:

$$u(30) - (u(510) - u(500)) = \kappa_{T|X_F} w(\hat{t}_F).$$

As discussed by Scholten and Read (2010), κ can, in any given local context, acquire two values, depending on “similarity.” The lower value applies when the time advantage of *SS* falls below a threshold, and the higher value applies when the time advantage falls above the threshold. This gives rise to a step function relating the time advantage to the weight assigned to the advantage (see Scholten & Read, 2010, Figure 3). Crucially, the threshold depends on the relative similarity of the outcomes and the delays. When the outcomes are treated as “dissimilar,” the threshold for the time advantage of *SS* rises, and the time advantage is therefore more likely to be *subthreshold*; this translates into the low value of κ , or less impatience. When the outcomes, on the other hand, are seen as “similar,” the threshold for the time advantage of *SS* falls, and the time advantage is thus more likely to be *suprathreshold*; this translates into the high value of κ , or more impatience.

We argue that introduction of the front-end amount increases the similarity between the options along the outcome attribute, and, because the outcomes are more likely to be treated as “similar,” the time advantage of SS is more likely to be suprathreshold, i.e., $\kappa_{T|X_F} > \kappa_{T|X}$. If this similarity effect outweighs duration, by which $\hat{t}_F < 2$, and concave utility, by which $u(510) - u(500) < u(10)$, it drives the preference relation between SS_F and LL_F toward indifference.

In the choice between SS_A and LL_A , however, a large majority preferred SS_A , and, indeed, the interest rate on offer was now “only” 1.9% per month (still 26.0% per year).¹⁰ According to TM,

$$u(530) - u(510) < \kappa_{T|X_A} w(2).$$

Upon introduction of the front-end amount, the majority preference for SS_F shrunk toward indifference between SS_{FA} and LL_{FA} , which we write as:

$$u(530) - (u(1,510) - u(1,000)) = \kappa_{T|X_{FA}} w(\hat{t}_{FA}),$$

where the similarity effect, i.e., $\kappa_{T|X_{FA}} > \kappa_{T|X_A}$, countervails, *but does not outweigh*, duration, by which $\hat{t}_{FA} < 2$, and concave utility, by which $u(1,510) - u(1,000) < u(510)$. One reason why similarity no longer prevails may be that \$1,510 and \$1,000 (in the reverse front-end amount effect) are less likely to be judged as “similar” than \$510 and \$500 (in the front-end amount effect). The other reason is that, by concave utility, the difference between $u(510)$ and $u(1,510) - u(1,000)$ is larger than the difference between $u(10)$ and $u(510) - u(500)$, meaning that concave utility becomes more prevalent. Thus, TM can accommodate the front-end amount effect and its reversal by combining duration and concave utility with similarity effects.

Relocation effect

Read and Scholten (2012) also reported what we call a *relocation effect*. Consider a choice between “\$510 today” (SS_F) and “\$500 today and \$30 in two months” (LL_F), and another between

“\$10 today and \$500 in two months” (SS_B) and “\$530 in two months” (LL_B). Thus, the common amount of \$500 is relocated from the front end in the first choice to the back end in the second choice. Relocation decreases patience, i.e., the probability of choosing LL_B is lower than the probability of choosing LL_F , although it leaves the interest rate unchanged. Formally, we might have:

$$u(30) - (u(510) - u(500)) = \kappa_{T|X_F} w(\hat{t}_F),$$

and

$$(u(530) - u(500)) - u(10) < \kappa_{T|X_B} (w(2) - w(\hat{t}_B)).$$

Assuming that \$530 and \$500 are not less likely to be treated as “similar” than \$510 and \$500, the relocation effect arises when concave utility outweighs duration. The time disadvantage of LL_B is smaller than that of LL_F , i.e., $2 - \hat{t}_B < 1$ while $\hat{t}_F > 1$. Countervailing duration is concave utility, by which the outcome advantage of LL_B is smaller than that of LL_F , i.e., $u(530) < u(30) + u(500)$ and $-u(500) - u(10) < -u(510)$. The effect of concave utility can be expected to be substantial, because the utility decrements from $u(30)$ to $u(530) - u(500)$, and from $-(u(510) - u(500))$ to $-u(10)$, are substantial (see also Read & Scholten, 2012, Figure 2). Thus, TM can accommodate the relocation effect through concave utility, and the desire for spreading it generates.

Violations of independence

Three additional anomalies to the Net Present Value model are *violations of independence* (Koopmans, 1960), a condition that is invoked in the axiomatic derivation of the discounted utility model. Independence states that time periods yielding identically valued outcomes regardless of one’s choice (a *common consequence*) should cancel out, and therefore not affect preference. The evidence shows that adding a common consequence *does* affect preference. Loewenstein (1987) provided a first demonstration in the domain of consumption, which, as shown by Loewenstein

and Prelec (1993), can be accommodated by the sequences model. We continue to concentrate on unlabeled monetary gains, and discuss violations of independence in that restricted domain.

Read and Scholten (2012) reported a violation that we call the *common-consequence effect*. Consider a choice between “\$300 today” (SS) and “\$400 in 50 weeks” (LL), and another between “\$300 today and \$350 in 4 weeks” (SS_C) and “\$350 in 4 weeks and \$400 in 50 weeks” (LL_C). In the choice between SS and LL , few participants preferred LL , but, in the choice between SS_C and LL_C , more participants preferred LL_C , a small but reliable effect. Therefore, the common consequence increases patience, i.e., the probability of choosing LL_C is higher than the probability of choosing LL . Formally, we might have:

$$u(400) - u(300) < \kappa_{T|X} w(50),$$

and

$$(u(400) - u(300)) + (u(350) - u(350)) \leq \kappa_{T|X_C} (w(\hat{t}_{LL_C}) - w(\hat{t}_{SS_C})).$$

Obviously, the common consequence does not affect the utility difference between the options, but the common consequence is written down so as to imply that it increases, by the similarity effect, the marginal impact of time, i.e., $\kappa_{T|X_C} > \kappa_{T|X}$. Outweighing the similarity effect, however, is the effect of duration: As a result of the common consequence, the duration of SS_C is longer than the delay of SS , i.e., $\hat{t}_{SS_C} > 0$, and the duration of LL_C is shorter than the delay of LL , i.e., $\hat{t}_{LL_C} < 50$, thus increasing patience. As mentioned earlier, the effect was small, and this analysis indeed suggests a delicate balance between similarity and duration.

The remaining violations of independence in the literature involve common consequences located *before*, rather than *between*, the smaller-sooner and the larger-later outcome. Urminsky and Kivetz (2011) reported what they call a *mere-token effect*, by which the common consequence increases patience, but does not increase patience much more as the token becomes larger. However,

Read and Scholten (2012) showed, drawing on their formulation of the tradeoff model, that this depends on how much larger the token becomes. Consider a choice between “\$200 in one week” (SS) and “\$400 in one year” (LL), another between “\$50 tomorrow and \$200 in one week” (SS_M) and “\$50 tomorrow and \$400 in one year” (LL_M), and yet another between “\$5,000 tomorrow and \$200 in one week” (SS_{AM}) and “\$5,000 tomorrow and \$400 in one year” (LL_{AM}). A minority preferred LL over SS , participants were equally divided SS_M and LL_M , and a majority preferred LL_{AM} over SS_{AM} . Therefore, the probability of choosing LL_M was higher (by a narrow but reliable margin) than the probability of choosing LL , and the probability of choosing LL_{AM} was higher (by a wider margin) than the probability of choosing LL_M . This result, which we call the *amplification effect*, can be written as follows:

$$u(400) - u(200) < \kappa_{T|X} (w(365) - w(7)),$$

$$(u(400) - u(200)) + (u(50) - u(50)) = \kappa_{T|X_M} (w(\hat{t}_{LL_M}) - w(\hat{t}_{SS_M})), \text{ and}$$

$$(u(400) - u(200)) + (u(5,000) - u(5,000)) > \kappa_{T|X_{AM}} (w(\hat{t}_{LL_{AM}}) - w(\hat{t}_{SS_{AM}})).$$

The observed effects are, according to TM, carried entirely by duration. Figure 7 shows how duration, under a cumulative weighing of time, decreases as the token increases, separately for SS (which the tokens transform into SS_M and SS_{AM}) and LL (which the tokens transform into LL_M and LL_{AM}). With the \$50 token being introduced, the duration of SS decreases, under linear utility, from seven to six days, whereas the duration of LL decreases more, from 365 to 329 days. If duration outweighs the similarity effect, by which $\kappa_{T|X_M} > \kappa_{T|X}$, and diminishing sensitivity, by which the marginal impact of delay or duration decreases with their length, i.e., $w'(365) < w'(7)$, we obtain the mere-token effect. With the \$50 token being replaced by a \$5,000 token, the duration of SS decreases, under linear utility, from six to four days, while the duration of LL decreases much more, from 329 to 190 days. If, as is likely, duration outweighs diminishing

sensitivity, we obtain the amplification effect. Therefore, the violations of independence seen in the literature can be comprehensively dealt with by a cumulative weighing of time, which yields the duration of utility accumulation.

<Insert Figure 7 about here>

Violations of dominance

A final piece of evidence concerns *violations of dominance*, as reported by Scholten and Read (2014). One violation involved monetary gains. Consider a choice between “\$75 today” (SS) and “\$100 in one year” (LL), and another between “\$75 today *and* \$5 in one year” (SS_D), which dominates SS , and “\$100 in one year” (LL_D), which is the same as LL . A large majority preferred SS , but a *minority* preferred SS_D . Another violation of dominance involved monetary losses. Consider a choice between “paying \$75 today” (SS) and paying \$100 in one year” (LL), and another between “paying \$75 today *and* paying \$5 in one year” (SS_D), which is dominated by SS , and “paying £100 in one year” (LL_D), which is the same as LL . A majority preferred SS , but a *larger* majority preferred SS_D . The explanation offered by TM is that the duration of SS_D is longer than the delay of SS ; with \$75 today accumulating to \$80 in one year, the duration will be “a bit more than half a year,” *significantly* longer than “today.” With gains, this drastically decreases the appeal of SS_D , and, with losses, it increases the appeal of SS_D (not as drastically, because, as we will point out in the General Discussion, people differ in their time preference for losses). These violations of dominance, involving losses as well as gains, can therefore also be dealt with by a cumulative weighing of time.

General Discussion

Our theoretical and empirical analysis follows up on rather scarce research that has been done on preferences for sequences of unlabeled monetary gains. We extended the intertemporal tradeoff model (Read & Scholten, 2012) from two-period to multiple-period sequences, and

proposed that people choose between sequences by trading off accumulated utilities of gains against the duration of utility accumulation. In Experiment 1, aggregate choice behavior provided qualitative support for the revised tradeoff model over its original specification, and over two other candidate models from the literature: The discounted instantaneous utility model (which holds that individuals discount the instantaneous utility of monetary outcomes, rather than, as often assumed in the literature, the outcomes themselves) and Loewenstein and Prelec's (1993) sequences model. In three subsequent experiments, the last one incentivized, disaggregate choice behavior provided quantitative support for the tradeoff model in a Bayesian model contest. In Experiment 3, we showed that the asymmetric hidden-zero effect, which was a prime motivation for revising the tradeoff model, was indeed driven by those best described by the tradeoff model. Finally, we showed how the revised tradeoff model can accommodate remaining evidence from the literature, if it makes room for similarity effects, which was a prime motivation for developing the tradeoff model (Scholten & Read, 2010).

Paramorphic representations of choice

The cumulative weighing of time should be appreciated as a paramorphic feature of the tradeoff model, like other “gestalt” properties in models of decision making. Tversky and Kahneman (1992), in proposing cumulative prospect theory, probably did not contend that people actually go through the mathematically ingenious computation of decision weights; rather, the computations were meant to capture the intuition that people are more sensitive to extreme outcomes in a probability distribution than to intermediate ones (Fennema & Wakker, 1997). Similarly, Loewenstein and Prelec (1993), in proposing the sequences model, probably did not contend that people actually go through laborious comparisons between cumulative utilities afforded by a sequence and those that would be afforded by a uniform sequence; rather, the comparisons were meant to capture the intuition that people are sensitive to improvement and

spreading of outcomes in a sequence. On a less speculative note, Loomes (2010), in proposing his *perceived relative argument model* of risky choice, stressed that it is “a model of decisions often made quite quickly and on the basis of perceptions rather than after long deliberation involving complex calculation. The structure of the model is therefore intended to capture tendencies in the ways perceptions are formed and judgments are made: it is not suggested that people actually make calculations strictly according to the formulae but, rather, that the formulae capture key features of the ways in which decision parameters are perceived and processed.” In the same spirit, the cumulative weighing of time is meant to capture the intuition that people are sensitive to how quickly the outcomes in a sequence accumulate: “How much utility do I get until this point in time, how much do I get until that point, and how long does it take?”

The tradeoff model as an attribute-based choice model

In contrast with choice between single dated outcomes, which, according to Scholten et al. (2014), follows a stochastic *ratio* rule, choice between outcome sequences was assumed to follow a stochastic *difference* rule. If, as in Equation (3b), the difference between weighted durations is traded off against differences in accumulated utilities by a currency parameter κ , rather than a *nonlinear* tradeoff function Q (Scholten & Read, 2010; Scholten et al., 2014), the tradeoff model might be rearranged into an (arithmetically) “discounted utility” model, in which choice would be *alternative based* rather than attribute based. However, as mentioned in the introduction, this interpretation of the tradeoff model has the undesirable characteristic that a sequence of *positive* outcomes might receive a *negative* “discounted utility.” Moreover, as discussed in the previous section, even choice between sequences can be susceptible to similarity effects, which expose attribute-based comparisons.

Features of sequences

We examined the feasibility of our approach to preferences for sequences in experiments that adopted standard practice in research on this topic, which is that the options afford the same total amount of money, and that the money is distributed over contiguous future time periods. Future model contests should be run under less restrictive conditions. As discussed earlier, the sequences model holds that discounting of individual outcome utilities interacts with the “gestalt” properties of outcomes sequences (improvement and spreading) when a sequence starts in (or close to) the present (e.g., 0, 1, 2) or when time periods are noncontiguous (e.g., 1, 2, 10). The tradeoff model would continue to operate as it does with contiguous future time periods, although it would be necessary to allow the time-weighting function w to exhibit diminishing sensitivity to delays, so as to capture the disproportionate impact of shorter durations of outcome sequences. Similarly, the discounted instantaneous utility model would continue to operate as it does with contiguous future time periods, although it would be necessary to generalize the exponential discount function, which exhibits constant sensitivity, into a hyperbolic discount function, which exhibits diminishing sensitivity, and thus captures the disproportionate impact of shorter delays to the outcomes in a sequence.

Future model contests should also consider choices between options that afford *different* total amounts of money. In the introduction, we gave the example of a choice between the rising profile $\{\$100, \$150\}$ and the falling profile $\{\$150, \$100\}$, in which the options afford the same total amount of money, and therefore imply a 0% interest rate. One alternative to this would be a choice between $\{\$100, \$250\}$ and $\{\$200, \$100\}$, which, as the choice between \$100 in one year and \$150 in two years, implies a 50% interest rate. We used options with the same total payoffs not only to ensure continuity with past research, but also, and much more importantly, to ensure identification of choice tasks in which the Net Present Value model would dictate a certain choice

regardless of the discount rate, and the tradeoff model, by the duration of outcome accumulation, would dictate the opposite choice. We thus obtained a particularly strong contrast between the tradeoff model and the model of economic rationality, but future research might be guided by other motives.

Money and consumption

Most empirical research on preferences for sequences has used sequences of consumption experiences or labeled monetary money starting either in the present or in the future (Chapman, 1996; Guyse et al., 2002; Loewenstein & Prelec, 1993; Loewenstein & Sicherman, 1991; Matsumoto et al., 2000; Read & Powell, 2002). We tried to tap into “raw” preferences for sequences, or preferences purely driven by the magnitudes of, and delays to, the outcomes in the sequence, by using sequences of *unlabeled future money*, thus reducing the role of factors like *expectations*, which might exist in preferences for profiles of wage payments or life-time health (Chapman, 1996), the difference in *uncertainty* between present and future money (Epper, Fehr-Duda, & Bruhin, 2011; Halevy, 2008), and *satiation* with consumption experiences that money can buy (Reuben, Sapienza, & Zingales, 2010; Baucells & Sarin, 2007). Loewenstein and Prelec (1993) partly motivated the sequences model with the evidence on preferences for sequences of income from labor or rent (Loewenstein & Sicherman, 1991), but applied the model (equivalent the Zero Sequences Model) to preferences for sequences of consumption experiences, and found that it outperformed a consumption analogue of the discounted instantaneous utility model. In our application of the models to preferences for sequences of unlabeled future money, it was the other way around. More must be done to better understand the variability of relative model performance across settings, and thus come to further insight into the psychology of preferences for sequences.

It must be recognized, however, that consumption decisions, although a legitimate focus of inquiry, do complicate a *controlled* analysis of intertemporal preferences. Reuben et al. (2010)

identified a number of hurdles for a controlled analysis of intertemporal consumption decisions: Not only do consumption goods easily lead to satiation (e.g., 100 candy bars in one year; see Estle et al., 2007), they also differ from individual to individual in their desirability (some love candy bars, others hate them), they introduce intertemporal uncertainty about tastes (how much will I like a candy bar in one year?), and they are not easily divisible into an arbitrary number of units (trading off two candy bars in one year against one candy bar today already entails a 100% yearly interest rate). Money removes all hurdles in one sweep. As Abdellaoui et al. (2010) conclude: “It is hard to see how one could construct sequences of, say, candy bars like we did for money amounts and, to the best of our knowledge, no study has been able to measure the utility of consumption for intertemporal choice. Developing methods to measure the intertemporal utility for consumption is an important topic for future research but as yet this cannot be fulfilled” (p. 861).

The distinction between money and consumption is also important from a theoretical perspective, and Abdellaoui et al. (2010) actually made the above observations in discussing economic theory. Economic analyses of intertemporal choice have been dominated by Samuelson’s (1937) Discounted Utility (DU) model (Loewenstein & Prelec, 1992). Like the Net Present Value (NPV) model, the DU model assumes that the future is discounted exponentially, but unlike the NPV model, which is defined over monetary outcomes, the DU model is defined over consumption. The complete picture of the connection between the NPV model and the DU model is as follows: (1) Given two monetary sequences, choose the one with the highest NPV at the prevailing interest rate, because it places you on the highest attainable intertemporal budget constraint (the income decision), and, (2) given the intertemporal budget constraint, choose the consumption stream with the highest DU (the consumption decision). Analyses of choices between monetary sequences place us in stage (1), and psychological models, *including* the

Discounted Instantaneous Utility model, question the NPV model as an accurate description of how people make these choices. Abdellaoui et al. (2010) basically argue that, given the problems in measuring intertemporal utility for consumption, this is the best we can do, which, according to the authors, “seems to lead to the undesirable conclusion that economics is unable to quantify one of its principal models” (p. 861).

Sequences of losses

A natural thing to do is to extend the analysis from preferences for sequences of gains to preferences of sequences of *losses* (Guyse & Simon, 2011). As our discussion of Scholten and Read’s (2014) violations of dominance (in the previous section) suggests, the tradeoff model, with its cumulative weighing of time, may apply equally well to preferences for sequences of losses as to preferences for sequences of gains. The aggregate data showed positive time preference, in that a longer duration was better. However, it is likely that these aggregate data conceal disaggregate diversity. Yates and Watts (1975) showed that only half of their participants preferred to pay an amount of money later rather than sooner (positive time preference); the other half preferred to pay it sooner rather than later (*negative* time preference). Hardisty, Appelt, and Weber (2013) showed that negative time preference for monetary losses may transpire *even* in aggregate data. In response to this evidence, Scholten et al. (2016) extended the tradeoff model to accommodate heterogeneity in preferences for single dated losses; when applied to preferences for *sequences* of losses, the tradeoff model must take that heterogeneity into consideration as well, as, indeed, any model should.

Sequences of gains and losses

Further complications arise if we broaden the scope to sequences that *combine* gains and losses. These come even closer to the real-life examples with which we started this paper. The versatile discounted instantaneous utility model, with its atomistic approach to preferences for

sequences, can readily be extended to mixed sequences. It is not immediately obvious how the tradeoff model and the sequences model can be extended to mixed sequences; perhaps, drawing on the property of gain-loss separability from prospect theory (Wu & Markle, 2008), it may be that people evaluate the gain and loss parts of the sequences separately; thus, in the tradeoff model, the cumulative weighing of time would occur separately in the two domains, and so would, in the sequences model, the computation of improvement and spreading scores.

The issue of mixed sequences lies wide open as an avenue for future research. There is *some* evidence (Jiang, Hu, & Zhu, 2014; Rao & Li, 2011; Scholten & Read, 2014; Sun & Jiang, 2015), and a common pattern is that the introduction of mixed sequences increases patience. It may, more generally, be that sequences reduce the salience of time: In choice between single dated outcomes, one only needs to attend to two different outcomes occurring at two different delays, but, when choice involves sequences, the *evolution* of outcomes must be tracked over time, which may draw attention away from delays, and toward the outcomes (Jiang et al., 2014; Shu & Jiang, 2015). However, our analysis shows that there is much more to preferences for sequences than just an overall rise in patience, and past evidence on preference for unmixed sequences already attests to this, in the form of the *asymmetric* hidden-zero effect, the front-end amount effect, the relocation effect, and the amplification effect (see Table 6). The evidence presented in this paper further exposes the rich psychology behind preferences for sequences.

References

- Abdellaoui, M., Attema, A. E., & Bleichrodt, H. (2010). Intertemporal tradeoffs for gains and losses: An experimental measurement of discounted utility. *Economic Journal*, *120*, 845-866.
- Baucells, M., & Sarin, R. K. (2007). Satiation in discounted utility. *Operations Research*, *55*, 170-181.
- Brealey, R. A., Myers, S. C., Allen, F., & Mohanty, P. (2012). *Principles of corporate finance*. New York, NY: McGraw-Hill Education.
- Camerer, C. F., & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty*, *19*, 7-42.
- Chapman, G. B. (1996). Expectations and preferences for sequences of health and money. *Organizational and Human Decision Processes*, *67*, 59-75.
- Dai, J., & Busemeyer, J. R. (2014). A probabilistic, dynamic, and attribute-wise model of intertemporal choice. *Journal of Experimental Psychology: General*, *143*, 1489-1514.
- Duffy, S., & Smith, J. (2013). Preference for increasing wages: How do people value various streams of income? *Judgment and Decision Making*, *8*, 74-90.
- Duffy, S., Smith, J., & Woods, K. (2015). How does the preference for increasing payments depend on the size and source of the payments? *Management Science Letters*, *5*, 1071-1080.
- Epper, T., Fehr-Duda, H., & Bruhin, A. (2011). Viewing the future through a warped lens: Why uncertainty generates hyperbolic discounting. *Journal of Risk and Uncertainty*, *43*, 169-203.
- Estle, S. J., Green, L., Myerson, J., & Holt, D. D. (2007). Discounting of monetary and directly consumable rewards. *Psychological Science*, *18*, 58-63.
- Fennema, H., & Wakker, P. (1997). Original and cumulative prospect theory: A discussion of empirical differences. *Journal of Behavioral Decision Making*, *10*, 53-64.
- Fisher, I. (1930). *The theory of interest*. New York, NY: Macmillan.

- Gigliotti, G., & Sopher, B. (1997). Violations of present-value maximization in income choice. *Theory and Decision*, 43, 45-69.
- Guyse, J. L., Keller, L. R., & Eppel, T. (2002). Valuing environmental outcomes: Preferences for constant or improving sequences. *Organizational Behavior and Human Decision Processes* 87, 253-277.
- Guyse, J. L., & Simon, J. (2011). Consistency among elicitation techniques for intertemporal choice: A within-subjects investigation of the anomalies. *Decision Analysis*, 8, 233-246.
- Halevy, Y. (2008). Strotz meets Allais: Diminishing impatience and the certainty effect. *American Economic Review*, 98, 1145-1162.
- Hardisty, D. J., Appelt, K. C., & Weber, E. U. (2013). Good or bad, we want it now: Fixed-cost present for gains and losses explains magnitude asymmetries in intertemporal choice. *Journal of Behavioral Decision Making*, 26, 348-361.
- Hertwig, R., & Ortmann, A. (2001), Experimental practices in economics: A methodological challenge for psychologists? *Behavioral and Brain Sciences*, 24, 383-451.
- Hey, J. D. (2003). *Intermediate microeconomics*. New York, NY: McGraw-Hill Education.
- Jeffreys, H. (1961). *Theory of probability*. Oxford, UK: Oxford University Press.
- Jiang, C.-M., Hu, F.-P., & Zhu, L.-F. (2014). Introducing upfront losses as well as gains decreases impatience in intertemporal choices with rewards. *Judgment and Decision Making*, 9, 297-302.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 363-391.
- Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90, 773-795.
- Koopmans, T. C. (1960). Ordinal utility and impatience. *Econometrica*, 28, 287-309.

- Leland, J. W. (2002). Similarity judgments and anomalies in intertemporal choice. *Economic Inquiry, 40*, 574-581.
- Loewenstein, G. (1987). Anticipation and the valuation of delayed consumption. *Economic Journal, 97*, 666-684.
- Loewenstein, G., & Prelec, D. (1992). Anomalies in intertemporal choice: Evidence and an interpretation. *Quarterly Journal of Economics, 107*, 573-597.
- Loewenstein, G., & Prelec, D. (1993). Preferences for sequences of outcomes. *Psychological Review, 100*, 91-108.
- Loewenstein, G., & Sicherman, N. (1991). Do workers prefer increasing wage profiles? *Journal of Labor Economics, 9*, 67-84.
- Loomes, G. (2010). Modeling choice and valuation in decision experiments. *Psychological Review, 117*, 902-924.
- Macaulay, F. R. (1938). *Some theoretical problems suggested by the movements of interest rates, bond yields and stock prices in the United States since 1856*. New York, NY: National Bureau of Economic Research.
- Magen, E., Dweck, C. S., & Gross, J. J. (2008). The hidden-zero effect. Representing a single choice as an extended sequence reduces impulsive choice. *Psychological Science, 19*, 648-649.
- Manzini, P., Mariotti, M., & Mittone, L. (2010). Choosing monetary sequences: Theory and experimental evidence. *Theory and Decision, 69*, 327-354.
- Matsumoto, D., Peecher, M. E., & Rich, J. S. (2000). Evaluations of outcome sequences. *Organizational Behavior and Human Decision Processes, 83*, 331-352.
- Meng, X. (1994). Posterior predictive *p*-values. *The Annals of Statistics, 22*, 1142-1160.

- Myung, J. I., Karabatsos, G., & Iverson, G. J. (2005). A Bayesian approach to testing decision making axioms. *Journal of Mathematical Psychology, 49*, 205-225.
- Rao, L. L., & Li, S. (2011). New paradoxes in intertemporal choice. *Judgment and Decision Making, 6*, 122-129.
- Read, D., Frederick, S., & Scholten, M. (2013). DRIFT: An analysis of outcome framing in intertemporal choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 573-588.
- Read, D., Olivola, C. Y., & Hardisty, D. (in press). The value of nothing: Asymmetric attention to opportunity costs drives intertemporal decision making. *Management Science*.
- Read, D., & Powell, M. (2002). Reasons for sequences preferences. *Journal of Behavioral Decision Making, 15*, 433-460.
- Read, D., & Scholten, M. (2012). Tradeoffs between sequences: Weighing accumulated outcomes against outcome-adjusted delays. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*, 1675-1688.
- Reuben, E., Sapienza, P., & Zingales, L. (2010). Time discounting for primary and monetary rewards. *Economics Letters, 106*, 125-127.
- Rosenthal, J. (2011). Optimal proposal distributions and adaptive MCMC. In: S. Brooks, A. Gelman, G. Jones, & X.-L. Meng (Eds.), *Handbook of Markov Chain Monte Carlo* (pp. 93-112). London: Chapman and Hall.
- Rubinstein, A. (2003). "Economics and psychology"? The case of hyperbolic discounting. *International Economic Review, 44*, 1207-1216.
- 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). The psychology of intertemporal tradeoffs. *Psychological Review*, *117*, 925-944.
- Scholten, M., & Read, D. (2013). Time and outcome framing in intertemporal tradeoffs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 1192-1212.
- Scholten, M., & Read, D. (2014). Better is worse, worse is better: Violations of dominance in intertemporal choice. *Decision*, *1*, 215-222.
- Scholten, M., Read, D., & Sanborn, A. (2014). Weighing outcomes *by* time or *against* time? Evaluation rules in intertemporal choice. *Cognitive Science*, *38*, 399-438.
- Scholten, M., Read, D., Walters, D. J., Erner, C., & Fox, C. R. (2016). *A diagnosis of tolerance and aversion to debt*. Forthcoming.
- Stevens, J. R. (2016). Intertemporal similarity: Discounting as a last resort. *Journal of Behavioral Decision Making*, *29*, 12-24.
- Sun, H.-Y., & Jiang, C.-M. (2015). Introducing money at any time can reduce discounting in intertemporal choices with rewards: An extension of the upfront money effect. *Judgment and Decision Making*, *10*, 564-570.
- Thaler, R. (1981). Some empirical evidence on dynamic inconsistency. *Economics Letters*, *8*, 201-207.
- Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business*, *59*, S251-S278.
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297-323.
- Urminsky, O., & Kivetz, R. (2011). Scope insensitivity and the “mere token” effect. *Journal of Marketing Research*, *48*, 282-295.

Wakker, P. P. (2010). *Prospect theory for risk and ambiguity*. Cambridge, MA: Cambridge University Press.

Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E.-J. (2011). Statistical evidence in experimental psychology: An empirical comparison using 855 t tests. *Perspectives on Psychological Science*, 6, 291-298.

Wu, G., & Markle, A. B. (2008). An empirical test of gain-loss separability in prospect theory. *Management Science*, 54, 1322-1335.

Wu, C.-Y., & He, G.-B. (2012). The effects of time perspective and salience of possible monetary losses on intertemporal choice. *Social Behavior and Personality*, 40, 1645-1654.

Yates, J. F., & Watts, R. A. (1975). Preference for deferred losses. *Organizational Behavior and Human Performance*, 13, 294-306.

Appendix:**Parallel Contest**

In this Appendix, we report the results of a contest, paralleling the main contest, between five candidate models. There are the four substantive psychological models from the main contest. Then there is the Random-choice Model (RM), which is technically a reduced version of any of the four substantive psychological models (when ε is zero). Each of the five candidate models was specified either with or without a parameter for dominance detection (ϑ), making for 10 models in total. Figure A1 shows the posterior probabilities of the five candidate models in the parallel contest, averaged across participants, with the dark and light shades indicating the parts of the support obtained with and without the parameter for dominance detection. Table A1 shows how the supporters of the four candidate models in the main contest distribute across the five candidate models in the parallel contest. Each of the five candidate models takes the posterior probabilities of its variant with adjustment for dominance detection and its variant without adjustment for dominance detection, averaged across the posterior probabilities of all 10 models in the parallel contest. Averaging posteriors is the same as having an average prior for ϑ , i.e., an average of the prior with a spike at zero and the prior with a uniform distribution over the interval from 0 to 1. Table A2 shows the proportions of Bayesian p -values lower than .05, an aggregate measure of badness of fit, among those lending anecdotal support to the respective models in the parallel contest.

<Insert Figure A1 about here>

<Insert Table A1 about here>

<Insert Table A2 about here>

Acknowledgments

Marc Scholten, associate professor with habilitation at the Universidade Europeia, Lisbon, Portugal, Daniel Read, professor of Behavioural Science at Warwick Business School, and Adam Sanborn, associate professor at the University of Warwick, Coventry, United Kingdom. The authors thank Neil Stewart, Francisco Garcia dos Santos, and João Nogueira.

In this work, Marc Scholten received support from the Fundação para a Ciência e Tecnologia [grant number PTDC/ MHC-PCN/3805/2012], Daniel Read from the Economic and Social Research Council [grant number ES/K002201/1], and the Leverhulme Trust [grant number RP2012-V-022], and Adam Sanborn from the Economic and Social Research Council [grant number ES/K004948/1].

Correspondence concerning this article should be addressed to Marc Scholten, Universidade Europeia, Quinta do Bom Nome, Estrada da Correia 53, 1500-210 Lisbon, Portugal. E-Mail: marc.scholten@universidadeeuropeia.pt.

TABLE 1. Threads of evidence as dealt with by the Discounted Instantaneous Utility Model (DIUM), the Sequences Model (SM), and the Tradeoff Model (TM).

Model	Thread of evidence		
	Undecided preference between decreasing and constant sequences	Asymmetric hidden-zero effect	Preference for faster accumulation, i.e., $\{1,000, 0, 1,000\} \succ \{0, 500, 0\}$
DIUM	Concave utility countervails discounting	Unaccounted for	Convex utility outweighs discounting
SM	(1) Distaste for improvement countervailed by concave utility and a desire for spreading; (2) distaste for improvement and spreading countervailed by concave utility; (3) convex utility and distaste for improvement countervailed by a desire for spreading; (4) a desire for improvement countervailed by convex utility and distaste for spreading; (5) a desire for improvement and spreading countervailed by convex utility; (6) concave utility and a desire for improvement countervailed by distaste for spreading	Concealed zeroes are not inferred; with the later zero (SS_0), a desire for improvement reinforces a desire for spreading; with the sooner zero (LL_0), a desire for improvement countervails a desire for spreading	(1) Concealed zeroes are not inferred; convex utility; (2) concealed zeroes are inferred; convex utility reinforces distaste for spreading or outweighs a desire for spreading; (3) concealed zeroes are inferred; distaste for spreading outweighs concave utility
TM	Cumulative weighing of time countervails concave utility	Cumulative weighing of time	Cumulative weighing of time outweighs concave utility

TABLE 2. Sequences, duration under a linear utility function, and 95% confidence intervals (LOWER, UPPER) for percentages of participants ($\%_H$) choosing the high NPV option when zeroes are not revealed ($N = 356$) and when they are revealed ($N = 347$), in Experiment 1.

No-zero condition											
#	High NPV				Low NPV				Confidence interval		
	t_1	t_2	t_3	\hat{t}	t_1	t_2	t_3	\hat{t}	LOWER	$\%_H$	UPPER
1	300		300	2.33		600		2.00	38.11	43.26	48.40
2	400		200	2.20		600		2.00	47.34	52.53	57.72
3	400		400	2.33		800		2.00	38.11	43.26	48.40
4	600		200	2.14		800		2.00	49.04	54.21	59.39
Zero condition											
#	High NPV				Low NPV				Confidence interval		
	t_1	t_2	t_3	\hat{t}	t_1	t_2	t_3	\hat{t}	LOWER	$\%_H$	UPPER
1	300	0	300	2.25	0	600	0	2.50	56.26	61.38	66.51
2	400	0	200	2.14	0	600	0	2.50	57.15	62.25	67.35
3	400	0	400	2.25	0	800	0	2.50	55.38	60.52	65.66
4	600	0	200	2.10	0	800	0	2.50	61.31	66.28	71.26

TABLE 3. Choice pairs and percentages of participants choosing the high NPV option among those supporting the Tradeoff Model (TM, $n = 317$), the Discounted Instantaneous Utility Model (DIUM, $n = 80$), the No-Zero Sequences Model (SMNZ, $n = 78$), and the Zero Sequences Model (SMZ, $n = 45$), and among all those participating in Experiment 2 (Sample, $N = 520$).

#	High NPV			Low NPV			Model Supported				Sample
	t_1	t_2	t_3	t_1	t_2	t_3	TM	DIUM	SMNZ	SMZ	
1	800	400	200	200	400	800	77.92	81.25	65.38	68.89	75.77
2	600	400	200	200	400	600	77.29	77.50	65.38	68.89	74.81
3	600	400	200	400	400	400	51.42	61.25	65.38	44.44	54.42
4	400	400	400	200	400	600	90.22	85.00	74.36	82.22	86.35
5	800		800	600	200	800	48.90	40.00	65.38	35.56	48.85
6	800		200	200	800		26.50	73.75	66.67	62.22	42.88
7	600		200	200	600		20.82	65.00	58.97	66.67	37.31
8	400		200	200	400		10.41	60.00	47.44	62.22	28.08
9	600		200		800		37.22	91.25	50.00	55.56	49.04
10	400		200		600		31.55	95.00	53.85	46.67	45.96
11	400		400		800		22.71	90.00	44.87	53.33	39.04
12	200		200		400		21.14	85.00	34.62	48.89	35.38
13	200	200	200		600		42.27	90.00	37.18	64.44	50.77
14	200	400	200		800		35.96	83.75	38.46	62.22	45.96
15	400	200	200		800		40.06	92.50	43.59	64.44	50.77
16	400	400	200	200	800		28.08	80.00	53.85	68.89	43.46
17	600	400	200	400	800		27.44	82.50	50.00	64.44	42.50

TABLE 3 TO BE CONTINUED

TABLE 3. Continued.

18	800	400	200	600	800		26.18	73.75	52.56	71.11	41.35
19	800		200	400	600		20.82	53.75	60.26	53.33	34.62
20	800		200	600	400		11.36	47.50	43.59	57.78	25.77
21	600		200	400	400		10.73	43.75	38.46	53.33	23.65
22	600	200	200	400	600		25.87	85.00	50.00	75.56	42.88
23	600	200	200	200	800		39.75	88.75	62.82	66.67	53.08
24	800		600	600	400	400	18.93	38.75	44.87	37.78	27.50
25	600		600	400	400	400	9.78	27.50	29.49	28.89	17.12
26	400		800	200	400	600	17.67	36.25	42.31	24.44	24.81

^aChoice pairs #10, #11, and #9 correspond to choice pairs #2, #3, and #4 in Table 2.

TABLE 4. Model recovery exercise. Proportions of occasions, out of 1,000, in which a generating model was recovered as a candidate model, out of four.

Generating model	Recovered model			
	TM	DIUM	SMNZ	SMZ
TM	0.94	0.01	0.02	0.03
DIUM	0.12	0.80	0.06	0.02
SMNZ	0.01	0.08	0.89	0.03
SMZ	0.06	0.04	0.03	0.87

TABLE 5. Medians (*Md*), along with the Lower bounds (*L*) and Upper bounds (*U*), of the 95% Highest Density Intervals of model parameters in Experiments 2, 3, and 4 (E2, E3, and E4), among those supporting the Tradeoff Model (TM, $n = 312, 498,$ and 91), the Discounted Instantaneous Utility Model (DIUM, $n = 80, 137,$ and 81), the No-Zero Sequences Model (SMNZ, $n = 80, 142,$ and 23), and the Zero Sequences Model (SMZ, $n = 48, 71,$ and 41).

	TM			DIUM			SMNZ			SMZ						
	<i>L</i>	<i>Md</i>	<i>U</i>	<i>L</i>	<i>Md</i>	<i>U</i>	<i>L</i>	<i>Md</i>	<i>U</i>	<i>L</i>	<i>Md</i>	<i>U</i>				
	ε	3.90	4.15	4.43	ε	0.12	0.29	0.52	ε	0.00	0.80	4.90	ε	0.17	0.61	1.25
	γ	0.95	0.97	0.99	γ	0.46	0.55	0.64	γ	0.00	0.13	0.45	γ	0.00	0.15	0.28
E2	ϑ	0.11	0.16	0.22	ϑ	0.00	0.04	0.13	ϑ	0.14	0.27	0.40	ϑ	0.06	0.19	0.31
	κ	18.73	63.66	100.00	δ	0.39	0.50	0.63	β	-99.99	-58.56	-14.45	β	-29.20	-10.02	6.03
									σ	-25.51	67.47	99.99	σ	27.06	72.43	100.00
	ε	4.20	4.46	4.83	ε	0.32	0.58	0.92	ε	0.00	0.10	1.18	ε	0.18	0.54	1.01
	γ	0.97	0.99	1.00	γ	0.48	0.54	0.60	γ	0.18	0.45	1.31	γ	0.00	0.12	0.26
E3	ϑ	0.29	0.33	0.38	ϑ	0.04	0.17	0.30	ϑ	0.20	0.34	0.43	ϑ	0.24	0.34	0.44
	κ	5.58	37.52	84.41	δ	0.24	0.30	0.36	β	-99.15	-78.31	-31.55	β	-29.43	-8.86	7.90
									σ	1.57	24.00	54.74	σ	25.50	72.28	100.00
	ε	4.21	5.34	8.29	ε	0.71	0.84	1.21	ε	0.03	3.09	8.66	ε	1.58	2.18	2.83
	γ	0.90	0.96	0.99	γ	0.57	0.74	0.94	γ	0.09	0.24	0.66	γ	0.23	0.29	0.36
E4	ϑ	0.00	0.02	0.08	ϑ	0.00	0.05	0.13	ϑ	0.00	0.14	0.31	ϑ	0.00	0.08	0.18
	κ	0.70	4.10	8.88	δ	0.77	0.89	0.98	β	-19.05	-7.26	3.18	β	-62.00	-38.11	-3.38
									σ	36.63	76.87	100.00	σ	9.71	62.25	99.99

TABLE 6. The tradeoff model with a cumulative weighing of time and anomalies to the Net Present Value model reported in the literature.

Designation	Comparison		Effect	Accommodated by Equations (3a) and (3b)?
	Control	Experimental		
Hidden-zero effect	$SS: (x_S, t_S)$	$SS_0: (x_S, t_S; 0, t_L)$	More patience	Yes
	$LL: (x_L, t_L)$	$LL_0: (0, t_S; x_L, t_L)$		
Asymmetric hidden-zero effect	$SS: (x_S, t_S)$	Sooner $SS: (x_S, t_S)$	No effect	Yes
	$LL: (x_L, t_L)$	zero $LL_0: (0, t_S; x_L, t_L)$		
Front-end amount effect	$SS: (x_S, t_S)$	Later $SS_0: (x_S, t_S; 0, t_L)$	More patience	No
	$LL: (x_L, t_L)$	zero $LL: (x_L, t_L)$		
Reverse front-end amount effect	$SS_A: (x_A + x_S, t_S)$	$SS_{FA}: (x_F + x_A + x_S, t_S)$	More patience	Yes
	$LL_A: (x_A + x_L, t_L)$	$LL_{FA}: (x_F, t_S; x_A + x_L, t_L)$		
Relocation effect	$SS_F: (x_R + x_S, t_S)$	$SS_B: (x_S, t_S; x_R, t_L)$	Less patience	Yes
	$LL_F: (x_R, t_S; x_L, t_L)$	$LL_B: (x_R + x_L, t_L)$		
Common-consequence effect ^a	$SS: (x_S, t_S)$	$SS_C: (x_S, t_S; x_C, t_C)$	More patience	Yes
	$LL: (x_L, t_L)$	$LL_C: (x_C, t_C; x_L, t_L)$		
Mere-token effect ^a	$SS: (x_S, t_S)$	$SS_M: (x_M, t_M; x_S, t_S)$	More patience	Yes
	$LL: (x_L, t_L)$	$LL_M: (x_M, t_M; x_L, t_L)$		
Amplification effect ^a	$SS_M: (x_M, t_M; x_S, t_S)$	$SS_{AM}: (x_A + x_M, t_M; x_S, t_S)$	More patience	Yes
	$LL_M: (x_M, t_M; x_L, t_L)$	$LL_{AM}: (x_A + x_M, t_M; x_L, t_L)$		
Violations of dominance	$SS: (x_S, t_S)$	$SS_D: (x_S, t_S; x_D, t_D)$	More patience	Yes
	$LL: (x_L, t_L)$	$LL_D: (x_L, t_L)$		

^aViolations of independence (Koopmans, 1960).

TABLE A1. Distribution (row-wise percentages) of those lending support to the four candidate models in the main contest across the five candidate models in the parallel contest, in Experiments 2, 3, and 4.

Main contest		Parallel contest																	
		Experiment 2						Experiment 3						Experiment 4					
		TM	DIUM	SMNZ	SMZ	RM	<i>n</i>	TM	DIUM	SMNZ	SMZ	RM	<i>n</i>	TM	DIUM	SMNZ	SMZ	RM	<i>n</i>
TM		212	0	0	0	100	312	334	0	0	1	163	498	64	0	0	0	27	91
		66%	0%	0%	0%	32%		67%	0%	0%	0%	33%		70%	0%	0%	0%	30%	
DIUM		0	62	0	0	18	80	0	102	0	3	32	137	0	61	0	1	19	81
		0%	77.5%	0%	0%	22.5%		0%	74.5%	0%	2%	23.5%		0%	75%	0%	1%	24%	
SMNZ		2	0	22	0	56	80	1	0	40	0	101	142	0	0	13	0	10	23
		2.5%	0%	27.5%	0%	70%		1%	0%	28%	0%	71%		0%	0%	56.5%	0%	43.5%	
SMZ		1	0	0	37	10	48	2	1	0	56	12	71	0	2	2	31	6	41
		2%	0%	0%	77%	21%		3%	1%	0%	79%	17%		0%	5%	5%	75.5%	14.5%	

TABLE A2. Proportions of Bayesian p-values lower than .05 among those lending anecdotal support to the respective models in the parallel contest.

Experiment	Model									
	Adjustment for dominance detection					No adjustment for dominance detection				
	TM	DIUM	SMNZ	SMZ	RM	TM	DIUM	SMNZ	SMZ	RM
2	0%	16.67%	7.14%	0%	0%	0%	9.09%	14.29%	3.70%	100%
3	0%	15.09%	3.85%	0%	0%	0%	11.76%	7.14%	9.68%	100%
4	0%	0%	0%	0%	0%	0%	0%	9.09%	4.00%	100%

Figure Captions

Figure 1. Screen shot from choice pair #11.

Figure 2. Proportions of participants (above bars) lending increasingly stronger support for candidate models in Experiment 2 (top panel), Experiment 3 (center panel), and Experiment 4 (bottom panel).

Figure 3. Observed proportions of participants choosing the high NPV option among those lending anecdotal support to the respective models (horizontal), plotted against the average predicted probabilities of choosing the high NPV option generated by the models (vertical), in Experiment 2.

Figure 3. Observed proportions of participants choosing the high NPV option among those lending anecdotal support to the respective models (horizontal), plotted against the average predicted probabilities of choosing the high NPV option generated by the models (vertical), in Experiment 3.

Figure 5. The effect of zero outcomes among those providing anecdotal support for the candidate models (top panel), and among those providing increasingly stronger support for the Tradeoff Model (bottom panel), in Experiment 3.

Figure 6. Observed proportions of participants choosing the high NPV option among those lending anecdotal support to the respective models (horizontal), plotted against the average predicted probabilities of choosing the high NPV option generated by the models (vertical), in Experiment 4.

Figure 7. Duration of *SS* and *LL* as a function of token magnitude. Duration is given in days, and computed with $u(x) = x^\gamma$, where $\gamma = 1$ (solid line), $.75$ (dashed line), and $.50$ (dotted line). Numbers along the ordinate are the duration under linear utility for the \$50, \$500, and \$5,000 tokens. The tokens, along the abscissa, are logarithmically spaced.

Figure A1. Posterior probabilities for the parallel model contest, averaged across participants, in Experiment 2 (top panel), Experiment 3 (center panel), and Experiment 4 (bottom panel).

FIG. 1

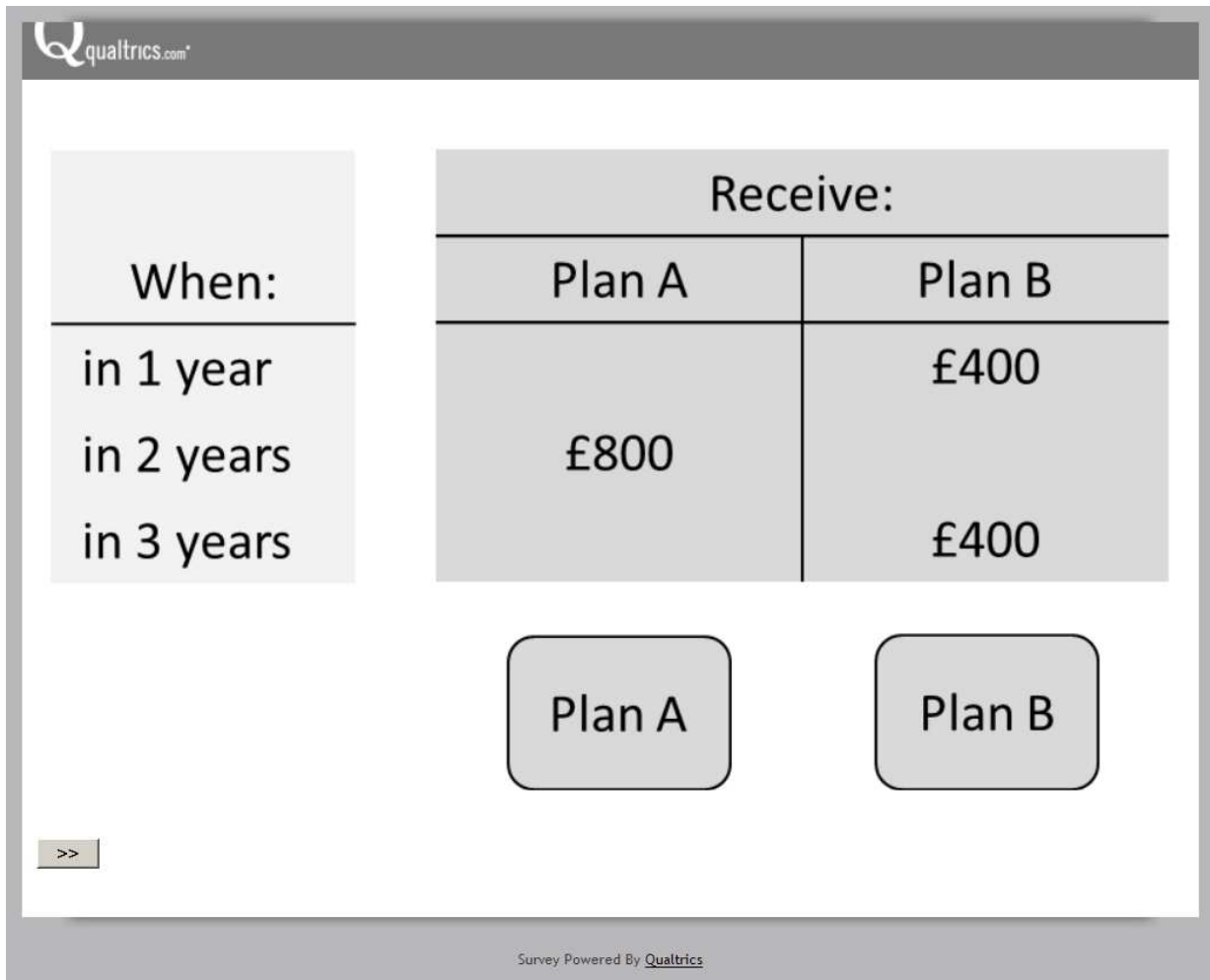


FIG. 2

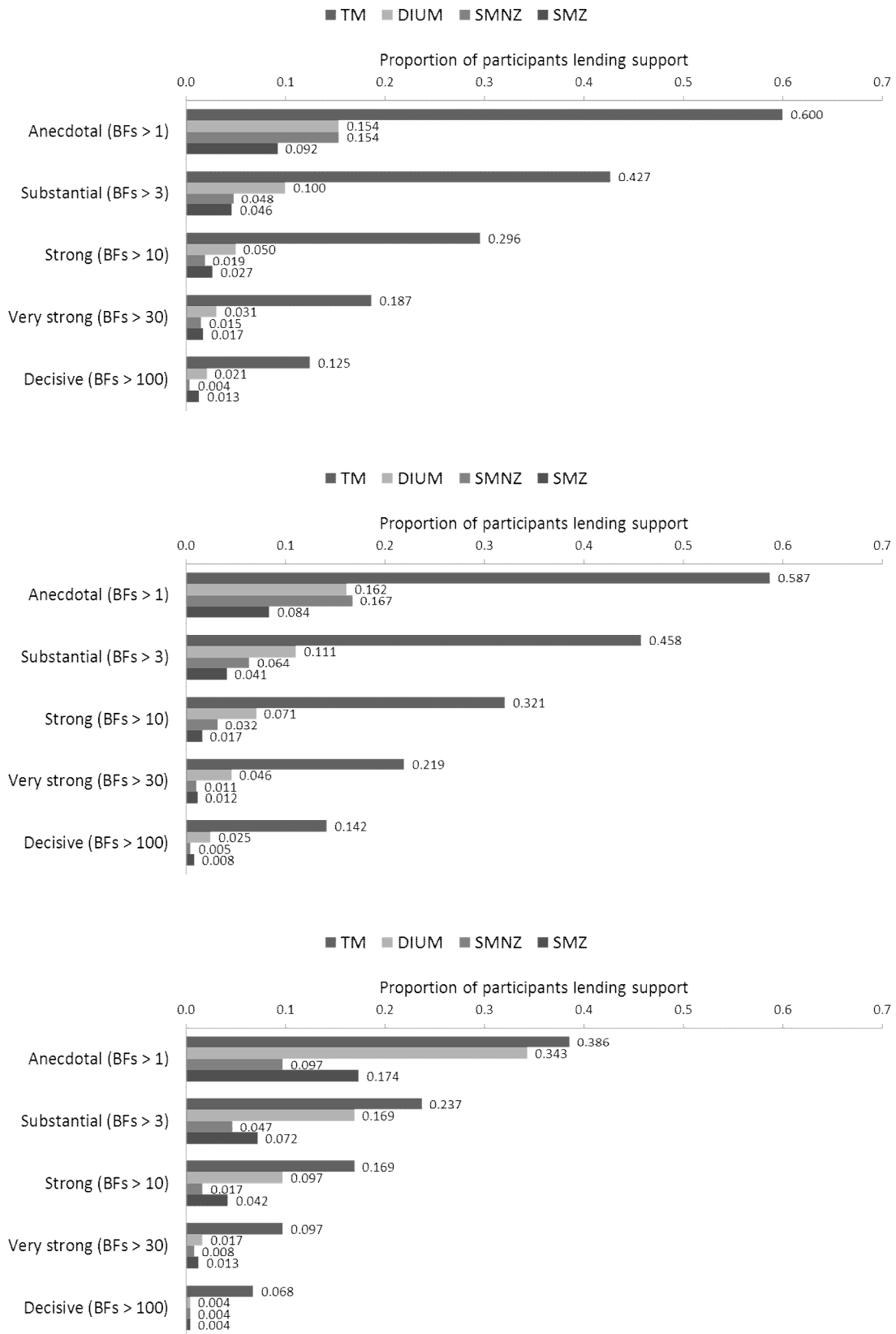
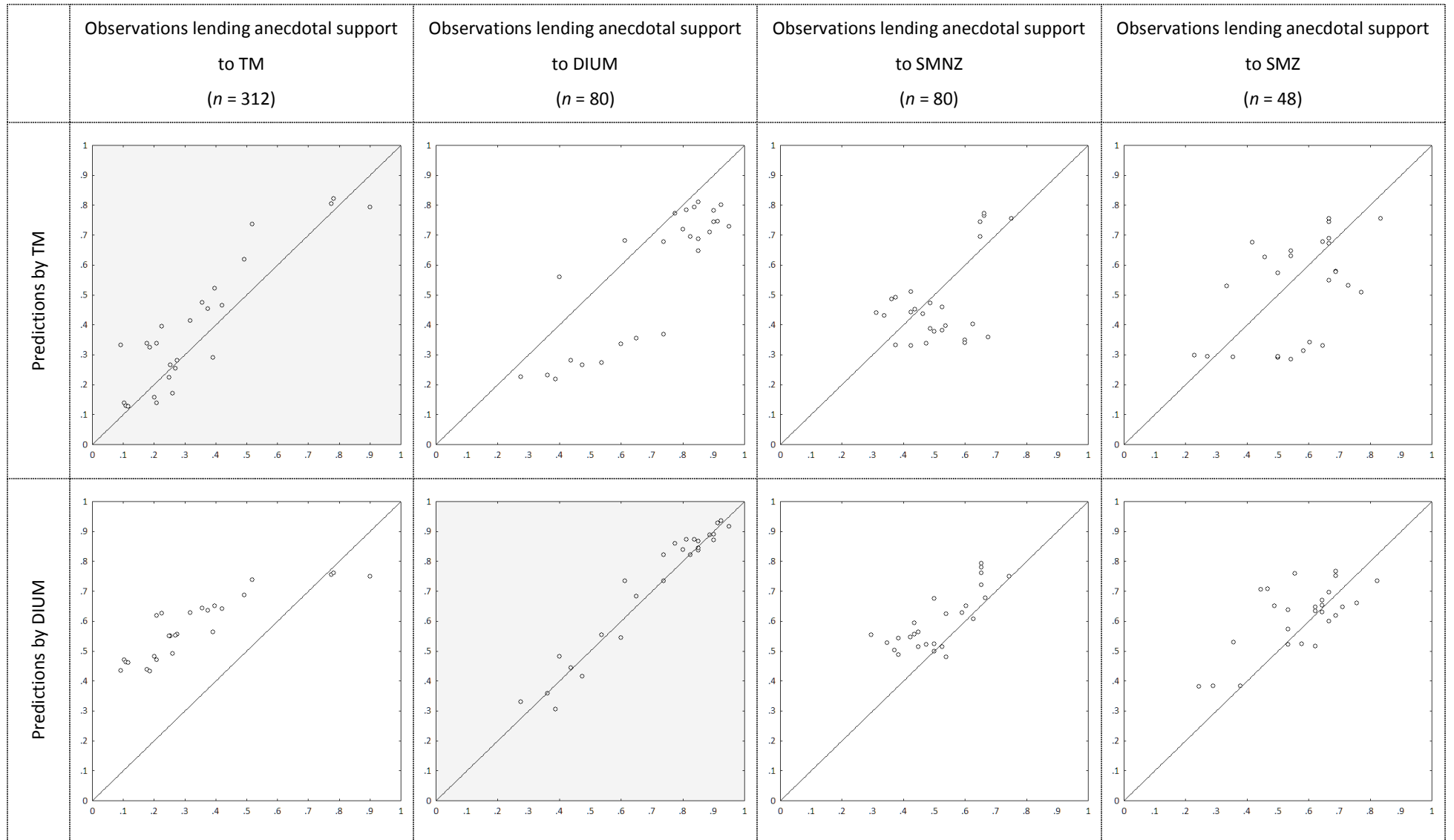


FIG. 3



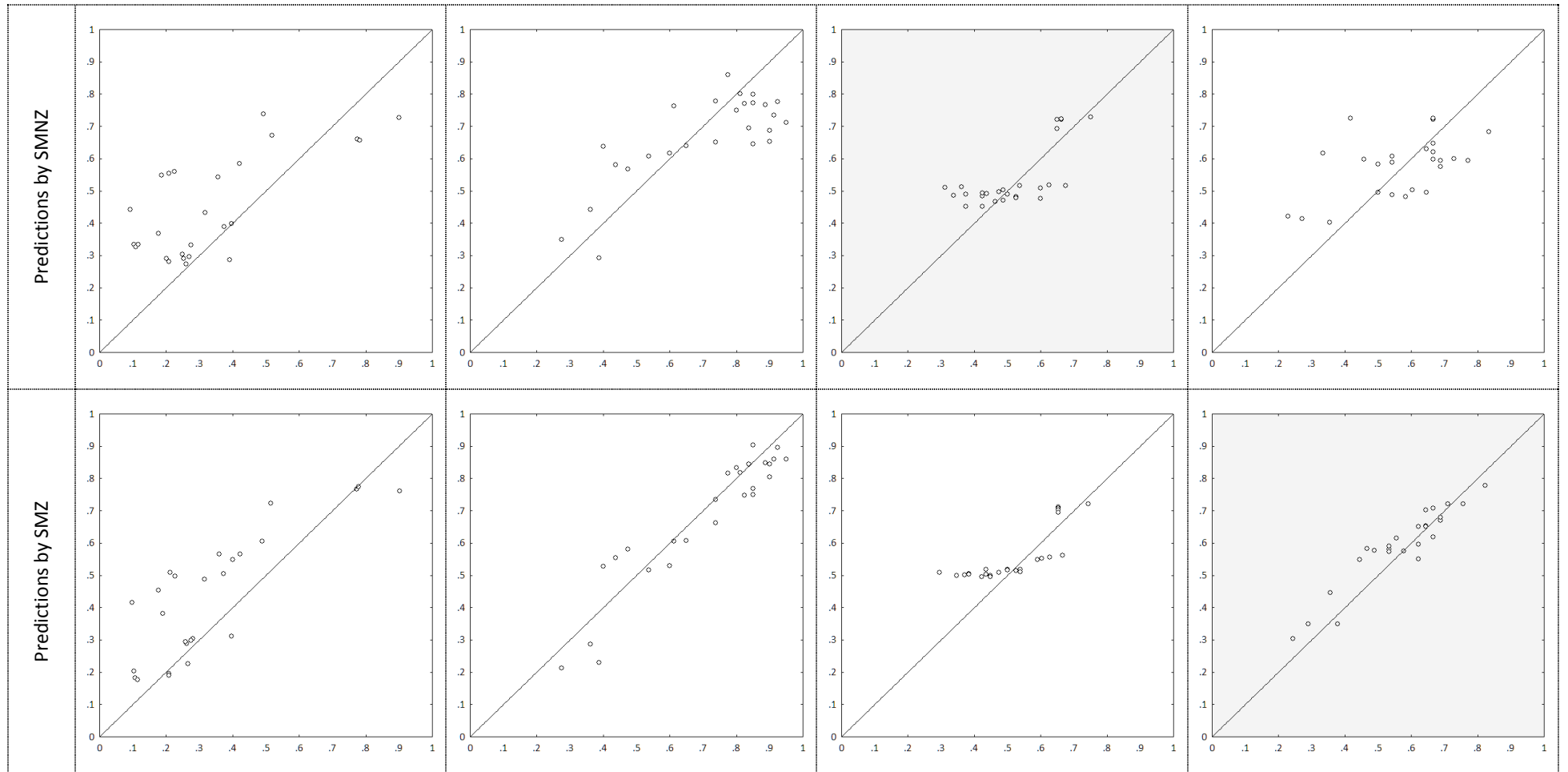
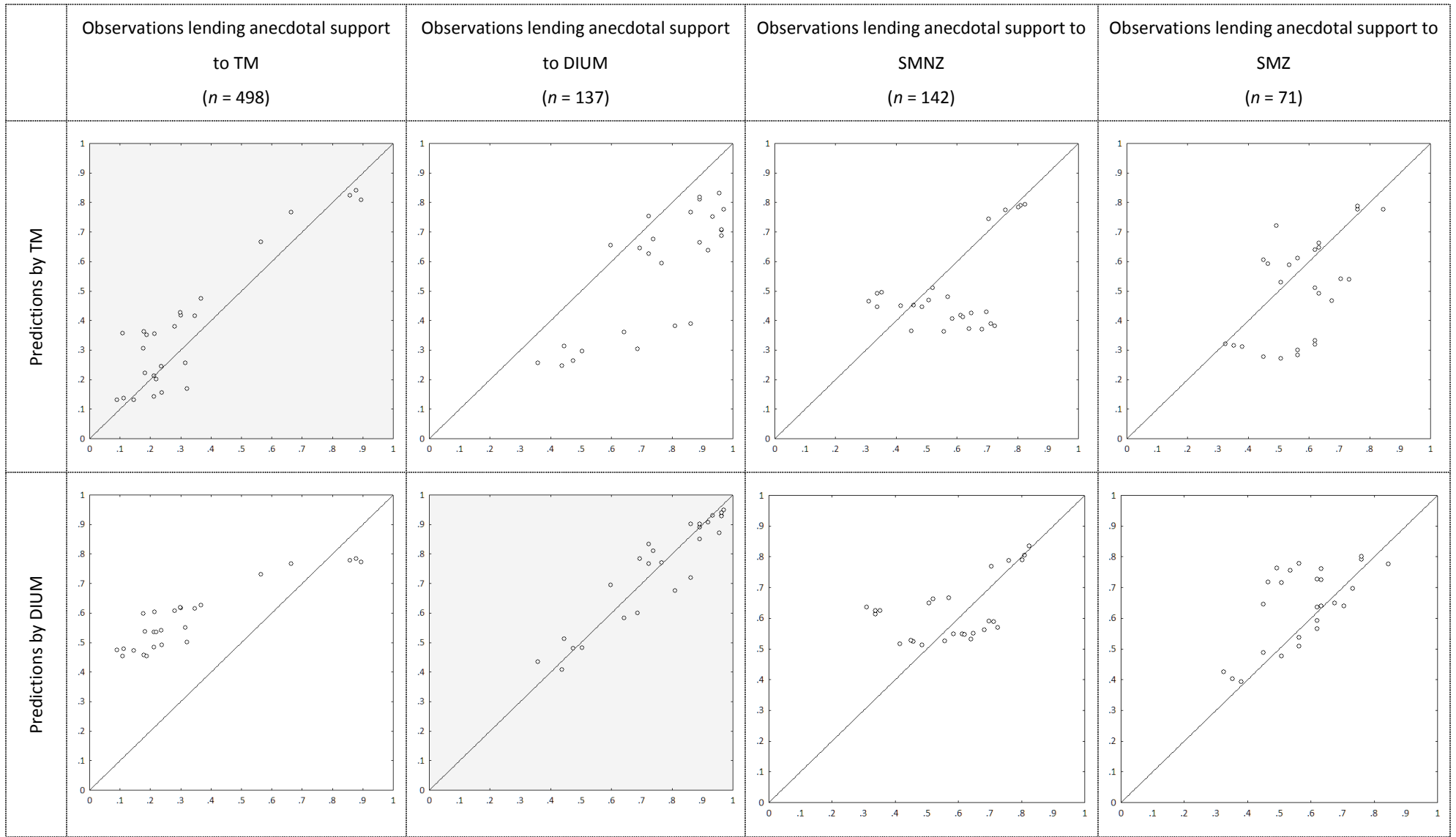


FIG. 4



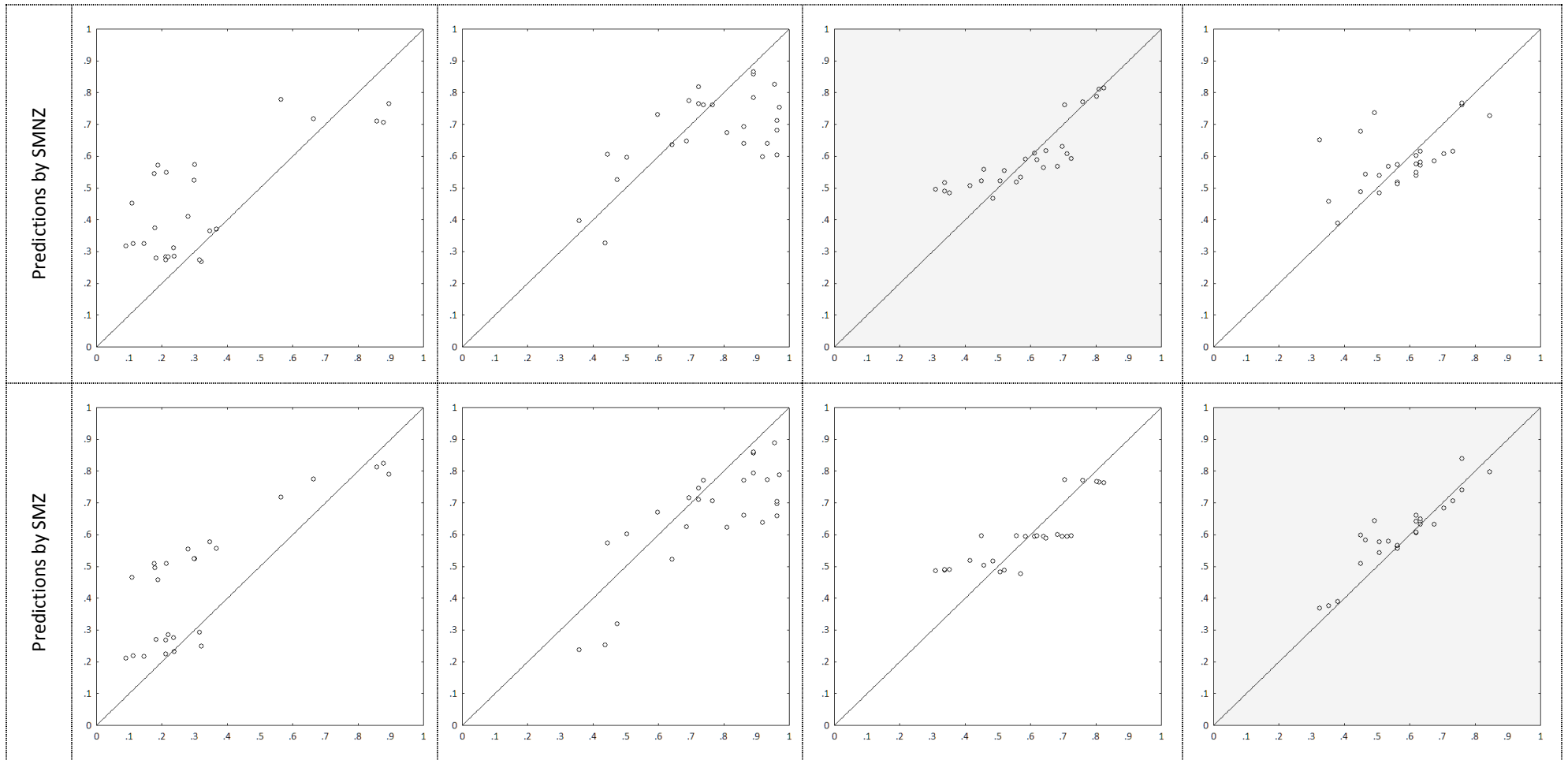


FIG. 5

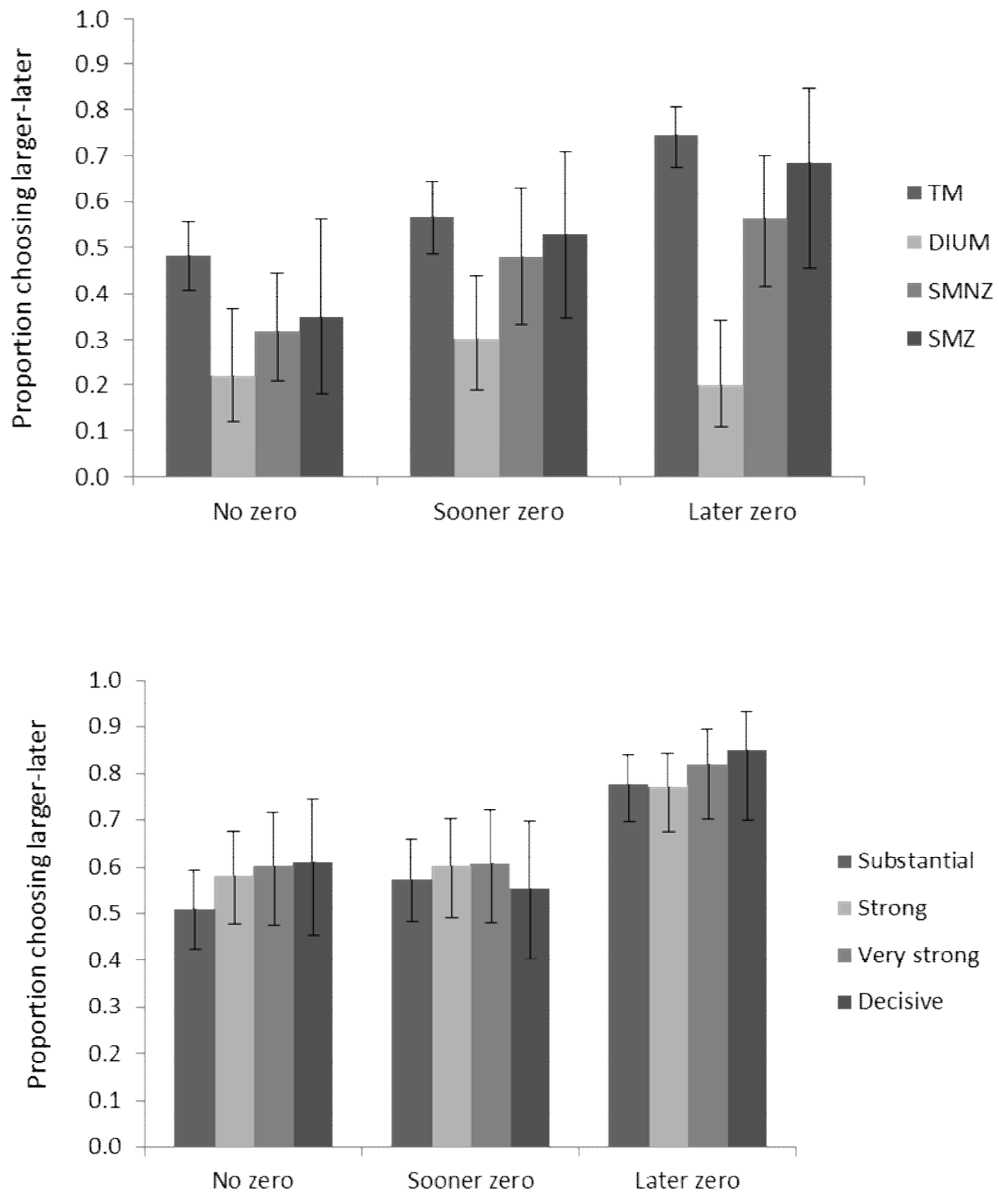
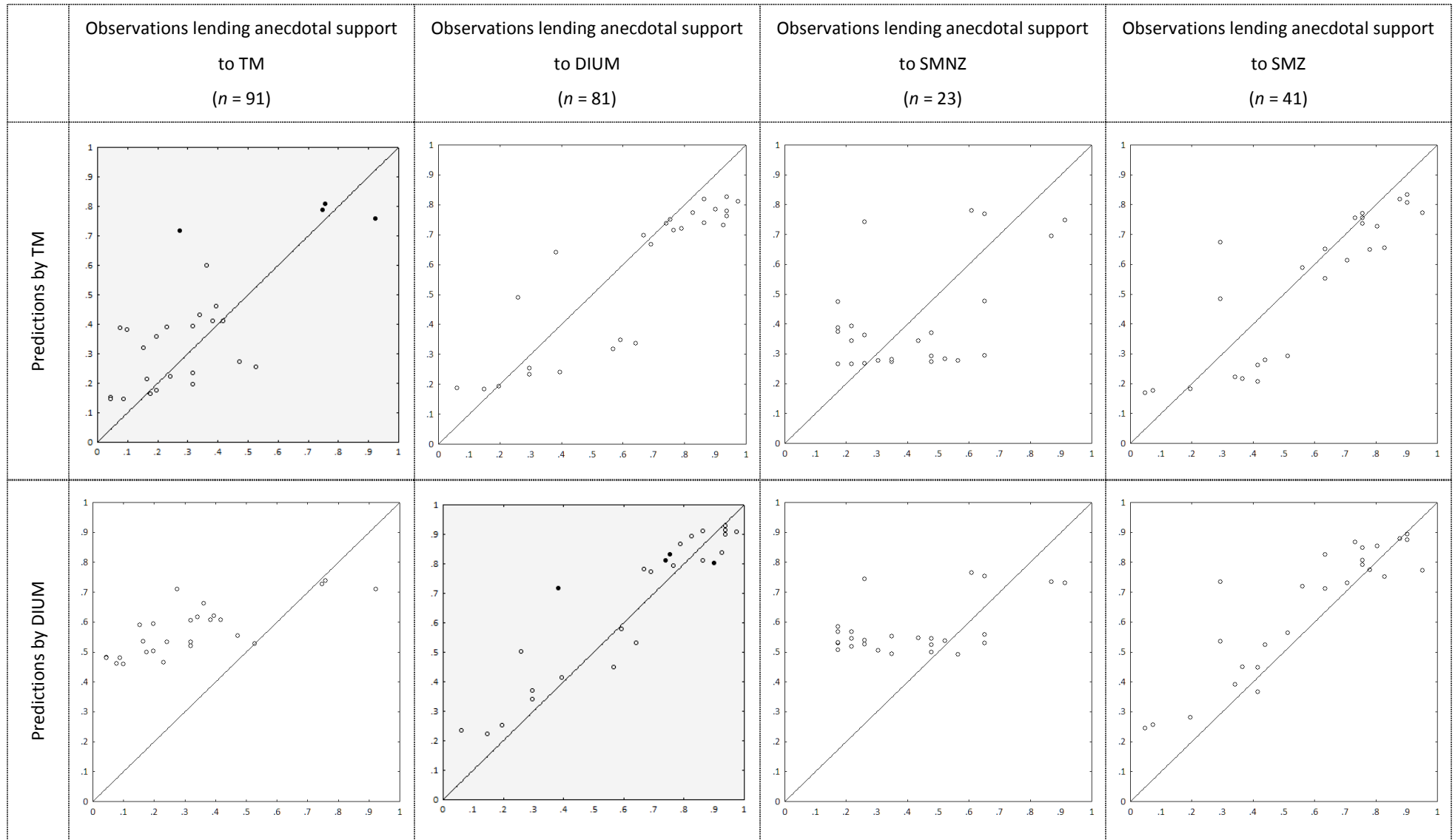


FIG. 6



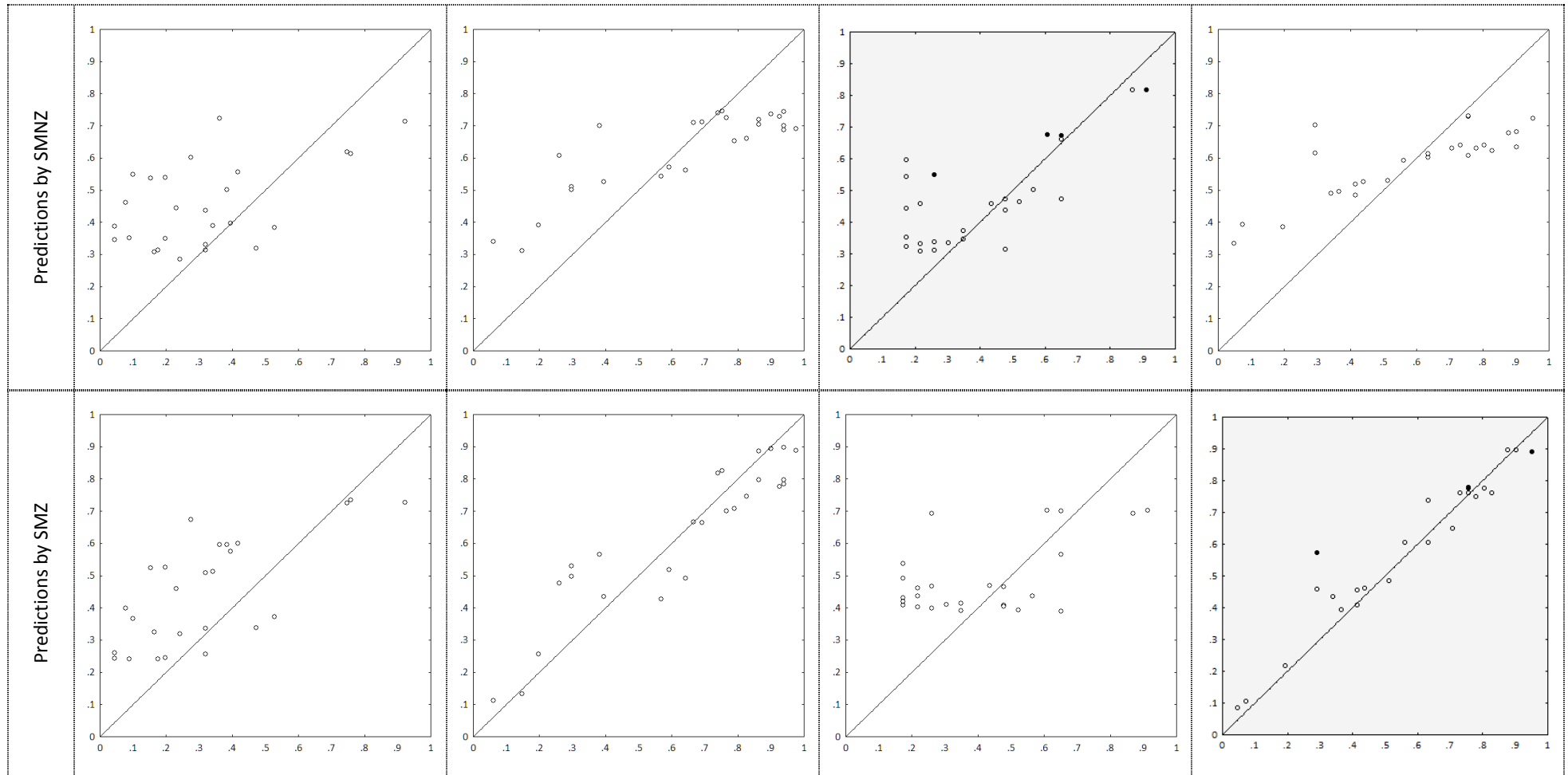


FIG. 7

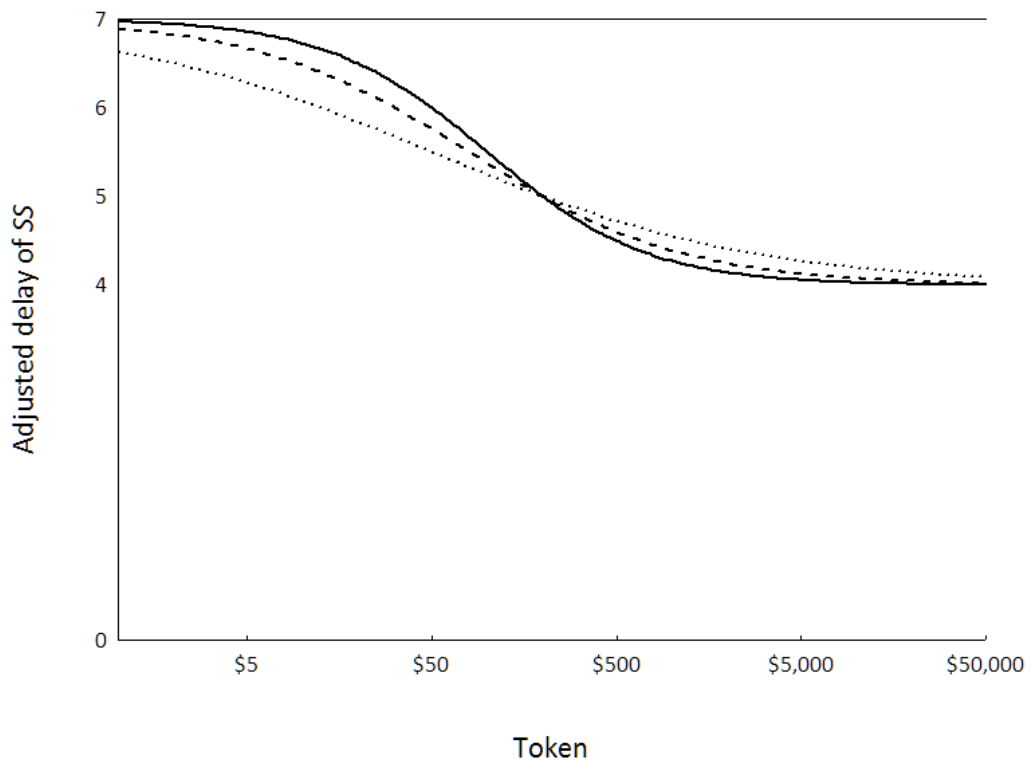
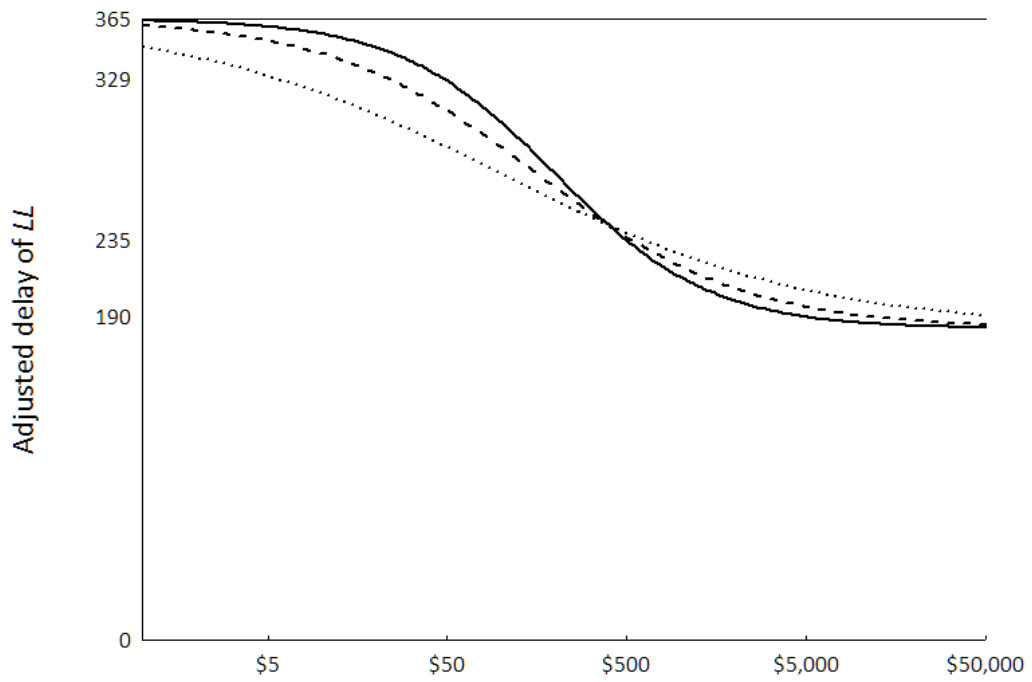
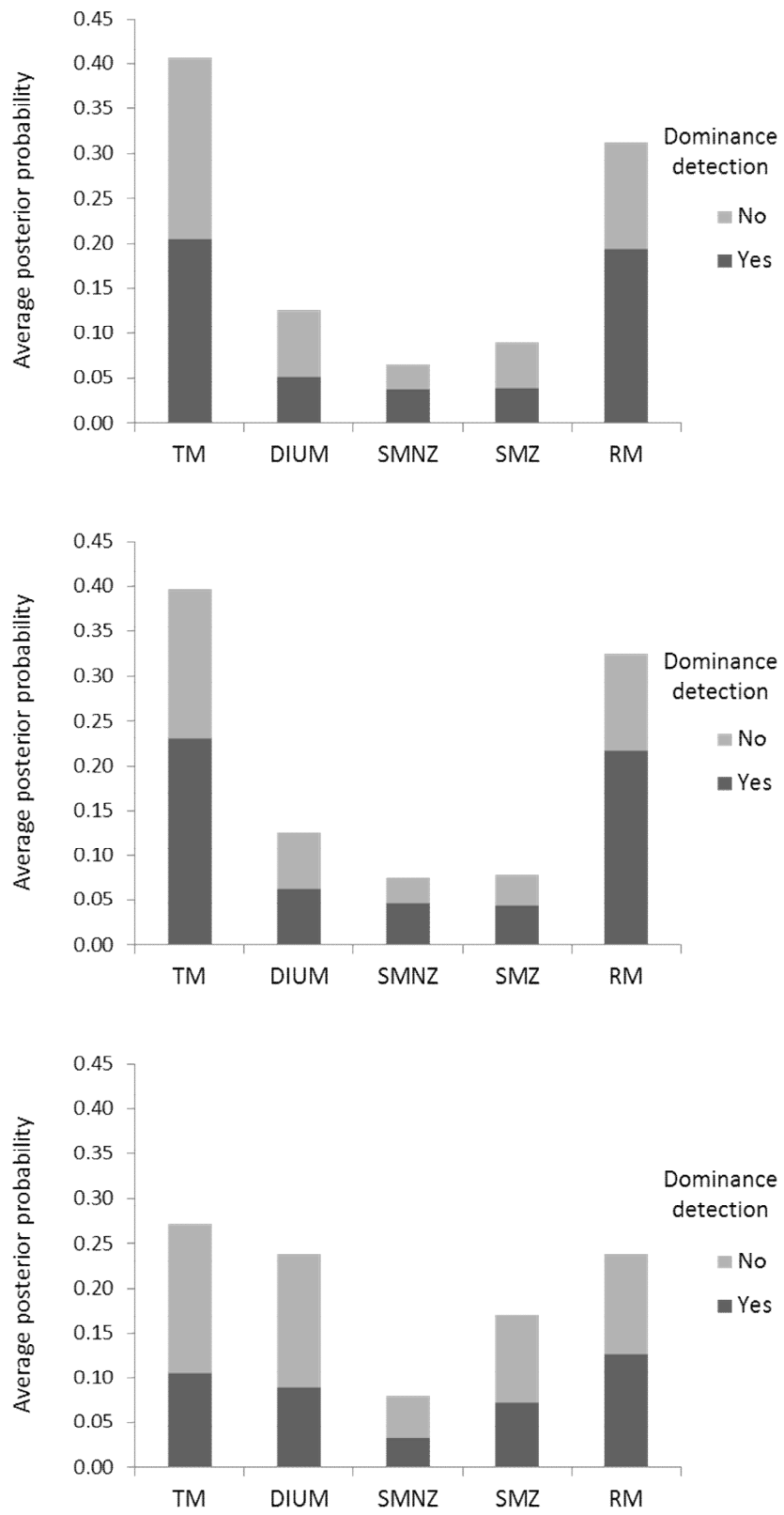


FIG. A1



Footnotes

¹The MTurk survey included several sections on intertemporal choice and risky choice, but the choice between the low and high NPV options was always the first section. The sample was 52% female, with an average age of 34. A majority (58%) were employed (46% full time), and 23% were student. A majority (62%) had an academic degree (47% bachelors, 13% masters, 2% PhD). The second survey included several sections on intertemporal choice, and the choice between the low and high NPV options was always the fourth section. The sample was 57% female, with an average of 52. A majority (54%) were employed (40% full time), and 3% were student. This sample was not as well educated as the first, with 42% having concluded high school, and 37% having concluded university (27% bachelors, 8% masters, 2% PhD).

²When outcomes are distributed over three contiguous time periods and outcome sequences afford the same total amount of money, as in our experiments, distaste for improvement in the sequences model is analogous to discounting in the discounted instantaneous utility model. To see the analogy, compare the sequences $\{x_1, x_2, x_3\}$ and $\{y_1, y_2, y_3\}$, both summing to c . In terms of monetary values, the difference between weighted improvement scores is $\beta(x_3 - x_1) - \beta(y_3 - y_1)$. At the same time, the difference between Net Present Values is $(\delta x_1 + \delta^2(c - x_1 - x_3) + \delta^3 x_3) - (\delta y_1 + \delta^2(c - y_1 - y_3) + \delta^3 y_3)$, or $((\delta - \delta^2)x_1 - (\delta^2 - \delta^3)x_3) - ((\delta - \delta^2)y_1 - (\delta^2 - \delta^3)y_3)$, which, given distaste for improvement, i.e., $\beta < 0$, is analogous to the difference between weighted improvement scores, but with discounting giving a greater weight to period 1 than to period 3.

³In Loewenstein and Prelec's (1993) statement of the model, the spreading scores are weighted $+\sigma$. We reversed this to $-\sigma$, so that the sign of σ indicates whether an individual has a positive ($\sigma > 0$) or negative ($\sigma < 0$) attitude toward spreading.

⁴Duration is also known in the world of finance, where D is the duration of a bond (Macaulay, 1938). In the *Macaulay duration*, a_j is given as:

$$a_j = \frac{x_j / (1+r)^{t_j}}{\sum_{i=1}^n x_i / (1+r)^{t_i}}.$$

In comparison with the preceding calculation of duration, the outcomes are replaced by the Net Present Values of the outcomes. Thus, when discounting the future ($r > 0$), an outcome loses impact with its distance from the present; if it does not, in the absence of discounting ($r = 0$), we obtain the weighted average again.

⁵A widely discussed topic in the literature on intertemporal choice is whether discounting is “exponential,” as in the discounted instantaneous utility model, or “hyperbolic.” The discount function of the discounted instantaneous utility model can be given as $d(t) = \delta^t = e^{-\kappa t}$, where $\kappa > 0$ is constant discounting. Loewenstein and Prelec (1992) generalized this into $d(t) = (1 + \tau t)^{-\kappa/\tau}$, where $\kappa > 0$ is discounting, and $\tau > 0$ is the departure from constant discounting. Rearranging the terms in Equation (3b), and reducing each outcome sequence to a single dated outcome (x, t) , the (arithmetically) “discounted utility” of each outcome would be $u(x) - \kappa w(t)$, where $\kappa > 0$ is “discounting.” When the time-weighing function is given as $w(t) = (1/\tau) \log(1 + \tau t)$, where $\tau > 0$ is diminishing sensitivity to delays (Scholten & Read, 2010), Loewenstein and Prelec’s (1992) generalized hyperbolic discount function follows from exponential discounting over weighted delays, i.e., $d(t) = e^{-\kappa w(t)}$, meaning that hyperbolic discounting is equivalent to diminishing sensitivity to delays (see also Scholten & Read, 2006). In our investigation, sequences span three future and contiguous time periods (i.e., 1, 2, and 3), so that any diminishing sensitivity to delays

is difficult to identify. We therefore let w be the identity function (i.e., $\tau \rightarrow 0$), which also ensures formal equivalence to the discounted instantaneous utility model.

⁶We used a power value function, because it is widely used in formal modeling, and because it is versatile in allowing for both diminishing sensitivity ($\gamma < 1$) and augmenting sensitivity ($\gamma > 1$) to outcomes. The tradeoff model, however, was originally proposed to have a logarithmic value function, analogous to the time-weighting function in Footnote 5, i.e., $v(x) = (1/\gamma)\log(1 + \gamma x)$, where $\gamma > 0$ is diminishing sensitivity to outcomes (Scholten & Read, 2010). Because our results strongly argue in favor of diminishing sensitivity, the logarithmic value function may be revisited in future modeling.

⁷We ran two additional analyses. One used uniform priors from 0 to 1 for γ , δ , and ϑ , the density function $1/(x + 1)^2$ for the priors on κ and ε , which were restricted to positive values, and the density function $0.5/(|x| + 1)^2$ for the priors on β and σ , which can take on negative as well as positive values. Another analysis used the same priors as in the main contest, except that γ was uniform from 0 to 1 rather than 0 to 5. These analyses yielded the same headline result as the main contest, which is that TM comes out as the winning model.

⁸Plots of posteriors of model parameters in the individual-level and group-level analyses are available as supplemental material at: [ADDRESS].

⁹For more discussion of similarity effects in intertemporal choice, see Leland (2002), Rubinstein (2003), Scholten and Read (2010), and Stevens (2016).

¹⁰In a choice-based matching study by Scholten, Read, and Sanborn (2014), participants demanded on average an interest rate of 6.7% per month, so one might expect that the typical participant will regard an interest rate of 1.9% to be low.