The Cost of Keeping Track

Johannes Haushofer*

First draft: August 9, 2014
This version: March 24, 2015

Abstract

People frequently decide between completing transactions (e.g. paying a bill, cashing a check) immediately or in the future. When they opt for the future, the transaction may fail: people may forget about it, or external factors may get in the way. I show that when such failures are associated with a lump-sum cost, this predicts several known departures from the standard discounting model: an agent who anticipates such costs will exhibit decreasing impatience with dynamic inconsistency; a magnitude effect, i.e. discounting large amounts less than small amounts; a reversal of this magnitude effect in the loss domain; a sign effect, i.e. discounting gains more than losses; and an Andreoni-Sprenger (2012) type reduction of discounting and decreasing impatience when money is added to existing payoffs. In addition, agents of this type may prefer to "pre-crastinate", i.e. incur losses sooner rather than later. The model further predicts a gain-loss asymmetry similar to that which leads to loss aversion. The results do not depend on whether the cost arises for internal (e.g. forgetting) or external reasons (e.g. "acts of God"), and whether the cost of keeping track is incurred every period (e.g. a psychological hassle cost) or once (e.g. as a result of forgetting combined with a penalty). Naïveté causes agents to behave more like standard exponential discounters in most respects, except that it leads to a welfare loss when agents choose delayed outcomes because they underestimate their cost. Sophisticated agents are willing to pay for reminders. Empirical tests conducted in Nairobi, Kenya confirm that people are likely to fail to act on profitable future opportunities, and that they are partly aware of this fact, preferring to pre-crastinate on losses and being willing to pay for reminders. I discuss applications in development economics, such as low takeup of savings technology, water chlorination, vaccination, fertilizer, and insurance.

JEL codes: D9, D11, D60, D91, E21, D03, D81, C9

Keywords: intertemporal choice, temporal discounting, time preference, memory, limited attention

*Department of Psychology, Department of Economics, and Woodrow Wilson School of Public and International Affairs, Princeton University, Princeton, NJ. I thank Sam Asher, Abhijit Banerjee, Roland Bénabou, Seth Blumberg, Stefano Caria, Janet Currie, Alain de Janvry, Marcel Fafchamps, Marc Fleurbaey, Ben Golub, Pam Jakiela, Chaning Jang, Anett John, Shachar Kariv, Ilyana Kuziemko, Jeremy Magruder, Ted Miguel, Paul Novosad, Simon Quinn, Christopher Roth, Martha Ross, Betty Sadoulet, Josh Schwartzstein, Dan Sacks, Jeremy Shapiro, Charlie Sprenger, Tom Vogl, and seminar and conference participants at Berkeley, Princeton, Oxford (CSAE), NEUDC, and ideas42 for comments. I am further grateful to Lara Fleischer, Allan Hsiao, Chaning Jang, Lucy Rimmington, and James Vancel for excellent research assistance. This research was supported by NIH R01AG039297.

1 Introduction

Individuals often have to keep track of future transactions. For instance, when a person makes a decision to pay a bill not now, but later, she has to keep this plan in mind and perform an action later to implement the decision (e.g. logging into her online bank account and making the transfer). If she fails to make the transfer, she may face a late fee. Similarly, when she expects a payment, she is likely to keep track of the incoming payment by verifying whether it arrived in her bank account. If she fails to keep track of the incoming payment, it may get lost and she may have to pay a hassle cost to follow up on it.

Similarly, future transactions may fail for reasons external to the individual. In the above examples, paying the bill may fail because the bank's website is down (and of course it's Sunday so the branch is closed); an incoming payment may fail because of a computer error. In both cases, costs arise for the agent, who may again face a fee for late payment of the bill, and a hassle cost for following up on the incoming payment.

This paper presents a simple model of intertemporal choice whose main assumption is that such "keeping track" generates costs for the agent. This cost can come in several forms: agents may forget about the future transaction, resulting in a penalty such as a late fee, or a hassle cost to salvage the transaction. Additionally or alternatively, simply having to keep the task in mind may generate a psychological cost. The cost may be avoidable through reminders, but these in turn may be costly to set up. In either case, the agent integrates into her decision today the cost of keeping track of future transactions. In most of what follows, I model this cost as a one-time, lump sum cost that is substracted from any future payoff, both in the gain and in the loss domain.[1] For instance, receiving a check worth \$100 today may be worth 100 utils; receiving it two weeks from now may be worth 90 utils after standard exponential discounting, but may additionally incur a cost of keeping track of 20 utils, resulting in an expected utility of 70. Analogously, the disutility from paying a \$100 credit card bill today may be -100 utils; paying it in two weeks (by the deadline) may have a disutility of only -90 due to standard exponential discounting, but may also incur a cost of keeping track of 20 utils, resulting in an expected utility of -110. Everything else in the model is standard: agents have perfect foresight, are sophisticated, and discount the future exponentially, i.e. have time-consistent preferences.

This simple addition to the standard model turns out to predict several stylized facts of temporal discounting that the literature has documented. First, in standard laboratory experiments where discount rates are estimated by asking individuals to decide between an amount of money available now and a larger amount available later, people often exhibit very high discount rates for gains, which are difficult to reconcile with commonly prevailing interest rates (Frederick et al., 2002). The model predicts such steep discounting for future outcomes because they are discounted both by a discount factor *and* by the cost of keeping track. In the extreme, the cost of keeping track may

[1] Later sections of the paper discuss how the results change for different formulations of the cost; briefly, most results hold when the cost is proportional instead of lump-sum, and when it is paid every period instead of once.

even result in future gains having *negative* utility, causing agents to prefer forgoing future gains; for instance, agents who face a high cost of keeping track of future gains may elect not to take up free trial periods (because they risk having to subscribe to the service if they fail to cancel in time), or mail-in rebates (because the cost of keeping track of the transaction is larger than its expected value).

Second, people often discount losses less than interest rates would predict; this fact is predicted by the model because future losses, too, generate a cost of keeping track, which makes them less desirable than they otherwise would be, leading to *less* discounting of future losses.

Third, this decrease in discounting of losses can lead to a reversal of discounting in the loss domain: if the cost of keeping track is large enough, agents who would otherwise prefer delayed to immediate losses may now prefer to incur losses sooner rather than later. I will refer to this phenomenon as *pre-crastination*. It captures the familiar feeling of wanting to "get it over with", and has recently been demonstrated empirically in the effort domain: Rosenbaum et al. (2014) asked participants to carry one of two buckets down an alley, and found that a large proportion of participants preferred to pick up the bucket sooner rather than later, even though this meant that they had to carry it a greater distance. Rosenbaum et al. attribute this preference to a desire to reduce working memory load, an interpretation that is in line with the suggestion that keeping track of the goal is costly. Existing models of temporal discounting currently do not account well for such behaviors. O'Donoghue & Rabin (1999) show that the well-known quasi-hyperbolic discounting model (Strotz, 1956; Laibson, 1997) predicts "pre-properation" under certain conditions. However, this occurs only when costs are increasing over time; this is precisely the core assumption of the present model, which produces pre-crastination even without present bias. Note that I consciously use a different term for the phenomenon: pre-properation is defined as doing things *too* soon, with a time-consistent individual as the benchmark. In my model, agents *are* time-consistent; their preference for sooner over later losses stems simply from correctly taking all associated costs into account, and therefore they cannot be said to act *too* soon.

Fourth, beacuse a cost of keeping track leads to increased discounting in the gains domain and decreased discounting in the loss domain, agents who face a cost of keeping track will, all else equal, discount gains more than losses. This phenomenon is a frequently documented empirical regularity about discounting, and is commonly known as the sign effect (Thaler, 1981; Loewenstein, 1987, 1988; Benizon et al., 1989).

Fifth, as a consequence of increased discounting of gains and decreased discounting of losses, agents who face a cost of keeping track will exhibit an asymmetry in discounting similar to that observed in loss aversion: future losses are more painful than future gains are pleasurable. Thus, a corollary of the model is that it predicts atemporal loss aversion, and in particular, that loss aversion for future outcomes should be more pronounced than that for immediate outcomes.[2]

Sixth, depending on the specific nature of the cost of keeping track, the model predicts that large

[2] I am not familiar with empirical tests of this hypothesis.

gains will be discounted less than small gains. In particular, a lump-sum cost of keeping track associated with any future transaction is proportionally smaller for large compared to small amounts, and thus the additional discounting that arises form the cost of keeping track is proportionally greater for small gains than or large gains. This phenomenon is commonly known as the magnitude effect, and has been extensively documented in empirical studies (Thaler, 1981; Loewenstein, 1987; Benzion et al., 1989). It should be noted that a magnitude effect can also be produced with concavity of the utility function (Loewenstein & Prelec, 1992). However, as Noor (2011) points out, the curvature of the utility function that would be required to account for the magnitude of this effect that is typically observed empirically is extreme, and thus alternative models are needed to account for this effect. The present model suggests one plausible source for the magnitude effect.

Seventh, Hardisty (2011) showed recently that the magnitude effect often does not exist in the loss domain, or is in fact reversed, with *more* discounting of large compared to small losses. It turns out that the cost of keeping track model also predicts this somewhat obscure finding. Recall from above that the cost of keeping track is subtracted from the expected utility of both gains and losses, and thus increases the disutility that arises from future losses; in other words, it *decreases* discounting for losses. If the cost is lump-sum, this decrease in discounting is proportionally larger for small losses; as a result, small losses are discounted *less* than large losses.

Eighth, probably the most frequently studied empirical fact about temporal discounting is decreasing impatience, i.e. higher discount rates in the near future compared to the distant future.[3] The reason why this particular feature of discounting has received so much attention is that it predicts dynamic inconsistency: individuals with decreasing impatience may prefer the larger, later payoff over a smaller, sooner payoff when both payments are in the distant future, but may change their mind and instead opt for the smaller, sooner payoff as the payments approach. The consequence of this prediction is that individuals do not follow through on plans made in the past when they face an opportunity to reconsider them; in the gains domain, this is frequently referred to as *impulsivity;* in the loss domain, as *procrastination*. The present model generates a discount function that exhibits decreasing impatience for gains by imposing a cost on any future outcomes, but not present outcomes. As such, it is similar in spirit to the quasi-hyperbolic model (Laibson, 1997; Strotz, 1956), which also distinguishes between the present and all future periods. However, it generates different predictions in the loss domain as described above; in particular, the model predicts *increasing* impatience in the loss domain. If the cost of keeping track is large enough, this leads to pre-crastination on losses, i.e. preferring to incur losses sooner rather than later. Note that the model therefore does not predict procrastination on losses.

Ninth, it has recently emerged that people discount the future at much lower rates when money

[3]Decreasing impatience is commonly referred to as hyperbolic discounting, myopia, or present bias. I use the term decreasing impatience here because the other terms can be formulated as special cases of decreasing impatience (e.g., hyperbolic discounting implies higher discount rates in the near future than in the distant future, but additionally imposes a specific functional form; similarly, present bias implies that the present receives particular weight, but this focus on the present is not necessary to obtain the result most commonly associated with present bias, i.e. dynamic inconsistency).

is added to existing payoffs. Andreoni & Sprenger (2012) estimate time preferences by asking individuals to make convex allocations between two timepoints. Crucially, these allocations were added to two "thank-you" payments of $5 each, of which one was delivered at the sooner and one at the later of the timepoints in the experiment. Andreoni & Sprenger find much lower discount rates using this method than other studies which use the traditional "multiple price list" approach. This is encouraging because the rates estimated from such "money now vs. later" experiments are often so high that they are hard to reconcile with prevailing interest rates. At the same time, extant models of discounting have trouble predicting such different behavior depending on the method used to elicit discount rates. The present model offers an explanation for the difference in estimated discount rates with the convex budget method relative to others: the claim of the model is that observed discount rates in the gains domain arise from standard exponential discounting and a cost of keeping track. In Andreoni & Sprenger's experiment, by the time subjects make the allocation to the sooner vs. later timepoint, they have already been told that they will receive a "thank-you" payment at both timepoints. As a result, they already know that they will pay a cost of keeping track at each of the timepoints in question. If this cost is a lump sum, it does not change when money is added to the thank-you payment at one or the other timepoint. Thus, any money added to the existing "thank-you" payments at all timepoints is discounted only with the standard discount factor and does not incur an additional cost of keeping track. The model therefore offers an intuitive explanation for the finding that adding money to existing payoffs results in a lower observed discount rate.

Finally, using the same method, Andreoni & Sprenger were surprised to find that subjects exhibit less decreasing impatience (which predicts dynamic inconsistency) than with standard experimental protocols; in fact, they classify most of their subjects as dynamically consistent. This prediction also falls naturally out of a cost of keeping track model when there are pre-existing payoffs at all timepoints: when one period is immediate and the other in the future, a pre-existing payoff at both timepoints ensures that the decision to allocate additional payoffs to either the immediate or the future timepoint is governed only by standard discounting, because the cost of keeping track of the future payment is already sunk in the "thank-you" payment (cf. above). But importantly, this is true *a forteriori* when both timepoints are in the future, because again the cost of keeping track of the payments at both timepoints is already sunk by the time the individual decides to which timepoint to allocate additional payments. As a consequence, when agents already anticipate to pay a cost of keeping track for existing payoffs, adding money to these payoffs is governed only by standard discounting regardless of time horizon, and therefore agents will exhibit no decreasing impatience with a cost of keeping track.

Thus, the simple additional assumption that any future transaction – both for gains and losses – carries a cost of keeping track predicts a number of the stylized empirical facts that the literature on temporal discounting has established. It is interesting to ask how the model relates to standard accounts of discounting such as the quasi-hyperbolic model (Strotz, 1956; Laibson, 1997). In the gains domain, the two models produce some similar results: the quasi-hyperbolic model predicts discounting at a rate higher than the interest rate if $\delta = \frac{1}{1+r}$ (where δ is the exponential discount factor in the quasi-hyperbolic model and r is the interest rate) by additionally discounting all future outcomes by β. Standard assumptions about the utility function, in particular concavity, can also produce the magnitude effect in the quasi-hyperbolic model (as well as the standard exponential model). However, neither the quasi-hyperbolic nor the exponential model predict less discounting of losses compared to the interest rate or compared to gains; pre-crastination; the sign effect; a gain-loss asymmetry; or the Andreoni-Sprenger results of reduced discounting and decreasing impatience when money is added to future payoffs. These models also do not predict a willingness to pay for reminders without additional assumptions about imperfect memory (e.g. Ericson, 2014).

A further point of divergence between the present model and quasi-hyperbolic discounting is that the latter predicts procrastination in the loss domain, whereas a cost of keeping track model does not make this prediction. In other words, putting off costs until a later time is only captured by the cost of keeping track model if it is combined with quasi-hyperbolic preferences. In this case, agents procrastinate when β is small enough, and pre-crastinate when c is large enough. I specify the exact condition for this divergence in Section 2.

Finally, a significant difference between the present model and the quasi-hyperbolic model is that the latter is defined over *consumption*; in contrast, the cost of keeping track model insists only that the *cost* be incurred in a particular period, while the transfers themselves are fungible and agents can borrow and lend against them. As a consequence, the quasi-hyperbolic model predicts that agents should exhibit dynamic inconsistency only in the consumption and not in the money domain; in contrast, the cost of keeping track model suggests that agents exhibit dynamic inconsistency even for monetary payoffs, to the extent that they are associated with a cost of keeping track. In support of the former view, recent studies have demonstrated dynamic inconsistency for effort and consumption decisions (Augenblick et al., 2013; Sadoff et al., 2014), while not finding evidence of inconsistency for tradeoffs between money (Augenblick et al., 2013). However, I note that Augenblick finds no evidence of dynamic inconsistency in the money domain in an experimental setup in which individuals choose between adding money to existing payoffs at different timepoints; under these conditions, the cost of keeping track model predicts no dynamic inconsistency, since the cost of keeping track is already sunk by the time agents consider the intertemporal tradeoff. It remains to be elucidated whether dynamic inconsistency exists in the money domain when transaction costs are kept constant, but the cost of keeping track is not zero or sunk.

Thus, the cost of keeping track model predicts a number of empirical regularities that are not well accounted for by the quasi-hyperbolic model. However, augmentations of such models do somewhat better at accounting for these stylized facts. In particular, Loewenstein & Prelec (1992) show that combining hyperbolic discounting with a value function that exhibits loss aversion, a reference point, and concavity can predict a larger number of the stylized facts that the present model seeks to account for; in particular, decreasing impatience, the magnitude effect, and gain-loss asymmetry (in addition, their model explains the delayed-speedup asymmetry in discounting that has been described by Loewenstein (1988). However, their model requires a relatively large number of non-standard assumptions because it combines non-standard discounting with non-standard preferences; in addition, it leaves several other empirical regularities unexplained (e.g. precrastination, reversed magnitude effect in the loss domain, reduced decreasing impatience and dynamic inconsistency when money is added to existing payoffs).

The particular form of the cost of keeping track matters for some of these results, but not for others. In Section 4, I consider four different formulations of the cost of keeping track: a one-time lump-sum cost; a per-period lump-sum cost; a one-time proportional cost; and a per-period proportional cost. The intuition behind the choice of these particular cost structures is as follows. First, a lump-sum, one-time cost of keeping track might arise in the gains domain when individuals forget to act on an incoming payment (e.g. a check) and as a result face a fixed hassle or time cost to salvage the transaction. In the loss domain, lump-sum, one-time costs may be the result of communications providers or banks imposing lump-sum penalties for late bill payment. A per-period lump-sum cost might consist of the mental effort to keep the upcoming task, whether gain or loss, in mind over a period of time. A one-time proportional cost might consist simply of the probability of forgoing an incoming payment altogether by forgetting about it, or by having to pay a proportional penalty for forgetting about a payment. Finally, a per-period proportional cost for might consist of interest forgone (gains) or interest to be paid (losses).

Section 4 presents a more general version of the model, in which the utility function has a more flexible shape (monotonic and concave), the probability of remembering a task over time is exponential, and the cost of keeping track allows for the four components described above. I find that all results hold when the cost of keeping track is a lump-sum, regardless of whether it is paid once or every period. The same is true when the cost is proportional, except that some results hold only for particular parameter values, and the reduction in decreasing impatience when money is added to existing payoffs (Andreoni & Sprenger, 2012) does not hold in general. Thus, the model predicts empirically observed regularities in discounting behavior best when there is a lump-sum element in the cost of keeping track, although many of the results hold even with a proportional cost.

The basic exposition of the model assumes that agents do not have a reminder technology at their disposal. In Section 3, I extend the results to include the availability of a reminder technology. When reminders for future transactions are available, agents are willing to pay for them up to the discounted cost of keeping track of the future transaction in question. The basic results of the model described above hold when reminders are available, with the exception of the results on decreasing impatience and dynamic inconsistency, and the decrease in discounting and decreasing impatience when money is added to existing payoffs. It is easy to see why agents will not exhibit dynamic inconsistency after buying a reminder. As an example, consider an agent who decides in period 0 to choose a larger payoff in period 2 over a smaller payoff in period 1. If the agent has an opportunity to reconsider her decision in period 1, then the payoff in period 1 is no longer subject to the cost of keeping track, while the payoff in period 2 still incurs a cost of keeping track. As a result, the payoff in period 1 is relatively more attractive, which may cause the agent to reverse her previous decision in favor of the payoff in period 1. In contrast, assume again that the agent in period 0 opted for the payoff in period 2, but bought a reminder for it. If she now reconsiders her decision in period 1, the payoff in period 2 is no longer subject to a cost of keeping track, because that cost was sunk in the reminder in period 0. Thus, the agent has no motive to reverse her previous decision in favor of the payoff in period 2; the reminder acts as a commitment device. When reminders are available, agents may therefore not exhibit dynamic inconsistency.

A similar argument illustrates why agents with a reminder technology at their disposal will not discount less than otherwise, and will not exhibit less decreasing impatience than otherwise, when money is added to existing payoffs. If agents buy reminders in period 0, then the cost of keeping track of the future transactions in question is sunk, and therefore the only factor agents take into account when trading off different future payoffs is standard discounting, both when money is added to existing payoffs and when it is not. Thus, the two cases do not differ, and agents therefore do not discount more when money is added to existing payoffs. The argument in the preceding paragraph illustrates the result on decreasing impatience.

The basic formulation of the model also assumes that agents are sophisticated, i.e. they have correct beliefs about their own cost of keeping track (or probability of forgetting). Section 3 relaxes this assumption and allows for the possiblity that agents underestimate their own probability of forgetting about future transactions. If this is the case, agents will behave like time-consistent exponential discounters: they discount gains and losses equally and with the standard exponential discount rate, do not show a sign effect, pre-crastination, or a magnitude effect (unless concavity is assumed). They also do not show decreasing impatience or dynamic inconsistency. However, they incur a welfare loss, because they will choose delayed outcomes without appreciating the cost associated with those outcomes.

Probably the most policy-relevant contribution of the model is that it provides an account for why individuals may fail to adopt profitable technologies. Lack of demand for profitable technologies is a common phenomenon especially in developing countries. The model suggests one possible reason for this phenomenon: when people face the decision of adopting technology, they often cannot act on it immediately, but instead have to make a plan to do it later. For instance, when people fetch water at a source, they might be reminded that they want to chlorinate their water. However, when opportunities to chlorinate are not available at the source, they have to make a plan to do it later, e.g. when they reach their home where they store their chlorine bottle. However, at the later timepoint when they can act on their plan, they may have forgotten about it – in this example, by the time they reach the homestead, they do not remember to use the chlorine bottle. The model captures this phenomenon, and predicts that when reminders are provided at the right time – e.g., at the water source – adoption should be high. Indeed, Kremer et al. (2009) show that dispensers at the water source can dramatically increase takeup. Other authors have documented similar failures to adopt technology, and shown that reminders at the right time – i.e., when people can act on them – can increase takeup: timed discounts after the harvest can increase fertilizer usage among farmers (Duflo et al., 2009), vaccination camps and small gifts can increase vaccination rates for children (Banerjee et al., 2010), and text message reminders can increase savings rates (Karlan et al., 2010).

Note that a (quasi-)hyperbolic model discounting model has trouble explaining this behavior: for instance, in the chlorination example, the account offered by the quasi-hyperbolic model is that people do not chlorinate their water because the immediate cost of doing so outweighs the delayed benefits. However, if this is the reason why individuals do not chlorinate their water when they have a chlorine bottle kept in the household, they should be even less inclined to chlorinate at the source, because at that time they can still procrastinate and decide to chlorinate at home. In other words, in the quasi-hyperbolic world, individuals postpone chlorinating until the last possible moment (consumption). The (quasi-)hyperbolic model does not explain the pre-crastination observed by Kremer et al. (2009), *unless* an increased cost for doing the task later is invoked. However, this is precisely the core assumption of the present model (and also the condition under which O'Donoghue & Rabin obtain "preproperation").

The remainder of the paper is organized as follows. Section 2 presents a simple version of the model in which utility is linear and the cost of keeping track is constant and lump-sum. Section 3 generalizes the results to the case where a reminder technology is available, and where agents are not fully sophisticated about their future cost of keeping track. Section 4 shows that most results hold for any concave and monotonically increasing utility function. Section 5 presents the design and results of several experiments conducted in Kenya which test predictions of the model. Section 6 describes a few applications of the model in the developing world. Section 7 concludes.

2 The Model: A simple example

We begin with a simple two-period model in which "tasks" arise at the beginning of period 0. Tasks consist either of payments to be made, or payments to be received. When a task arises, the agent either decides to act on it in period 0, or to act on it in period 1. Acting on a task consists in performing a costless action a. For instance, in the case of paying a bill, acting on the task consists in making the required bank transfer or writing a check; in the case of receiving a payment, acting on the task might consist in first sending one's bank details to the sender and verifying that the payment has arrived. Each task is defined by the payoff of acting on it immediately, x_0, and the payoff of acting on it later, x_1, with $x_1 \geq x_0$. These payoffs accrue in the period in which the agent acts on the task, although note that this is not required; the results shown below hold even when agents can borrow or save against the transfers, as long as the cost of keeping track is incurred in the correct period.

The core assumption of the model is that when an agent decides to act on a task in period 1, or any other period apart from period 0, the transaction may fail with probability p (because the agent forgets to act, or because external factors get in the way), in which case she incurs a cost c. In what follows, we will assume that $p = 1$ (i.e. future transactions always fail) to illustrate the main results. I begin by modeling the cost as a lump-sum cost; Section 4 extends the framwork to other formulations. Similarly, I begin by using linear utility; this choice is motivated by the fact that for the relatively small magnitude of the transactions which this model concerns, linear utility is a reasonable approximation. Section 4 provides a more general treatment.

Gains I now ask how agents behave when they face a task consisting of receiving a transfer on which they can act now or later. For instance, they may receive a check in the mail which they can cash immediately or later; they may decide between selling stock now or later; or in an analogous experiment, they may face a choice between a smaller amount available sooner or a larger amount available later.

The utility of acting in period 0 is:

$$u_0^+ = x_0 \tag{1}$$

The utility of acting in period 1 is a convex combination of the payoff when the transaction succeeds, and the payoff when the transaction fails:

$$\begin{aligned} u_1^+ &= \delta \left[(1-p)x_1 + p(x_1 - c)\right] \\ &= \delta(x_1 - pc) \end{aligned} \tag{2}$$

Thus, for acting in period 1, agents anticipate the discounted period 1 payoff, x_1, less the discounted cost of keeping track of the task. In this section, we set $p = 1$, without loss of generality, to illustrate the main results. With $p = 1$, the condition for preferring to act in period 1 is $u^1 > u^0$, which simplifies to:

$$x_1 > \frac{x_0}{\delta} + c \tag{3}$$

Thus, agents prefer to act in period 1 if the future gain exceeds the future-inflated value of the small-soon gain and the cost of keeping track.

Losses I now extend this framework to losses, by asking how agents behave when they are presented with a task consisting of making a transfer which they can complete now or later. For instance, they may receive a bill in the mail which they can pay immediately or later; farmers may decide between buying fertilizer now or later; or participants in an experiment may face a choice between losing a smaller amount of money immediately, or losing a larger amount later. To preserve the analogy to the framework for gains, I consider the utility of a smaller loss $-x_0$ incurred in period 0, and that of a larger loss $-x_1$ incurred in period 1. If they choose the larger loss in period 1, agents additionally pay the cost of keeping track. The utilities are thus as follows:

$$u_0^- = -x_0 \tag{4}$$

$$
\begin{aligned}
u_1^- &= \delta\left[(1-p)(-x_1) + p(-x_1 - c)\right] \\
&= \delta(-x_1 - pc)
\end{aligned}
\tag{5}
$$

Again setting $p = 1$, the condition for acting in period 1 is again given by $u_1 > u_0$, which simplifies to:

$$
x_1 < \frac{x_0}{\delta} - c \tag{6}
$$

Thus, agents prefer to delay losses if the future loss $-x_1$ is sufficiently small relative to the immediate loss net of the cost of keeping track.

I now discuss the implications of this framework for choice behavior.

Claim 1. Steeper discounting of gains: With a positive cost of keeping track, agents discount future gains more steeply than otherwise.

From 2, it is easy to see that $\frac{\partial u_1^+}{\partial c} = -\delta$. Thus, the discounted value of future gains decreases in the cost of keeping track, c; agents discount future gains more steeply the larger the cost of keeping track.

One implication of this result is that agents discount future outcomes at a higher rate than given by their time preference parameter. For instance, even agents who discount at the interest rate will exhibit choice behavior that looks like much stronger discounting when the cost of keeping track is high. The high discount rates frequently observed in experiments may partly be accounted for by participants correctly anticipating the cost of keeping track of the payment. For instance, in a standard discounting experiment, participants may be given a voucher to be cashed in in the future; with a positive probability of losing these vouchers, or of automatic payments not arriving, the future will be discounted more steeply than otherwise.

Claim 2. Shallower discounting of losses: With a positive cost of keeping track, agents discount future losses less steeply than otherwise.

As above, it follows from 5 that $\frac{\partial u_1^-}{\partial c} = -\delta$. Thus, the discounted utility of future losses decreases in the cost of keeping track, c; put differently, the *dis*utility of future losses *in*creases in c, i.e. future losses are discounted *less* as the cost of keeping track rises.

Intuitively, both delayed losses and delayed gains become less attractive because of the penalty for forgetting, which corresponds to steepr discounting for gains and shallower discounting for losses.

Claim 3. When agents choose between an equal-sized immediate vs. delayed loss, they prefer to delay when the cost of keeping track is zero, but may prefer to "pre-crastinate" with a positive cost of keeping track.

When the payoffs of acting now vs. acting later are both $-\bar{x}$, and $c = 0$, the condition for acting later on losses given in Equation 6 simplifies to $\bar{x} < \frac{\bar{x}}{\delta}$, which is always true with $\delta < 1$. Thus, when agents choose between equal-sized immediate vs. delayed losses and $c = 0$, they prefer to act in period 1. However, when $c > 0$, agents may prefer to act in period 0: the condition for acting in period 0 implied by 4 and 5 is $-\bar{x} > \delta(-\bar{x} - c)$, which simplifies to

$$c > \frac{1-\delta}{\delta}\bar{x}$$

When this condition is met, i.e. the cost of keeping track of having to act later is large enough, agents prefer to incur the loss in period 0 rather than period 1, i.e. they "pre-crastinate".

Under standard discounting, agents want to delay losses: a loss is less costly when it is incurred in the future compared to today. However, if the risk and penalty for forgetting to act in period 1 are sufficiently large relative to the payoff, agents prefer to act in period 0, i.e. they "pre-crastinate". For instance, such individuals may prefer to pay bills immediately because making a plan to pay them later is costly. This phenomenon corresponds well to everyday experience, and has recently been empirically demonstrated (Rosenbaum et al., 2014). However, it is not captured by standard discounting models, under which agents weakly prefer to delay losses.

It should be noted that this reasoning implies that a cost of keeping track model does not predict *pro*-crastination in the loss domain. I define pro-crastination as dynamic inconsistency in the loss domain, where agents decide to incur a loss in the earlier of two periods when those periods are in the future, but reverse their decision when the first of these periods arrives.[4] To see why a cost of keeping track does not predict this type of dynamic inconsistency, consider an agent who decides in period 0 between a loss of x_0 at $t = 1$ and a loss of x_1 at $t = 2$. The condition for choosing the large, late loss is $\delta^2(-x_1 - c) > \delta(-x_0 - c)$, which simplifies to $x_1 < \frac{x_0}{\delta} + \frac{1-\delta}{\delta}c$. When the agent reconsiders her decision at $t = 1$, the immediate loss is no longer subject to a cost of keeping track, and therefore the condition for choosing the large, late loss is $\delta(-x_1 - c) > -x_0$, which simplifies to $x_1 < \frac{x_0}{\delta} - c$. Note that this condition is harder to meet than the previous one; the agent therefore has no procrastination motive. In fact, she is motivated to incur the loss sooner rather than later – the pre-crastination described above.

Thus, a cost of keeping track model fails to generate one of the important predictions of the quasi-hyperbolic model, i.e. procrastination. However, both pre-crastination and procrastination can result when a cost of keeping track is added in a quasi-hyperbolic model. To see this, consider again an agent deciding in period 0 between a loss x_0 at $t = 1$ and a loss of x_1 at $t = 2$. Her preferences are quasi-hyperbolic and she faces a cost of keeping track. Thus, the utility of the early loss is $\beta\delta(-x_0 - c)$, and that of the delayed loss is $\beta\delta^2(-x_1 - c)$. The agent will prefer the delayed loss if $x_1 < \frac{x_0}{\delta} + \frac{1-\delta}{\delta}c$. In contrast, from the perspective of period 1, the utility of the now immediate loss is $-x_0$, that of the delayed loss is $\beta\delta(-x_1 - c) > -x_0$, and the condition for choosing the

[4] Note that this definition is somewhat asymmetrical with regard to my definition of pre-crastination, which is simply that agents prefer to incur a loss sooner rather than later; the reason for this asymmetry is that the analogous behavior to procrastination in the loss domain, i.e. dynamic inconsistency in the loss domain, has been called *pre-properation* by O'Donoghue & Rabin (1999). O'Donoghue and Rabin show that the quasi-hyperbolic model can account for this phenomenon; in contrast, it cannot account for precrastination as I define it here.

latter is $x_1 < \frac{x_0}{\beta\delta} - \beta c$. The agent will pre-crastinate when this condition is more difficult to meet than the previous one ($x_1 < \frac{x_0}{\delta} + \frac{1-\delta}{\delta}c$), and procrastinate when it is easier to meet. Thus, the agent procrastinates if $\frac{x_0}{\beta\delta} - \beta c > \frac{x_0}{\delta} + \frac{1-\delta}{\delta}c$, which simplifies to $c < x_0\frac{1-\beta}{(1-\delta+\delta\beta)\beta}$. Intuitively, if the cost of keeping track is small enough relative to the effect of hyperbolic discounting, the agent procrastinates, otherwise she precrastinates.

Claim 4. Sign effect: With a positive cost of keeping track, agents discount gains more than losses.

I show that the absolute value of the utility of a delayed loss is greater than that of a delayed gain, which corresponds to greater discounting of gains than losses. The absolute value of the utility of a delayed loss is

$$| u_1^- | = | \delta(-x_1 - c) | = \delta(x_1 + c)$$

Because $u_1^+ = \delta(x_1 - c)$, it is easy to see that $| u_1^- | > u_1^+$.

This result produces a the sign effect, a well-known departure of empirically observed time preferences from standard discounting models: agents discount losses less than gains.

Claim 5. Gain-loss asymmetry: With a positive cost of keeping track, agents exhibit a gain-loss asymmetry for future outcomes, similar to that observed in loss aversion.

Follows directly from Proposition 4.

Claim 6. Magnitude effect in the gains domain: With a positive cost of keeping track, agents discount large future gains less than small future gains.

Consider the utilities of acting now vs. later when both payoffs are multiplied by a constant $A > 1$:

$$u_0 = Ax_0$$

$$u_1 = \delta(Ax_1 - c)$$

The condition for acting in period 1 is now:

$$x_1 > \frac{x_0}{\delta} + \frac{c}{A}$$

Recall that the condition for acting on gains in period 1 with $c > 0$ is $x_1 > \frac{x_0}{\delta} + c$. Because $\frac{c}{A} < c$, the condition for acting in period 1 is easier to meet when the two outcomes are larger; thus, large outcomes are discounted less than small ones. It should be noted that concave utility is sufficient to produce a magnitude effect; a positive cost of keeping track exacerbates it. Note that this model predicts no magnitude effect when both outcomes are in the future.

Claim 7. Reversed magnitude effect in the loss domain: With a positive cost of keeping track, agents discount large losses more than small losses.

Consider the utilities of acting now vs. later when both losses are multiplied by a constant $A > 1$:

$$u_0 = -Ax_0$$

$$u_1 = \delta(-Ax_1 - c)$$

The condition for acting in period 1 is now:

$$x_1 < \frac{x_0}{\delta} - \frac{c}{A}$$

Recall that the condition for acting on losses in period 1 with $c > 0$ is $x_1 < \frac{x_0}{\delta} - c$. Because $\frac{c}{A} < c$, the condition for acting in period 1 is easier to meet when the two outcomes are larger. Because a preference for acting later corresponds to *more* discounting in the loss domain, this fact implies that large losses are discounted more than small ones. Thus, the magnitude effect is reversed in the loss domain. This prediction has recently been empirically confirmed by Hardisty (2011).

Claim 8. Decreasing impatience and dynamic inconsistency: With a positive cost of keeping track, agents exhibit decreasing impatience and dynamic inconsistency.

When both outcomes are moved one period into the future, they are both subject to the risk and penalty of forgetting; their utilities are:

$$u_1 = \delta(x_0 - c)$$

$$u_2 = \delta^2(x_1 - c)$$

The condition for acting later is

$$x_1 > \frac{x_0}{\delta} - \frac{1-\delta}{\delta}c$$

Note that this condition is easier to meet than condition 3 for choosing between acting immediately vs. next period, which is $x_1 > \frac{x_0}{\delta} + c$. Thus, when both outcomes are delayed into the future, the cost of waiting is smaller. As the future approaches, this will produce dynamic inconsistency.

Claim 9. Andreoni-Sprenger convex budgets, Effect 1: With a positive cost of keeping track, agents exhibit less discounting when adding money to existing payoffs than otherwise.

Assume a fixed initial payoff $\bar{x}$ in both periods 0 and 1. The lifetime utility of the agent in the absence of other transfers is

$$U(\bar{x}, \bar{x}) = \bar{x} + \delta(\bar{x} - c)$$

Now consider how this utility changes after adding x_0 in period 0 or x_1 in period 1:

$$U\left(\bar{x} + x_0, \bar{x}\right) = \bar{x} + x_0 + \delta(\bar{x} - c)$$

$$U(\bar{x}, \bar{x} + x_1) = \bar{x} + \delta(\bar{x} + x_1 - c)$$

The condition for acting later is $U(\bar{x}, \bar{x} + x_1) > U(\bar{x} + x_0, \bar{x})$, which simplifies to

$$x_1 > \frac{x_0}{\delta} \tag{7}$$

Note that this condition is again easier to meet than than condition 3 for choosing between acting immediately vs. next period *without* pre-existing payoffs at these timepoints. Thus, agents exhibit less discounting when money is added to existing payoffs than otherwise.

In their study on estimating time preferences from convex budgets, Andreoni & Sprenger (2012) pay the show-up fee of \$10 in two instalments: \$5 on the day of the experiment, and \$5 later. Even the payment on the day of the experiment is delivered to the student's mailbox rather than given at the time of the experiment itself, thus holding the cost of keeping track constant. The additional cost of payments now vs. later is thus minimal. Andreoni & Sprenger (2012) find much lower discount rates than most other studies on discounting. This finding is reflected in the result above.

Claim 10. Andreoni-Sprenger convex budgets, Effect 2: With a positive cost of keeping track, agents exhibit no decreasing impatience (such as hyperbolic discounting) when adding money to existing payoffs.

Assume again a fixed initial payoff $\bar{x}$ in both periods, but now move these periods one period into the future. The lifetime utility of the agent is

$$U = \delta(\bar{x} - c) + \delta^2(\bar{x} - c)$$

Now consider how this utility changes after adding x_0 in period 1, or x_1 in period 2:

$$U(0, \bar{x} + x_0, \bar{x}) = \delta(\bar{x} + x_0 - c) + \delta^2(\bar{x} - c)$$

$$U(0, \bar{x}, \bar{x} + x_1) = \delta(\bar{x} - c) + \delta^2(\bar{x} + x_1 - c)$$

The condition for acting later, $U(0, \bar{x}, \bar{x} + x_1) > U(0, \bar{x} + x_0, \bar{x})$, simplifies to

$$x_1 > \frac{x_0}{\delta}$$

Note that this condition is the same as that obtained in Claim 9. Thus, when money is added to existing payoffs now vs. next period, and when money is added to existing payoffs in two consecutive periods in the future, the conditions for preferring to act later are the same. This model therefore produces no decreasing impatience or dynamic inconsistency when adding money to existing payoffs. This mirrors the second result in Andreoni & Sprenger's (2012) study.

Together, these results predict many of the stylized facts that characterize empirically obtained discount functions. Figure 1 summarizes the magnitude and sign effects, decreasing impatience, and pre-crastination graphically.

3 Extensions

I now briefly consider two extensions to the basic setup: the availability of a reminder technology, and incomplete sophistication of the agent.

3.1 Reminders

A natural source of the cost of keeping track is a non-zero probability of forgetting to act on a task, combined with a non-zero cost incurred for forgetting. For what follows, it will be useful to make this intuition explicit. Assume that agents forget to act on future tasks with probability p, and remember them with probability $1 - p$. In the former case, they still earn the payoff, but pay cost c for salvaging the transaction. In the latter case, they simply earn the payoff. Thus, the utility of a future payoff is

$$\begin{aligned} u_1^+ &= \delta \left[(1-p)x_1 + p(x_1 - c) \right] \\ &= \delta(x_1 - pc). \end{aligned}$$

The above exposition can be seen as a special case of this setup, where the probability of forgetting is one.

So far we have assumed that no reminder technology is available: agents pay the cost of keeping track for any future transaction, and do not have an opportunity to avoid it by paying for a reminder. Now we relax this assumption: the agent has the opportunity to purchase a reminder when she makes a decision about when to act. For simplicity we set $p = 1$ again; the results extend easily to other values.

Consider again an agent who chooses at $t = 0$ between an immediate payoff of x_0 and a delayed payoff of x_1. Her utility for the immediate payoff is x_0, and that for the delayed payoff is $u_1^+ = \delta(x_1 - c)$.

Now assume that she can choose to buy a reminder at $t = 0$ which sets the cost of keeping track of the delayed payoff to zero[5]. Her utility in this case is:

$$u_1^+ = -r + \delta x_1$$

The condition for choosing to pay for the reminder vs. accepting the cost of keeping track is the following:

$$-r + \delta x_1 > \delta(x_1 - c)$$

$$r < \delta c$$

Thus, agents are willing to pay for reminders if their cost is smaller than the discounted cost of keeping track.

When choosing between an immediate payoff of x_0 in period 0, with utility $u_0^+ = x_0$, agents will therefore prefer the delayed outcome if

$$\max\{\delta(x_1 - c), -r + \delta x_1\} > x_0$$

If the price of the reminder is smaller than the discounted cost of keeping track, this corresponds to choosing the delayed outcome if

$$x_1 > \frac{x_0}{\delta} + \frac{r}{\delta}$$

Notice that because $\max\{\delta(x_1 - c), -r + \delta x_1\} \geq \delta(x_1 - c)$, reminders make the agent weakly more likely to choose the delayed outcome.

It is easy to see that most of the claims of the preceding section will still hold when agents can choose to pay for a reminder: she will discount gains more and losses less; if the cost of the reminder is large enough, she will prefer to incur losses sooner rather than later (pre-crastination); because the reminder cost is subtracted from both gains and losses, she will exhibit a sign effect and a gain-loss asymmetry similar to that observed in loss aversion; and because the cost of the reminder is lump-sum, she will exhibit a magnitude effect in the gains domain and a reversed magnitude effect in the loss domain. However, the results on decreasing impatience and less discounting and decreasing impatience when money is added to existing payoffs will change.

[5] Truly effective reminders are unlikely to be costless. First, even the small act of making a note in one's diary about the future task are *hassle costs* that can be cumbersome. Second, such simple reminders are not bullet-proof, and truly effective reminders are likely to be much more costly. For instance, consider the actions you would have to undertake to avoid forgetting an important birthday with a probability of one. Writing it in a diary is not sufficient because you might forget to look at it. Setting an alarm, e.g. on a phone, might fail because the phone might be out of battery at the wrong time. A personal assistant might forget himself to issue the reminder to you. Likely the most effective way of ensuring that the birthday is not forgotten would be to hire a personal assistant whose *sole* assignment is to issue the reminder. Needless to say, this would be rather costly. Cheaper versions of the same arrangement would come at the cost of lower probabilities of the reminder being effective.

3.1.1 Reminders and decreasing impatience

First consider the previous result on decreasing impatience. We had found earlier that from the perspective of $t = 0$, the condition to choose x_1 at $t = 2$ over x_0 at $t = 1$ was $x_1 > \frac{x_0}{\delta} - \frac{1-\delta}{\delta}c$. We now ask whether agents at $t = 0$ are willing to buy a reminder for the payoff at $t = 1$ or at $t = 2$, and whether agents at $t = 1$ are willing to buy a reminder for the payoff at $t = 2$. The agent can either choose the payoff at $t = 1$ with or without reminder, or the payoff at $t = 2$ with or without reminder; in the latter case, the reminder can be bought either immediately or in period 1, in which case it incurs a cost of keeping track of its own. The associated utilities are:

x_0 at $t = 1$:

$$\text{No reminder: } u^+_{1,\neg R} = \delta(x_0 - c)$$

$$\text{Reminder: } u^+_{1,R} = -r + \delta x_0$$

x_1 at $t = 2$:

$$\text{No reminder: } u^+_{2,\neg R} = \delta^2(x_1 - c)$$

$$\text{Reminder now: } u^+_{2,R_0} = -r + \delta^2 x_1$$

$$\left(\text{Reminder later: } u^+_{2,R_1} = \delta(-r - c) + \delta^2 x_1\right)$$

First note that for the large payoff in period 2, buying no reminder at all dominates buying a reminder in period 1: the cost of buying a reminder in period 1 is $\delta(r + c)$, which is strictly greater than its benefit $\delta^2 c$. Intuitively, the discounted cost of keeping track of the reminder is larger than the discounted cost of keeping track of the payoff in period 2, and therefore the agent would rather pay the cost of keeping track of the payoff itself. *The agent therefore always buys a reminder in period 0 if she buys one at all.*

Now consider under what circumstances the agent buys a reminder in period 0 when deciding between x_0 at $t = 1$ and x_1 at $t = 2$. It follows from the above exposition that for the payoff at $t = 1$, the agent wants to buy a reminder at $t = 0$ if $r < \delta c$. For the payoff at $t = 2$, the agent wants to buy a reminder at $t = 0$ if $r < \delta^2 c$. We can distinguish three cases:

1. No reminders First, assume that $r > \delta c$, i.e. the period 0 agent does not want to buy reminders at all. Recall from above that the condition for choosing the later payoff at $t = 0$ was

$$x_1 > \frac{x_0}{\delta} - \frac{1-\delta}{\delta}c.$$

2. Reminders for the payoff at $t = 1$, but not $t = 2$ Now assume that $\delta^2 c < r < \delta c$, i.e. the agent in period 0 prefers to buy a reminder for the payoff at $t = 1$, but not for that at $t = 2$. The condition for choosing the later outcome at $t = 0$ is

$$\delta^2(x_1 - c) > -r + \delta x_0,$$

which simplifies to

$$x_1 > \frac{x_0}{\delta} + c - \frac{r}{\delta^2}.$$

Because by assumption $\delta^2 c < r < \delta c$, this condition is harder to meet than the condition without reminders above. Thus, when reminders are available and cheaper than the cost of keeping track, the agent's decisions move closer to the condition for choosing the delayed payoff when deciding between an immediate payoff and a payoff delayed by one period; in other words, the agent moves closer to time consistency.

3. Reminders for both payoffs Finally, assume that both conditions for buying reminders are true, i.e. $r < \delta c$ (condition for wanting to buy a reminder in period 0 for the payoff at $t = 1$) and $r < \delta^2 c$ (condition for wanting to buy a reminder in period 0 for the payoff at $t = 2$). Because the first condition is strictly easier to meet than the second, this implies that both conditions are met when $r < \delta^2 c$. In this case, the agent wants to buy reminders for both payoffs at $t = 0$. The condition for choosing the later outcome at $t = 0$ is therefore

$$-r + \delta^2 x_1 > -r + \delta x_0,$$

which simplifies to

$$x_1 > \frac{x_0}{\delta}. \tag{8}$$

Thus, when reminders are available and cheaper than the discounted cost of keeping track, the agent's preference for x_1 vs. x_0 is undistorted by the cost of reminders and/or the cost of keeping track, and is instead only determined by the relative magnitude of the two payoffs and standard discounting.

No dynamic inconsistency with reminders Now assume that the agent has chosen to buy a reminder for the payoff at $t = 2$, and she reconsiders her decision at $t = 1$. If she had not already bought a reminder for the payoff at $t = 2$, she would buy one if $-r + \delta x_1 > x_0$, which implies $x_1 > \frac{x_0}{\delta} + \frac{r}{\delta}$. However, because she has already bought a reminder in period 0 for the payoff at $t = 2$, waiting from period 1 to period 2 for the delayed outcome is now costless, except for standard discounting. Thus, with the cost of the reminder sunk, in period 1 the agent chooses the period 2 outcome if

$$x_1 > \frac{x_0}{\delta}.$$

Notice that this condition is the same as condition 8 for choosing the delayed outcome from the perspective of period 0. Thus, the availability of reminders makes the agent time-consistent.

3.1.2 Reminders and discounting when money is added to existing payoffs

We had found above that agents discount less when money is added to existing future payoffs for which the agent has already paid a cost of keeping track. We now show that this effect disappears when the agent buys a reminder. Begin again by assuming that agents receive a fixed initial payoff $\bar{x}$ in both periods 0 and 1. The lifetime utility of the agent in the absence of other transfers is

$$U(\bar{x}, \bar{x}) = \bar{x} + \delta(\bar{x} - c)$$

Assume that the cost of reminders is such that agents buy a reminder at $t = 0$ for the payoff at $t = 1$, i.e. $r < \delta c$ (the cost of the reminder, incurred at $t = 0$, is smaller than the discounted cost of keeping track). The lifetime utility of the agent is now

$$U(\bar{x}, \bar{x}) = -r + \bar{x} + \delta\bar{x}$$

Now consider how this utility changes after adding x_0 in period 0 or x_1 in period 1:

$$U\left(\bar{x} + x_0, \bar{x}\right) = -r + \bar{x} + x_0 + \delta\bar{x}$$

$$U(\bar{x}, \bar{x} + x_1) = -r + \bar{x} + \delta(\bar{x} + x_1)$$

The condition for acting later is $U(\bar{x}, \bar{x} + x_1) > U(\bar{x} + x_0, \bar{x})$, which simplifies to

$$x_1 > \frac{x_0}{\delta}$$

Note that this condition is identical to condition 7, which specifies preferences when money is added to existing payoffs *without* the availability of reminders. Intuitively, when reminders are impossible or not desirable, agents discount less when money is added to existing payoffs because they incur a cost of keeping track of the delayed payoff regardless of whether they choose to add x_0 at $t = 0$ or x_1 at $t = 1$; analogously, when reminders are available, the result is identical because the cost of keeping track, paid in period 1, is now simply replaced with the cost of the reminder, paid in period 0, and again this cost is incurred regardless of whether agents choose to add x_0 at $t = 0$ or x_1 at $t = 1$ to the existing payoffs. Thus, both with and without reminders, agents discount less (i.e. only with δ) when money is added to existing payoffs.

Now consider the reduction in decreasing impatience we had described above when money is added to existing payoffs. Assume again a fixed initial payoff $\bar{x}$ in both periods, but now move these periods one period into the future. The lifetime utility of the agent is

$$U = \delta(\bar{x} - c) + \delta^2(\bar{x} - c)$$

Now assume that $r < \delta^2 c$, i.e. the agent want to buy reminders for both time periods. Because of the existing payoffs at these timepoints, she has to buy two reminders. Thus, her lifetime utility is

$$U = -2r + \delta\bar{x} + \delta^2\bar{x}.$$

Next, we consider how this utility changes after adding x_0 in period 1, or x_1 in period 2:

$$U(0, \bar{x} + x_0, \bar{x}) = -2r + \delta(\bar{x} + x_0) + \delta^2\bar{x}$$

$$U(0, \bar{x}, \bar{x} + x_1) = -2r + \delta\bar{x} + \delta^2(\bar{x} + x_1)$$

The condition for acting later, $U(0, \bar{x}, \bar{x} + x_1) > U(0, \bar{x} + x_0, \bar{x})$, simplifies to

$$x_1 > \frac{x_0}{\delta}.$$

Note that this condition is the same as that obtained in Propostion 10. Thus, when money is added to existing payoffs now vs. next period, and when money is added to existing payoffs in two consecutive periods in the future, the conditions for preferring to act later are the same, and this does not change when reminders are available.

Now assume that $\delta^2 c < r < \delta c$, i.e. the agent want to buy a reminder for the payoff at $t = 1$ but not for the payoff at $t = 2$. Her lifetime utility is

$$U = -r + \delta\bar{x} + \delta^2(\bar{x} - c).$$

Next, we consider how this utility changes after adding x_0 in period 1, or x_1 in period 2:

$$U(0, \bar{x} + x_0, \bar{x}) = -r + \delta(\bar{x} + x_0) + \delta^2(\bar{x} - c)$$

$$U(0, \bar{x}, \bar{x} + x_1) = -r + \delta\bar{x} + \delta^2(\bar{x} + x_1 - c)$$

The condition for acting later, $U(0, \bar{x}, \bar{x} + x_1) > U(0, \bar{x} + x_0, \bar{x})$, simplifies to

$$x_1 > \frac{x_0}{\delta}.$$

Again note that this condition is the same as that obtained in Propostion 10; thus, even when agents buy a reminder for only one of the two future periods, the conditions for preferring to act later are the same when the two periods are immediate compared to when they are in the future, and this does not change when reminders are available. This model therefore produces no decreasing impatience or dynamic inconsistency when adding money to existing payoffs, both with and without reminders.

3.2 Naïveté and Sophistication

We have so far assumed that agents have perfect foresight about the cost of forgetting. We now relax this assumption and ask how the results change when the agent underestimates her own cost of keeping track. This would naturally arise, for instance, if people are overconfident about their ability to remember future transactions, and thus their perceived probability of remembering to act on the task in the future might be larger than their true probability.

Consider again an agent who chooses at $t = 0$ between an immediate payoff of x_0 and a delayed payoff of x_1. Her utility for the immediate payoff is x_0, and that for the delayed payoff is $u_1^+ = \delta(x_1 - c)$. Now assume that her *perceived* utility of the delayed payoff is $u_1^+ = \delta\,(x_1 - \pi(c))$, where $\pi(\cdot)$ is the perceived cost of keeping track, with $\pi(c) < c$. Her utility for an immediate payoff remains $u_0^+ = x_0$. She will now choose the delayed payoff if

$$x_1 > \frac{x_0}{\delta} + \pi(c)$$

However, if she chooses the delayed outcome, her payoff at $t = 1$ will actually be $x_1 - c$, rather than (as she expected) $x_1 - \pi(c)$. Thus, there is a welfare loss associated with naïveté: if $\frac{x_0}{\delta} + \pi(c) < x_1 < \frac{x_0}{\delta} + c$, the agent will choose x_1 even though x_0 would have been preferred *ex post*.[6] The magnitude of the welfare loss from the perspective of period 0 is therefore $x_0 - \delta(x_1 - c)$.

The results are analogous for losses: the perceived utility of a delayed loss is $u_1^- = \delta(-x_1 - \pi(c))$; if the agent chooses the delayed outcome, the payoff in period 1 will be $-x_1 - c$, rather than the expected $-x_1 - \pi(c)$. In the case where $\frac{x_0}{\delta} - c < x_1 < \frac{x_0}{\delta} - \pi(c)$, the agent will choose the delayed loss even though the immediate loss of x_0 would have been preferred *ex post*, leading to a welfare loss of $-x_0 - \delta(-x_1 - c)$. Put differently, in this situation, naïveté leads the agent to *pro-crastinate* on losses when she would otherwise *pre-crastinate:* she will put off losses (e.g. a credit card bill)

[6] I follow Ericson (2014), O'Donoghue & Rabin (2006), Heidhues & Kőszegi (2010), and Gruber & Kőszegi (2002) in using *ex ante* welfare as the welfare criterion; i.e., welfare is defined by the period 0 preferences of the agent given correct beliefs about her preferences, costs, and constraints.

because she expects to remember it at the right time, even though she is in fact likely to forget about it. When she does, she incurs a cost.

4 The Model: General formulation

The previous discussion used linear utility and a lump-sum, one-time cost of keeping track to simplify the exposition. In the following, I make explicit when and how agents encounter tasks, how the cost of keeping track arises through the risk of forgetting, and how different formulations of the cost affect the main results.

4.1 When do agents think about tradeoffs? Memory and opportunity processes

In the exposition of the model, we have departed from the situation in which an agent faces a choice between tasks on which she can act now or later. In this section I briefly describe how these situations arise in the first place – i.e., under what circumstances does an agent consider a particular choice?

Let $\mathbf{s} \equiv (s_1, s_2, ...)$ be the agent's strategy, in which $s_t \in \{Y, N\}$ specifies for any period $t \in \{1, 2, ...\}$ whether the agents does or does not act on the task in that period. This formulation is similar to that of O'Donoghue & Rabin (1999).

Whether or not the agent acts on a task in a given period is governed by two additional processes. First, the agent's choice set in each period is constrained by an opportunity process $\mathbf{o} \equiv (o_1, o_2, ...)$, in which $o_t \in \{Y, N\}$ specifies for any period $t \in \{1, 2, ...\}$ whether or not the agent has an opportunity to act on the task in that period. Opportunities are exogenously given; for instance, Sarah might think of paying her credit card bill while at the gym, but has no opportunity to do it because her phone is at home. In the absence of an opportunity ($o_t = N$), agents cannot act in a given period, i.e. $s_t = \underline{N}$. This is true even when agents would otherwise have preferred to act in that period, i.e. $s_t^* = Y$.

Second, let $\mathbf{m} \equiv (m_1, m_2, ...)$ with $m_t \in \{Y, N\}$ be a *memory process* which specifies for any period $t \in \{1, 2, ...\}$ whether or not the agent remembers the task in that period. If and only if the agent remembers the task in a given period, i.e. $m_t = Y$, she makes a decision whether to act in that period. If she has previously made a decision, she revisits it. If the agent does not remember the task or opportunity ($m_t = N$), she will not act on it in that period, i.e. $s_t = \underline{N}$. This is true even when the agent would otherwise have preferred to act in that period, i.e. $s_t^* = Y$.

To make explicit how the memory process operates, let $\mathbf{r} = (r_1, r_2, ...)$ be a reminder process which specifies for any period $t \in \{1, 2, ...\}$ whether the agent receives a reminder about the task in period t. This reminder can be endogenous – i.e., the agent might have chosen in the past to set up a reminder to arrive in the particular period – or exogenous, i.e. the agent might be reminded by external stimuli in the environment, or by random fluctuations in her own memory process. We assume for simplicity that reminders are fully effective, i.e. a reminder always causes an agent to remember the task. Formally, $m_t = Y$ if and only if $r_t = Y$. Note that the reminder process is identical to the memory process in this formulation and thus is not needed for the results; however, we treat it separately because on the one hand it provides intuition for the circumstances under which agents remember tasks, and on the other hand it leaves room for future work to introduce the possibility of imperfect or probabilistic reminders.

Thus, agents will make a decision between acting vs. not acting on a task in period t if they both remember it and have an opportunity to act on it, i.e. $m_t = Y$ and $o_t = Y$. If either $m_t = N$ or $o_t = N$, agents do not act even though they might have done so otherwise. Note that the two cases are different in that agents still consider what they would ideally do when they remember the choice but do not have an opportunity to act ($m_t = Y$, $o_t = N$), while they do not formulate such a "shadow strategy" when they do not remember ($m_t = N$).

4.2 The risk of forgetting

Perhaps the most compelling argument in support of the assumption that agents incur a cost for keeping track of future transactions is the fallibility of human memory. If agents are more likely to forget acting on future transactions than current transactions, and if such forgetting entails to a cost, this setup naturally leads to the results derived in the previous section. Specifically, we assume that when an agent makes a plan to act on a task in the future, she forgets about this plan with probability $p \geq 0$ in each period, and remembers it with probability $1 - p \leq 1$. Thus, remembering an action from period 0 to period t entails a probability of remembering of $(1-p)^t$ and a probability of forgetting of $1 - (1 - p)^t$.[7]

If she forgets to perform the action required to act on the task, the agent receives a smaller payoff $x_D < x_1$; in the case of losses, she incurs a greater loss, $-x_D < -x_1$. We can think of the difference between x_D and x_1 as the cost $C^t_{x_1}$ of forgetting to keep track of the payment over t periods, with $C^t_{x_1} = u(x_1) - u(x_D) \geq 0$ for gains and $C^t_{x_1} = u(x_D) - u(x_1) \geq 0$ for losses. For instance, the agent can still cash a check that she forgot to cash by the deadline, but incurs a cost to "salvage" the

[7] From this discussion, naturally the question arises whether an exponential probability of forgetting for all future periods is justified. In fact, this particular choice is motivated by economic convention rather than by evidence, and it is conversative in that it makes the results weaker than they would be if the empirically observed shape of the forgetting function were used. The reason lies in the current consensus in psychology about the empirical shape of the forgetting function: as was first casually observed by the German psychologist Jost in 1897, and confirmed by Wixted and colleagues (e.g. Wixted, 2004) and many other studies since, the shape of the psychological forgetting function is not well described by an exponential function, but follows instead a power law, such as a hyperbola. It can intuitively easily be seen that some of the results described above which do not hold when the cost of keeping track is proportional to the payoff would hold with a hyperbolic rather than expoential forgetting function. As an example, notice that an exponential forgetting function with a proportional cost of keeping track essentially amounts to stronger exponential discounting; in contrast, introducing a hyperbolic forgetting function adds a hyperbolic element. Similarly, note that a *constant* probability of forgetting attached to future but not present transactions ($p_1 = p_1 = p_t = \bar{p}$) is a quasi-hyperbolic forgetting function similar to that used for discounting by Strotz (1956) and Laibson (1997).

transaction, e.g. by paying an administrative fee to have a new check issued. Similarly, if the agent fails to pay a bill by the deadline, the bill will still be paid, but the agent now has to pay a late fee.

4.3 Cost structure

In Section 2, we had considered a one-time, lump-sum cost of forgetting. To study how agents behave under different cost structures, we now incorporate up to four possible components of the cost: c is a lump-sum one-time cost; m is a lump-sum per-period cost; γ is a proportional one-time cost; and α is a proportional and compound per-period cost. We can therefore write the total cost of keeping track of a payment x over t periods as follows:

$$C_x^t = \begin{cases} 0, & x_t = 0 \\ c + mt + \gamma x_t + \left[(1+\alpha)^t x_t - x_t \right], & x_t \neq 0 \end{cases}$$

with

$$\begin{aligned} c &\geq 0 \\ m &\geq 0 \\ \gamma &\geq 0 \\ \alpha &\geq 0 \end{aligned}$$

and at least one strict inequality. The intuition here is that a cost of keeping track is incurred only when there is a payment x_t to be kept track of until period t. If this is not the case, the cost is zero. The following section makes explicit that this cost is borne only if the agent forgets to act on the task.

The intuition for the four components of the cost is as follows. First, one-time, lump-sum cost of keeping track might consist in the time and effort costs of setting up a reminder to keep track of the task, a lump-sum late fee for failing to pay a credit card bill on time, or the fixed cost of re-issuing a check after failing to cash it within its validity period. Second, a lump-sum per-period cost might consist of the cognitive effort expended on remembering the task over time; a "mental hassle cost".[8] Third, a one-time, proportional cost might consist in setting up a reminder whose cost depends on the magnitude of the expected payment (e.g. I might hire an assistant to ensure I do not forget about a large anticipated payment, but only set up a calendar reminder for a smaller payment).[9] Finally, a per-period, proportional cost with compounding might arise from having to pay interest after failing to pay a credit card bill on time. In what follows below, I ask which of the results

[8] Such psychological costs for keeping transactions in mind might be incurred even if agents have perfect memory. The psychological cost of juggling many different tasks has recently attracted increased interest in psychology and economics. Most prominently, Shafir & Mullainathan (2013) argue that the poor in particular may have so many things on their mind that only the most pressing receive their full attention. This argument implies that a) allocating attention to a task is not costless, and b) the marginal cost of attention is increasing in the number of tasks. Together, this reasoning provides an intuition for a positive *psychological* cost of keeping track of future transactions, even with perfect memory and no financial costs of keeping track.

[9] Appendix A deals with the possibility that the probability of forgetting is lower for large payments.

sketched above hold under these different cost structures. Most of the results hold as long as there is *any* lump sum component contained in the cost of forgetting (e.g., credit card late fees are often a percentage of the balance, but also have a lump-sum administrative fee).

4.4 Utility

The Section 2, we had assumed linear utility. To provide a more standard treatment, we now assume that agents have a utility function $u(\cdot)$ which is continuous, twice differentiable for $x \neq 0$, strictly monotonically increasing, concave, symmetric around $u(0)$, and $u(0) = 0$. Together with the probability of forgetting and the associated cost, we can now formulate a more general version of the cost of keeping track. The lifetime utility of the agent is given by:

$$
\begin{aligned}
U(x_0, x_1, ...) &= u(x_0) + \sum_{t=1}^{\infty} \delta^t \mathbb{E}\left[u(x_t)\right] \\
&= u(x_0) + \sum_{t=1}^{\infty} \delta^t \left[(1-p)^t u(x_t) + \left[1-(1-p)^t\right] u(x_t - C_{x_t}^t)\right]
\end{aligned}
$$

with

$$
C_x^t = \begin{cases} 0, & x_t = 0 \\ c + mt + \gamma x_t + \left[(1+\alpha)^t x_t - x_t\right], & x_t \neq 0 \end{cases}
$$

In most of what follows, we restrict ourselves to a two-period vesion of this infinite-horizon model, in which the agent chooses between acting on x_0 and x_1 in periods 0 and 1, respectively, with $x_1 \geq x_0$.

Gains For gains, the utility of acting in period 0 is:

$$
u_0^+ = u(x_0) \tag{9}
$$

The utility of acting in period 1 is:

$$
\begin{aligned}
u_1^+ &= \delta \mathbb{E}\left[u(x_1)\right] \\
&= \delta\left[(1-p)\, u(x_1) + pu(x_1 - C_{x_1}^1)\right] \\
&= \delta\left[(1-p)\, u(x_1) + pu(x_1 - c - m - \gamma x_1 - \alpha x_1)\right]
\end{aligned} \tag{10}
$$

The condition for preferring to act in period 1 is:

$$
(1-p)\, u(x_1) + pu\left(x_1 - C_{x_1}^1\right) > \frac{u(x_0)}{\delta} \tag{11}
$$

The optimal strategy of the agent is therefore

$$\mathbf{s}^* = \begin{cases} (Y,N), & u(x_0) \geq \delta\left[(1-p)u(x_1) + pu(x_1 - C^1_{x_1})\right] \\ (N,Y), & u(x_0) < \delta\left[(1-p)u(x_1) + pu(x_1 - C^1_{x_1})\right] \end{cases}$$

Losses The utilities of acting on losses in periods 0 and 1, respectively, are as follows:

$$u_0^- = u(-x_0) = -u(x_0) \tag{12}$$

$$\begin{aligned} u_1^- &= \delta\mathbb{E}\left[u(x_1)\right] \\ &= \delta\left[(1-p)\,u(-x_1) + pu(-x_1 - C^1_{x_1})\right] \\ &= \delta\left[(1-p)\,u(-x_1) + pu(-x_1 - c - m - \gamma x_1 - \alpha x_1)\right] \end{aligned} \tag{13}$$

The condition for acting in period 1 is again given by $u_1 > u_0$, which, invoking the symmetry of $u(\cdot)$ around $u(0)$, simplifies to:

$$(1-p)\,u(x_1) + pu(x_1 + C^1_{x_1}) < \frac{u(x_0)}{\delta} \tag{14}$$

The optimal strategy of the agent is therefore

$$\mathbf{s}^* = \begin{cases} (Y,N), & u(x_0) \leq \delta\left[(1-p)u(x_1) + pu(x_1 + C^1_{x_1})\right] \\ (N,Y), & u(x_0) > \delta\left[(1-p)u(x_1) + pu(x_1 + C^1_{x_1})\right] \end{cases}$$

4.5 Results

We can now formalize the claims made in Section 2 in turn.

Proposition 1. *Steeper discounting of gains: With a positive cost of keeping track, agents discount future gains more steeply than otherwise.*

Proof. From 10, it is easy to see that $\frac{\partial u_1^+}{\partial C^1_{x_1}} < 0$ regardless of which parameter c, m, γ, or α is strictly positive. Thus, agents discount future gains more steeply the larger any given component of the cost of keeping track. □

Proposition 2. *Shallower discounting of losses: With a positive cost of keeping track, agents discount future losses less steeply than otherwise.*

Proof. As above, it follows from 13 that $\frac{\partial u_1^-}{\partial C^1_{x_1}} < 0$. Thus, the disutility of future losses increases in $C^1_{x_1}$, which implies that future losses are discounted less as the cost of keeping track increases. □

Proposition 3. *Pre-crastination: When agents choose between an equal-sized immediate vs. delayed loss, they prefer to delay when the cost of keeping track is zero, but may prefer to "pre-crastinate" with a positive cost of keeping track.*

Proof. When the payoffs of acting now vs. acting later are both $-\bar{x}$, and $c = m = \gamma = \alpha = 0$, the condition for acting later on losses given in equation 14 simplifies to $u(\bar{x}) < \frac{u(\bar{x})}{\delta}$, which is always true with $\delta < 1$. Thus, when agents choose between equal-sized losses in periods 0 and 1, and the cost of keeping track is zero, they prefer to act in period 1. However, when $c \geq 0$, $m \geq 0$, $\gamma \geq 0$, and $\alpha \geq 0$ with at least one strict inequality (i.e., $C^1_{\bar{x}} > 0$), agents may prefer to act in period 0: the condition for acting in period 0 implied by 12 and 13 is $-u(\bar{x}) > -\delta(1-p)u(\bar{x}) - \delta p u\left(\bar{x} + C^1_{\bar{x}}\right)$. This expression simplifies to:

$$\frac{u(\bar{x})}{u\left(\bar{x} + C^1_{\bar{x}}\right)} < \frac{\delta p}{1 - \delta(1-p)}$$

Because $0 \leq \frac{u(\bar{x})}{u\left(\bar{x}+C^1_{\bar{x}}\right)} < 1$ by the strict monotonicity of $u(\cdot)$, this condition can be met with a sufficiently large cost of keeping track and sufficiently large δ. In this case, agents prefer to incur the loss in period 0 rather than period 1, i.e. they "pre-crastinate". □

Proposition 4. *Sign effect: With a positive cost of keeping track, agents discount gains more than losses.*

Proof. I show that the absolute value of the utility of a delayed loss is greater than that of a delayed gain, which corresponds to greater discounting of gains than losses. By symmetry of $u(\cdot)$ around $u(0)$, the absolute value of the utility of a delayed loss is

$$\mid u_1^- \mid = \mid -\delta(1-p)u(x_1) - \delta p u\left(x_1 + C^1_{x_1}\right) \mid = \delta(1-p)u(x_1) + \delta p u\left(x_1 + C^1_{x_1}\right)$$

Because $u_1^+ = \delta(1-p)u(x_1) + \delta p u\left(x_1 - C^1_{x_1}\right)$ and $u'(\cdot) > 0$, it is easy to see that $\mid u_1^- \mid > u_1^+$. Thus, the absolute value of the utility of a delayed loss is greater than that of a delayed gain. □

Proposition 5. *Gain-loss asymmetry: With a positive cost of keeping track, agents exhibit a gain-loss asymmetry for future outcomes, similar to that observed in loss aversion.*

Proof. Follows directly from Proposition 4. □

Proposition 6. *Magnitude effect in the gains domain: With a positive cost of keeping track, agents discount large future gains less than small future gains.*

Proof. The magnitude effect requires that the discounted utility of a large future payoff Ax ($A > 1$) be larger, as a proportion of the undiscounted utility of the same payoff, than that of a smaller payoff x:

$$\frac{\delta\left(1-p\right)u\left(Ax\right)+\delta pu\left(Ax-C_{Ax}^{1}\right)}{u(Ax)}>\frac{\delta\left(1-p\right)u\left(x\right)+\delta pu\left(x-C_{x}^{1}\right)}{u(x)} \tag{15}$$

By concavity of $u(\cdot)$, and given that $A>1$,

$$u(Ax+C_x^1)-u(Ax)<u(x+C_x^1)-u(x)$$

Because $u\left(\cdot\right)$ is monotonically increasing, $u(Ax)>u(x)$, and therefore:

$$\frac{u(Ax+C_x^1)-u(Ax)}{u(Ax)}<\frac{u(x+C_x^1)-u(x)}{u(x)}$$

Adding one on both sides and decreasing the arguments by C_x^1:

$$\frac{u(Ax)}{u(Ax-C_x^1)}<\frac{u(x)}{u(x-C_x^1)}$$

Inverting the fractions and multiplying both sides by δp, we obtain:

$$\frac{\delta pu(Ax-C_x^1)}{u(Ax)}>\frac{\delta pu(x-C_x^1)}{u(x)}$$

Adding $\delta\left(1-p\right)$ to both sides and combining each side into a single fraction, we obtain the desired result. □

It is natural to ask at this point whether this result would still hold if the probability of forgetting large payoffs were smaller than that of forgetting small payoffs. Indeed, this result can easily be derived; we show it in Appendix A.

Proposition 7. *Reversed magnitude effect in the loss domain: With a positive cost of keeping track, agents discount large losses more than small losses.*

This proof proceeds in an analogous fashion to that for Proposition 6. A reversed magnitude effect requires that the discounted utility of a large future loss Ax $(A>1)$ be *smaller*, as a propotion of the undiscounted utility of the same payoff, than that of a smaller payoff x:

$$\frac{-\delta\left(1-p\right)u\left(Ax\right)-\delta pu\left(Ax+C_{Ax}^{1}\right)}{-u(Ax)}<\frac{-\delta\left(1-p\right)u\left(x\right)-\delta pu\left(x+C_{x}^{1}\right)}{-u(x)} \tag{16}$$

Proof. Multiplying both sides by -1, we want to show that:

$$\frac{\delta\left(1-p\right)u\left(Ax\right)+\delta pu\left(Ax+C_{Ax}^{1}\right)}{u(Ax)}<\frac{\delta\left(1-p\right)u\left(x\right)+\delta pu\left(x+C_{x}^{1}\right)}{u(x)}$$

By concavity of $u(\cdot)$, and given that $A > 1$,

$$u(Ax + C_x^1) - u(Ax) < u(x + C_x^1) - u(x)$$

Because $u(\cdot)$ is monotonically increasing, $u(Ax) > u(x)$, and therefore:

$$\frac{u(Ax + C_x^1) - u(Ax)}{u(Ax)} < \frac{u(x + C_x^1) - u(x)}{u(x)}$$

Adding one on both sides and multiplying both sides by $-\delta p$, we obtain:

$$\frac{\delta p u(Ax + C_x^1)}{u(Ax)} < \frac{\delta p u(x + C_x^1)}{u(x)}$$

Adding $\delta(1-p)$ to both sides and combining each side into a single fraction, we obtain the desired result. □

Proposition 8. *Decreasing impatience and dynamic inconsistency: With a positive cost of keeping track, agents exhibit decreasing impatience and dynamic inconsistency except when the cost is a proportional and compound per-period cost.*

Proof. When both outcomes are moved one period into the future, they are both subject to the risk and penalty of forgetting; their utilities are:

$$\begin{aligned} u_1(x_0) &= \delta(1-p)\,u(x_0) + \delta p u(x_0 - C_{x_0}^1) \\ &= \delta(1-p)\,u(x_0) + \delta p u(x_0 - c - m - \gamma x_0 - \alpha x_0) \end{aligned}$$

$$\begin{aligned} u_2(x_1) &= \delta^2(1-p)^2\,u(x_1) + \delta^2\left[1-(1-p)^2\right]u(x_1 - C_{x_1}^2) \\ &= \delta^2(1-p)^2\,u(x_1) + \delta^2\left[1-(1-p)^2\right]u\left[x_1 - c - 2m - \gamma x_1 - \left((1+\alpha)^2 x_1 - x_1\right)\right] \end{aligned}$$

The condition for acting later can be written as follows:

$$u_2(x_1) > u_1(x_0)$$

$$\delta^2 (1-p)^2 u(x_1) + \delta^2 \left[1-(1-p)^2\right] u \left[x_1 - c - 2m - \gamma x_1 - \left((1+\alpha)^2 x_1 - x_1\right)\right]$$
$$> \delta (1-p) u (x_0) + \delta p u (x_0 - c - m - \gamma x_0 - \alpha x_0)$$

$$(1-p)^2 u(x_1) + \left[1-(1-p)^2\right] u \left[x_1 - c - 2m - \gamma x_1 - \left((1+\alpha)^2 x_1 - x_1\right)\right] \quad (17)$$
$$> \frac{(1-p) u (x_0) + p u (x_0 - c - m - \gamma x_0 - \alpha x_0)}{\delta}$$

$$(1-p)^2 u(x_1) + \left[1-(1-p)^2\right] u \left[(x_1 - c - m - \gamma x_1 - \alpha x_1) - m - (\alpha^2 + \alpha) x_1\right]$$
$$> \frac{(1-p) u(x_0) + p u(x_0 - c - m - \gamma x_0 - \alpha x_0)}{\delta}$$

To obtain decreasing impatience, we want to show that this condition is easier to meet than our original condition 11 for choosing between acting immediately vs. next period. When this is the case, impatience decreases when both outcomes are delayed into the future. We recall that our original condition is:

$$(1-p) u (x_1) + p u(x_1 - c - m - \gamma x_1 - \alpha x_1) > \frac{u(x_0)}{\delta} \quad (18)$$

We consider two cases: (1) $\alpha = 0$ and (2) $\alpha > 0$. Recall that $0 \leq p \leq 1$. We show that decreasing impatience holds when $\alpha = 0$, but only under some conditions when $\alpha > 0$.

We first show that we obtain decreasing impatience when $\alpha = 0$. With this restriction, the condition of interest and the original condition become, respectively:

$$(1-p^2) u(x_1) + p^2 u \left[(x_1 - c - m - \gamma x_1) - m\right] > \frac{(1-p) u(x_0) + p u(x_0 - c - m - \gamma x_0)}{\delta}$$

$$(1-p) u(x_1) + p u(x_1 - c - m - \gamma x_1) > \frac{u(x_0)}{\delta}$$

When $0 \leq p < 1$, the LHS of the condition of interest places more weight on the larger $u(x_1)$ term. Thus the LHS of the condition of interest is larger than the LHS of the original condition as long as there is a non-zero cost parameter. Similarly, the RHS of the condition of interest is a weighted average of $u(x_0)$ and $u(x_0 - c - m - \gamma x_0)$, which is smaller than the unweighted $u(x_0)$ of the RHS of the original condition. Thus, we establish our desired result. When $p = 1$, the two conditions simplify as follows:

$$u\left[(x_1 - c - m - \gamma x_1) - m\right] > \frac{u(x_0 - c - m - \gamma x_0)}{\delta}$$

$$u(x_1 - c - m - \gamma x_1) > \frac{u(x_0)}{\delta}$$

Compared to the original condition, the LHS of the condition of interest is reduced by m, and the RHS is reduced by $c + m + \gamma x_0$. Thus, by the concavity of $u(\cdot)$ we still obtain our desired result as long as one of the cost parameters is non-zero.

When $\alpha > 0$, the condition of interest and the original condition are as in equations 17 and 18. We compare term by term. The RHS terms are strightforward: the RHS of the condition of interest is smaller than the RHS of the original condition, as long as at least one cost parameter is non-zero. This helps toward our desired result. The LHS is ambiguous: the fact that $0 \leq p \leq 1$ means that $1 - (1-p)^2 > p$, which helps towards the desired result. However, we also have that $(1-p)^2 \leq 1 - p$, with the equality holding only at the corners. This does not help toward the desired result. Furthermore, given $\alpha > 0$ and $0 \leq p \leq 1$, we have that $u(x_1 - c - m - \gamma x_1 - \alpha x_1) > u\left[(x_1 - c - m - \gamma x_1 - \alpha x_1) - m - (\alpha^2 + \alpha)x_1\right]$, which again does not help towards the desired result. Thus, decreasing impatience is possible when $\alpha > 0$, but not guaranteed. Overall, we establish decreasing impatience, and the associated dynamic inconsistency, when $\alpha = 0$, but not necessarily when $\alpha > 0$. □

Proposition 9. *Andreoni-Sprenger convex budgets, Effect 1: With a positive cost of keeping track, agents exhibit less discounting when adding money to existing payoffs than otherwise under certain parameter values.*

Proof. Assume a fixed initial payoff $\bar{x}$ in both periods 0 and 1. The lifetime utility of the agent in the absence of other transfers is

$$U(\bar{x}, \bar{x}) = u(\bar{x}) + \delta(1-p)\,u(\bar{x}) + \delta p u(\bar{x} - C^1_{\bar{x}})$$

Now consider how this utility changes after adding x_0 in period 0 or x_1 in period 1:

$$U(\bar{x} + x_0, \bar{x}) = u(\bar{x} + x_0) + \delta(1-p)\,u(\bar{x}) + \delta p u(\bar{x} - C^1_{\bar{x}})$$
$$U(\bar{x}, \bar{x} + x_1) = u(\bar{x}) + \delta(1-p)\,u(\bar{x} + x_1) + \delta p u(\bar{x} + x_1 - C^1_{\bar{x}+x_1})$$

The condition for acting later is $U(\bar{x}, \bar{x} + x_1) > U(\bar{x} + x_0, \bar{x})$, which we rearrange as follows:

$$u(\bar{x}) + \delta(1-p)\,u(\bar{x} + x_1) + \delta p u(\bar{x} + x_1 - C^1_{\bar{x}+x_1}) > u(\bar{x} + x_0) + \delta(1-p)\,u(\bar{x}) + \delta p u(\bar{x} - C^1_{\bar{x}}) \quad (19)$$

Rearranging further, we obtain:

$$u(\bar{x}) - u(\bar{x} + x_0) + \delta(1-p)u(\bar{x} + x_1) - \delta(1-p)u(\bar{x}) + \delta pu(\bar{x} + x_1 - C^1_{\bar{x}+x_1}) - \delta pu(\bar{x} - C^1_{\bar{x}}) > 0$$

$$\begin{aligned} & u(\bar{x}) - u(\bar{x} + x_0) + \delta(1-p)\left[u(\bar{x} + x_1) - u(\bar{x})\right] \\ & + \delta pu(\bar{x} + x_1 - c - m - \gamma(\bar{x} + x_1) - \alpha(\bar{x} + x_1)) - \delta pu(\bar{x} - c - m - \gamma\bar{x} - \alpha\bar{x}) > 0 \end{aligned}$$

$$\begin{aligned} & u(\bar{x}) - u(\bar{x} + x_0) + \delta(1-p)\left[u(\bar{x} + x_1) - u(\bar{x})\right] \\ & + \delta p\left[u(\bar{x} + x_1 - c - m - \gamma\bar{x} - \alpha\bar{x} - \gamma x_1 - \alpha x_1) - u(\bar{x} - c - m - \gamma\bar{x} - \alpha\bar{x})\right] > 0 \end{aligned}$$

$$\begin{aligned} & (1-p)\left[u(\bar{x} + x_1) - u(\bar{x})\right] + pu(\bar{x} + x_1 - c - m - \gamma\bar{x} - \alpha\bar{x} - \gamma x_1 - \alpha x_1) \\ & - pu(\bar{x} - c - m - \gamma\bar{x} - \alpha\bar{x}) > \frac{u(\bar{x} + x_0) - u(\bar{x})}{\delta} \end{aligned}$$

We establish our result if this new condition is less strict that the original condition below:

$$(1-p)u(x_1) + pu(x_1 - c - m - \gamma x_1 - \alpha x_1) > \frac{u(x_0)}{\delta}$$

We distinguish four cases:

1. When $c > 0$, with all other cost parameters zero, the conditions simplify as follows:

$$(1-p)\left[u(\bar{x} + x_1) - u(\bar{x})\right] + p\left[u(\bar{x} + x_1 - c) - u(\bar{x} - c)\right] > \frac{u(\bar{x} + x_0) - u(\bar{x})}{\delta}$$

$$(1-p)u(x_1) + pu(x_1 - c) > \frac{u(x_0)}{\delta}$$

To establish the result, we require that:

$$\begin{aligned} & (1-p)\left[u(\bar{x} + x_1) - u(\bar{x})\right] + p\left[u(\bar{x} + x_1 - c) - u(\bar{x} - c)\right] \\ & - \frac{u(\bar{x} + x_0) - u(\bar{x})}{\delta} - (1-p)u(x_1) - pu(x_1 - c) + \frac{u(x_0)}{\delta} > 0 \end{aligned}$$

This implies that for large values of p, the result holds when $x_1 - x_0$ is very small, or when $\bar{x} - c$ is very small. For small values of p, the result holds only when $\bar{x} - c$ is very small.

2. When $m > 0$, with all other cost parameters zero, the conditions are symmetric to those of the $c > 0$ case.

3. When $\gamma > 0$, with all other cost parameters zero, our conditions simplify to:

$$(1-p)\left[u\left(\bar{x}+x_1\right)-u\left(\bar{x}\right)\right]+p\left[u((1-\gamma)\left(\bar{x}+x_1\right))-u\left((1-\gamma)\,\bar{x}\right)\right]>\frac{u(\bar{x}+x_0)-u(\bar{x})}{\delta}$$

$$(1-p)\,u\left(x_1\right)+pu\left((1-\gamma)\,x_1\right)>\frac{u\left(x_0\right)}{\delta}$$

It can be seen that for the result to hold, we require not only that x_1-x_0 is large, but also that $x_0-\bar{x}$ is very small, and that p is not too small.

4. When $\alpha>0$, with all other cost parameters zero, the conditions are symmetric to those of the $\gamma>0$ case.

In sum, whether or not the result holds when multiple cost parameters are non-zero depends on the relative values of the cost parameters. When p is large, the result holds for the $c>0$ and $m>0$ cases when x_1-x_0 is very small, but it holds for the $\gamma>0$ and $\alpha>0$ cases when x_1-x_0 is very large. □

Proposition 10. *Andreoni-Sprenger convex budgets, Effect 2: With a positive cost of keeping track, agents exhibit more hyperbolic discounting when adding money to existing payoffs under certain parameter values.*

Proof. We assume a fixed initial payoff $\bar{x}$ in both periods. The condition for acting in period 1 over period 0 is:

$$\begin{aligned}&u(\bar{x})+\delta\left(1-p\right)u\left(\bar{x}+x_1\right)+\delta pu\left(\bar{x}+x_1-C^1_{\bar{x}+x_1}\right)\\ >&u\left(\bar{x}+x_0\right)+\delta\left(1-p\right)u\left(\bar{x}\right)+\delta pu\left(\bar{x}-C^1_{\bar{x}}\right)\end{aligned} \tag{20}$$

When both periods are in the future, the condition for acting in period 2 over period 1 is:

$$\begin{aligned}&\delta\left(1-p\right)u\left(\bar{x}\right)+\delta pu(\bar{x}-C^1_{\bar{x}})+\delta^2\left(1-p\right)^2u\left(\bar{x}+x_1\right)+\delta^2\left[1-\left(1-p\right)^2\right]u\left(\bar{x}+x_1-C^2_{\bar{x}+x_1}\right)\\ >&\delta\left(1-p\right)u\left(\bar{x}+x_0\right)+\delta pu\left(\bar{x}+x_0-C^1_{\bar{x}+x_0}\right)+\delta^2\left(1-p\right)^2u\left(\bar{x}\right)+\delta^2\left[1-\left(1-p\right)^2\right]u\left(\bar{x}-C^2_{\bar{x}}\right)\end{aligned} \tag{21}$$

Rearranging 20 and 21, we obtain:

$$\begin{aligned}&u(\bar{x})+\delta\left(1-p\right)u\left(\bar{x}+x_1\right)+\delta pu\left(\bar{x}+x_1-c-m-\gamma\left(\bar{x}+x_1\right)-\alpha\left(\bar{x}+x_1\right)\right)\\ >&u\left(\bar{x}+x_0\right)+\delta\left(1-p\right)u\left(\bar{x}\right)+\delta pu\left(\bar{x}-c-m-\gamma\bar{x}-\alpha\bar{x}\right)\end{aligned} \tag{22}$$

$$\begin{aligned}&(1-p)\,u\left(\bar{x}\right)+pu(\bar{x}-c-m-\gamma\bar{x}-\alpha\bar{x})+\delta\left(1-p\right)^2u\left(\bar{x}+x_1\right)\\ &+\delta\left[1-\left(1-p\right)^2\right]u\left(\bar{x}+x_1-c-2m-\gamma\left(\bar{x}+x_1\right)-\left(\left(1+\alpha\right)^2\left(\bar{x}+x_1\right)-\left(\bar{x}+x_1\right)\right)\right)\\ >&(1-p)\,u\left(\bar{x}+x_0\right)+pu\left(\bar{x}+x_0-c-m-\gamma\left(\bar{x}+x_0\right)-\alpha\left(\bar{x}+x_0\right)\right)\\ &+\delta\left(1-p\right)^2u\left(\bar{x}\right)+\delta\left[1-\left(1-p\right)^2\right]u\left(\bar{x}-c-2m-\gamma\bar{x}-\left(\left(1+\alpha\right)^2\bar{x}-\bar{x}\right)\right)\end{aligned} \tag{23}$$

Our desired result requires that condition 23 be easier to meet than condition 22, such that acting later is more likely when both periods are in the future. We again distinguish four cases.

1. When $c > 0$, with all other cost parameters zero, our conditions become:

$$u(\bar{x}) + \delta(1-p)\,u(\bar{x}+x_1) + \delta p u(\bar{x}+x_1-c) > u(\bar{x}+x_0) + \delta(1-p)\,u(\bar{x}) + \delta p u(\bar{x}-c)$$

$$(1-p)\,u(\bar{x}) + pu(\bar{x}-c) + \delta(1-p)^2 u(\bar{x}+x_1) + \delta\left[1-(1-p)^2\right] u(\bar{x}+x_1-c)$$
$$> (1-p)\,u(\bar{x}+x_0) + pu(\bar{x}+x_0-c) + \delta(1-p)^2 u(\bar{x}) + \delta\left[1-(1-p)^2\right] u(\bar{x}-c)$$

It can be seen that the result holds regardless of the values of p, $\bar{x}-c$, and $x_1 - x_0$.

2. When $m > 0$, with all other cost parameters zero, our conditions become:

$$u(\bar{x}) + \delta(1-p)\,u(\bar{x}+x_1) + \delta p u(\bar{x}+x_1-m) > u(\bar{x}+x_0) + \delta(1-p)\,u(\bar{x}) + \delta p u(\bar{x}-m)$$

$$(1-p)\,u(\bar{x}) + pu(\bar{x}-m) + \delta(1-p)^2 u(\bar{x}+x_1) + \delta\left[1-(1-p)^2\right] u(\bar{x}+x_1-2m)$$

$$> (1-p)\,u(\bar{x}+x_0) + pu(\bar{x}+x_0-m) + \delta(1-p)^2 u(\bar{x}) + \delta\left[1-(1-p)^2\right] u(\bar{x}-2m)$$

It can be seen that the result holds when p is small.

3. When $\gamma > 0$, with all other cost parameters zero, our conditions become:

$$u(\bar{x}) + \delta(1-p)\,u(\bar{x}+x_1) + \delta p u((1-\gamma)(\bar{x}+x_1)) > u(\bar{x}+x_0) + \delta(1-p)\,u(\bar{x}) + \delta p u((1-\gamma)\,\bar{x})$$

$$(1-p)\,u(\bar{x}) + pu((1-\gamma)\,\bar{x}) + \delta(1-p)^2 u(\bar{x}+x_1) + \delta\left[1-(1-p)^2\right] u((1-\gamma)(\bar{x}+x_1))$$

$$> (1-p)\,u(\bar{x}+x_0) + pu((1-\gamma)(\bar{x}+x_0)) + \delta(1-p)^2 u(\bar{x}) + \delta\left[1-(1-p)^2\right] u((1-\gamma)\,\bar{x})$$

Given these conditions, the reuslt does not hold even for particular sets of parameter values. To see this, consider a counterexample: under log utility, both conditions simplify to $u(\bar{x}) + \delta u(\bar{x}+x_1) > u(\bar{x}+x_0) + \delta u(\bar{x})$. The conditions are equivalent, and the result does not hold.

4. When $\alpha > 0$, with all other cost parameters zero, our conditions become:

$$u(\bar{x}) + \delta(1-p)\,u(\bar{x}+x_1) + \delta p u((1-\alpha)(\bar{x}+x_1)) > u(\bar{x}+x_0) + \delta(1-p)\,u(\bar{x}) + \delta p u((1-\alpha)\,\bar{x})$$

$$(1-p)\,u(\bar{x}) + pu((1-\alpha)\,\bar{x}) + \delta(1-p)^2 u(\bar{x}+x_1) + \delta\left[1-(1-p)^2\right] u\left(\bar{x}+x_1-\left((1+\alpha)^2(\bar{x}+x_1)-(\bar{x}+x_1)\right)\right)$$

$$> (1-p)\,u(\bar{x}+x_0) + pu((1-\alpha)(\bar{x}+x_0)) + \delta(1-p)^2 u(\bar{x}) + \delta\left[1-(1-p)^2\right] u\left(\bar{x}-\left((1+\alpha)^2\bar{x}-\bar{x}\right)\right)$$

As above, we show by counterexample that the result does not hold even for particular parameter values. Under log utility, both conditions simplify to $u(\bar{x}) + \delta u(\bar{x} + x_1) > u(\bar{x} + x_0) + \delta u(\bar{x})$. The conditions are equivalent, and the result does not hold.

In sum, we see that the result always holds when $c > 0$; when $m > 0$, the result holds for small values of p; and the result does not hold when $\gamma > 0$ or $\alpha > 0$. □

4.6 Summary of results

Above we studied which of the claims described in Section 2 hold with a general formulation for the cost of keeping track, a more general formulation for the utility function, and an exponential forgetting function. We find that Claims 1 through 7 always hold, while propositions 8-10 hold under parameter-specific conditions, which are summarized in Table 1.

5 Empirical findings

In the following, we briefly discuss the findings of several field experiments conducted in Kenya bewteen October 2014 – February 2015.

5.1 Experiment 1: Pre-crastination

Design Experiment 1 aimed to test pre-crastination, i.e. whether respondents prefer to incur losses sooner rather than later. We made phone calls to 46 randomly selected respondents from the subject pool of the Busara Center for Behavioral Economics and told them that they would receive KES 100 (~USD 1.50) sent through the mobile money system *MPesa*. The purpose of this initial transfer was to give respondents an endowment from which they could draw the transfers that we requested in the next step. Once the money had been sent, respondents received another call and were offered a choice between three options. The first option required respondents to transfer KES 50 back to the experimenters on the same day; in return, they would receive a transfer of KES 500 two weeks later. The second option required respondents to transfer KES 50 back to the experiments exactly one week later; in return, they would recieve a transfer of KES 500 two weeks later. The third option offered respondents simply to keep the KES 100, not send any money back to the experiments, and not receive the KES 500. (No respondent chose this latter option.)

Thus, respondents were faced with a choice between sending KES 50 to the experimenters on the same day vs. a week later; if they did this, they would receive KES 500 two weeks later. Because the payment had to arrive on the correct day in either case, our the prediction was that sending KES 50 in one week would be subject to an additional cost of keeping track, and therefore would be less attractive than sending back KES 50 on the same day. Note that the standard discounting model predicts that respondents should unambiguously prefer to send back the KES 50 as late as possible – in this case, one week later. The timeline for the experiment is shown in Figure 2.

Results Figure 2 shows the results of the experiment. Not surprisingly, no participant chose not to send back KES 50 and forgo the KES 500 in two weeks. However, in deciding between whether to send back KES 50 on the same day or a week later, a majority ($35/46 = 76\%$) of respondents preferred to send back KES 50 on the same day instead of a week later. This result suggests that sending back KES 50 in a week is less desirable than sending back KES 50 on the same day, possibly because subjects are worried that they may forget about the transaction and thereby "lose" the KES 500 in two weeks. Recall that the standard model would predict that agents should prefer to send KES 50 next week, for two reasons: first, the transaction cost is incurred later; second, the loss is incurred later.

Note that this result cannot be explained by present bias in combination with sophistication: respondents who anticipate on the first day that choosing to send KES 50 in one week would result in them failing to send the KES 50 in one week should be even less likely to compensate for this by sending the KES 50 now. However, a more complicated alternative explanation is available: present bias in combination with both sophistication and budget constraints could potentially explain the finding. If respondents anticipate that they will spend the endowment of KES 100 in the first week and be unable to send KES 50 after a week, they might exhibit the behavior we observe. Experiment 2 aimed to control for this possibility.

5.2 Experiment 2: Pre-crastination with non-monetary outcomes

Design The results of Experiment 1 could potentially be explained by a combination of present bias, sophistication, and budget constraints, instead of a cost of keeping track. To control for this possibility, we conducted a second experiment in which the action to be performed by respondents was very low-cost, and thus any preference for performing the task sooner rather than later must be due to the costs of keeping track of the future transaction. Forty-four respondents of the subject pool of the Busara Center in Nairobi were again sent KES 100 on the first day, and were then offered the opportunity to send an SMS back to the experimenters, either on the same day or a week later. In return for sending the SMS, they would again receive KES 500 two weeks after the start of the study. Because an SMS costs KES 3, or about 1% of a daily wage, we are confident that budget constraints were not an obstacle for respondents to send the SMS at either timepoint. The standard model would again predict that respondents should prefer to perform this task as late as possible. To make sending the SMS after one week even more attractive, choosing this option additionally entailed receiving KES 50 immediately, i.e. on day 0. Conversely, to make the sending the SMS immediately less attractive, choosing this option additionally entailed receiving an additional KES 50 after one week. Thus, in this experiment, two forces should cause respondents to prefer to send the SMS late: a preference for incurring the cost of sending the SMS later rather than sooner, and a preference for receiving KES 50 sooner rather than later. The timeline of the experiment is shown in Figure 3.

Results Figure 3 shows the results of the experiment. Again not surprisingly, only one person out of 44 (2% of the sample) preferred to forgo the KES 500 two weeks later, and the extra payment of KES 50, and chose not to send an SMS on the same day or a week later. In deciding between sending the SMS on the same day and receiving KES 50 a week later vs. sending the SMS a week later and receiving KES 50 on the same day, the remainder of the participants were evenly split: 21/44 (48%) preferred to send the SMS on the same day and receive the extra KES 50 a week later, while the remaining 22 participants (50%) preferred to receive the KES 50 on the same day and send the SMS a week later. Thus, despite the fact that two factors (receiving KES 50 earlier rather than later, and being able to incur the cost of sending the SMS later rather than earlier) should have created a preference for sending the SMS later, half of participants prefer to send the SMS on the same day and receive the KES a week later. Thus, participants appear to be aware that they might forget to send the SMS a week later and thereby forgo the KES 500 two weeks later. Tthis experiment therefore demonstrates two things: first, there does appear to be an extra cost attached to future transactions; second, participants anticipate this cost, i.e. they are at least partly sophisticated.

It is natural to hypothesize that this cost attached to future transactions stems from participants' awareness of their own likelihood of forgetting to send the SMS in the future. To understand whether individuals really forget to act on future tasks, we could anlayze in the present experiment what proportion of participants actually manage to send the SMS request on the correct day; however, the groups of participants who had to send text messages on different days are self-selected, and therefore this analysis would yield biased results. The next experiment therefore aimed to quantify, with random assignment of participants to different timepoints at which they have to send a request, whether people in fact do forget about future transactions, and to what extent.

5.3 Experiment 3: The shape of the forgetting function

Design The goal of this experiment was to map the empirical forgetting function: is it true that individuals forget to act on future tasks, even when they stand to gain relatively large amounts of money by following through? If yes, to what extent? In addition, we aimed to make the *financial* cost of keeping track even closer to zero: in the previous experiment, respondents still had to pay for the SMS they had to send; here, we attempted to make this cost zero. The design of the experiment was similar to the two previous ones: 30 respondents in Nairobi were again called by phone, and received an initial transfer of KES 100. After this transfer had been made, they received another call in which they were told that they could receive an extra KES 500 five weeks later, in return for sending a "request" to the experimenters on the correct day. The request could consist of a phonecall, an SMS, or a "call me" request sent from the phone of the respondent to that of the experimenter. Importantly, this "call me" request is free of cost, and this fact was made known to participants. A total of 210 subjects were randomized into one of seven experimental conditions, with 30 subjects per condition. The conditions differed only in terms of the timepoint at which the request had to be sent to the experimenters: participants in condition 1 had to send the request *immediately*, i.e. within 5 minutes after the end of the call with the experimenter. If they managed to send the request within this time, they received the KES 500 five weeks later. In conditions 2 and 3, participants had to send the request *later on the same day* or *on the next day*, respectively, to receive the KES 500 five weeks later; sending the request at any other time would result in forfeiting the KES 500. In conditions 4–7, participants had to send the request *exactly* 1, 2, 3, or 4 weeks after the initial call. Again failing to send the request on the correct day would lead to not receiving the KES 500 five weeks after the initial call. The results from this experiment will allow plotting the "forgetting fucntion", i.e. an empirical estimate of how the probability of forgetting a task evolves over time.

Results The results of the experiment are shown in Figure 4. As we expected, the proportion of participants who send the request at the correct timepoint decreases over time: while 97% (29/30) of the participants who had to send the request immediately after the initial call successfully sent the request, this success rate dropped to 90% (27/30) when the request had to be sent later that day, and to 70% (21/30) when the request had to be sent on the following day. Thus, participants are much less likely to successfully keep track of transactions if these transactions have to be completed in the future relative to the present; this is true even when the transaction costs for completing the task are low. Together, these results suggest that participants indeed have a non-zero probability of forgetting about future transactions, despite the fact that they incur a cost for it (in this case, forfeiting the transfer of KES 500).

6 Applications

The framework described above unifies a number of anomalies that are observed in discounting behavior in the lab and the field. In addition, the framework speaks to a number of findings in development economics, which I briefly summarize here.

6.1 Chlorinating water at the source vs. the home

In many developing countries, access to clean water is difficult. Many households fetch water from a distant water source, where it is often contaminated. Purification through chlorination is relatively easy and cheap, but Kremer et al. (2009) show that chlorination levels in Kenya are low. In addition, providing households with free bottles of chlorine that they can keep in the home and use to treat water have little effect on chlorination levels. However, a slightly different intervention was much more successful: when Kremer and colleagues equipped the water source where people fetched the water with chlorine dispensers – simple containers from which individuals can release enough chlorine to treat the water they fetch at the source – the prevalence of chlorination increased dramatically. This finding can be explained in the framework described above: when anindividual is at a water source and considers whether or not to chlorinate water now – i.e. while still at the source – or later – i.e., after returning to the homestead – she previously had no real choice: chlorination was not available at the source, and "later" was the only option. In practice, once she returned to the household, she would often have forgotten her plan to chlorinate water, and would therefore not do it. In contrast, the chlorine dispenser at the source fulfills two functions. First, it reminds individuals about the chlorination and its benefits; second, it provides an opportunity to act immediately and thereby save the cost of keeping track. Thus, the model predicts that households may prefer to perform the (probably cumbersome) task of chlorinating water sooner rather than later, in the knowledge that a decision to do it later might cause it to be forgotten altogether.

6.2 Nudging farmers to use fertilizer

Developing economies are often characterized by surprisingly low levels of adoption of profitable technologies. A prime example is agricultural fertilizer, which is widely available and affordable in countries such as Kenya, and increases yields substantially. Nevertheless, few farmers report using fertilizer, even though many report *wanting* to use it. Duflo et al. (2009) show that a simple intervention can substantially increase fertilizer use: by offering timed discounts on fertilizer to farmers immediately after the harvest (when liquidity is high), rather than before the next planting season (when fertilizer would normally be bought, but liquidity is low), Duflo and colleagues provided farmers with both a reminder of their desire to use fertilizer, as well as with an opportunity to act immediately. As in the chlorine example, buying fertilizer is an expense, and therefore, assuming constant prices, standard discounting models would predict that farmers should delay it as much as possible. However, in Duflo et al.'s study, farmers pre-crastinate because they anticipate that they will fail to buy fertilizer with a non-zero probability if they delay it. This cost of keeping track is somewhat different than those discussed above – in particular, it is unlikely to stem from forgetting because the need to purchase fertilizer probably becomes ever more salient as the planting season approaches. Instead, it is likely that keeping track in this case consists of keeping track of the income from the last harvest and avoiding to spend it on other goods, with the result that nothing is left to buy fertilizer before the planting season.

6.3 Getting children vaccinated

Many children in developing countries do not receive the standard battery of vaccinations, even when these vaccinations are safe and available free of charge. Banerjee et al. (2010) organized and advertised immunization camps in 130 villages in rural India, to which mothers could bring their children to have them immunized. In a subset of villages, mothers additionally received a small incentive when they brought their children to get vaccinated. Banerjee and colleagues find that vaccination rates increase dramatically as a result of this program. Interpreted in the framework presented in this paper, we might suspect that women remember at random times that they value vaccinations and want to get their children vaccinated. However, these thoughts may often occur when no good opportunity exists to act on the thought – e.g., while performing other work, or at night. The vaccination camps might combine a reminder of the desire to get children vaccinated with a concrete opportunity to follow through on this desire.

6.4 Reminders to save

The savings rates of the poor are generally low, despite the fact that they often have disposable income that could in principle be saved. Karlan et al. (2010) show that savings rates among poor individuals in the Philippines, Peru, and Bolivia can be substantially increased through simple text message (SMS) reminders to save. This finding confirms that the poor in fact do have disposable income which they can save, and that they have a desire to do so. The fact that reminders alone can make them more successful in reaching this goal suggests that they may on occasion simply forget their savings goal and instead spend on other goods. The reminders transiently reduce the cost of keeping track to zero and thus allow households to follow through on their goal.

Note that the model predicts that reminders work because they decrease the cost of keeping track: if an individual is credibly offered a reminder, her cost of keeping track is reduced, so she is more likely to wait and successfully perform the task. However, the model also predicts that timing is crucial: if a reminder comes at a time when the agent can act on it, the probability that it is successful should be very high. In contrast, when the agent currently cannot act on it, the agent is in the previous situation of having to make a plan to act on the reminder later, making it less likely to happen because of the cost of keeping track.

7 Conclusion

This paper has argued that a number of features of empirically observed discount functions can be explained with a lump-sum cost of keeping track of future transactions. Such a cost will cause agents to discount gains more and losses less than they otherwise would; as a result, they will exhibit a sign effect in discounting, a gain-loss asymmetry in valuing future outcomes similar to that observed in loss aversion, and pre-crastination in the loss domain, i.e. a preference for incurring losses sooner rather than later. If the cost of keeping track is lump-sum, it also creates a magnitude effect in the gains domain, i.e. discounting large gains less than small gains; and a reversed magnitude in the loss domain, i.e. discounting large losses *more* than small losses. Finally, the model predicts decreasing impatience and dynamic inconsistency, and a decrease in discounting and decreasing impatience when money is added to existing payoffs, similar to that documented empirically by Andreoni & Sprenger (2012).

In addition to describing these stylized facts about temporal discounting, the model also predicts status quo bias and the choice of defaults: agents may appear to be unwilling to adopt profitable technologies or stick to disadvantageous defaults despite the presence of dominating alternatives. The model suggests that these behaviors need not reflect preferences, but either an inability to act on such tasks at the time when individuals think about them (e.g. no chlorine dispenser at the source while fetching water), or, in the case where the cost of keeping track is small enough that agents make plans to act later, the risk of forgetting to act on them (forgetting to chlorinate water in the home after returning from the water source). Finally, the model predicts that simple reminders might cause individuals to act on tasks that they previously appeared to dislike, and that reminders and/or creation of opportunities to act on tasks, such as bill payments, loan repayments, or taking medication, will increase payment reliability and adherence. Indeed, a number of studies have shown positive effects of reminders on loan repayment (e.g. Karlan et al., 2010).

A limitation of the current model is that it predicts that only sophisticates will exhibit the anomalous discounting behaviors summarized above, while naïve types will exhibit exponential discounting. To be sure, this leads to a welfare loss for the naïve type if they underestimate their probability of forgetting future tasks and therefore are more likely to incur the associated penalties; but it generates the somewhat surprising prediction that sophisticated decision-makers will appear more "anomalous" in their discounting behavior than naïve types. A further shortcoming of the model is that it does not predict procrastination on losses.

Together, these results unify a number of disparate features of empirically observed discounting behavior, as well as behavior of individuals in domains such as loan repayment, medication adherence, and technology adoption. The model makes quantitative predictions about the effectiveness of reminders, which should be experimentally tested.

References

1. Andreoni, James, and Charles Sprenger. 2012. "Estimating Time Preferences from Convex Budgets." American Economic Review 102 (7): 3333–56. doi:10.1257/aer.102.7.3333.

2. Augenblick, N., Niederle, M., & Sprenger, C. (2013). "Working Over Time: Dynamic Inconsistency in Real Effort Tasks". NBER Working Paper 18734.

3. Benzion, U., Rapoport, A. and Yagil, J. (1989): "Discount Rates Inferred from Decisions: An Experimental Study," Management Science, Vol. 35, No. 3, pp.270-284

4. Banerjee, Abhijit Vinayak, Esther Duflo, Rachel Glennerster, and Dhruva Kothari. 2010. "Improving Immunisation Coverage in Rural India: Clustered Randomised Controlled Evaluation of Immunisation Campaigns with and without Incentives." BMJ (Clinical Research Ed.) 340: c2220.

5. Duflo, Esther, Michael Kremer, and Jonathan Robinson. 2009. "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya." National Bureau of Economic Research Working Paper.

6. Ericson, K. M. M. 2013. On the Interaction of Memory and Procrastination: Implications for Reminders. NBER Working Paper.

7. Frederick, S., Loewenstein, G. and O'Donoghue, T. 2002. "Time Discounting and Time Preference: A Critical Review," Journal of Economic Literature, Vol. 40, No. 2, pp. 351-401.

8. Hardisty, David. 2011. "Temporal Discounting of Losses". PhD Dissertation, Columbia University.

9. Kahneman, D., and A. Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." Econometrica: Journal of the Econometric Society, 263–91.

10. Karlan, Dean, Margaret McConnell, Sendhil Mullainathan, and Jonathan Zinman. 2010. "Getting to the Top of Mind: How Reminders Increase Saving." National Bureau of Economic Research Working Paper.

11. Kremer M, Miguel E, Mullainathan S, Null C, Zwane A. 2009a. "Coupons, promoters, and dispensers: impact evaluations to increase water treatment." Harvard University Working Paper.

12. Laibson, David. 1997. "Golden Eggs and Hyperbolic Discounting." The Quarterly Journal of Economics 112 (2): 443–78. doi:10.1162/003355397555253.

13. Loewenstein, G. (1987): "Anticipation and the Valuation of Delayed Consumption," Economic Journal, Vol. 97, No. 387, pp.666-684

14. Loewenstein, G. (1988): “Frames of Mind in Intertemporal Choice,” Management Science, Vol. 34, No. 2, 200-214

15. Loewenstein, G. and D. Prelec (1992). Anomalies in Intertemporal Choice: Evidence and an Interpretation. Quarterly Journal of Economics 57(2), 573-598.

16. Mullainathan, S., and Eldar Shafir. Scarcity: Why Having Too Little Means So Much. New York: Times Books, 2013.

17. Noor, J. (2011). Intertemporal choice and the magnitude effect. Games and Economic Behavior 72 (1), 255-270.

18. O’Donoghue, Ted, and Matthew Rabin. 1999. “Doing It Now or Later.” American Economic Review 89 (1): 103–24. doi:10.1257/aer.89.1.103.

19. Rosenbaum, David A., Lanyun Gong, and Cory Adam Potts. 2014. “Pre-Crastination Hastening Subgoal Completion at the Expense of Extra Physical Effort.” Psychological Science, May, 0956797614532657. doi:10.1177/0956797614532657.

20. Sadoff, S., Samek, A., Sprenger, C. 2014. “Dynamic Inconsistency in Food Choice: Experimental Evidence from a Food Desert”. University of California, San Diego, Working Paper.

21. Strotz, Robert H., “Myopia and Inconsistency in Dynamic Utility Maximization,” Review of Economic Studies, XXIII (1956), 165–80.

22. Thaler, R. (1981): “Some Empirical Evidence on Dynamic Inconsistency,” Economic Letters, Vol. 8, pp.201-207

23. Wixted, John T. 2004. “On Common Ground: Jost’s (1897) Law of Forgetting and Ribot’s (1881) Law of Retrograde Amnesia.” Psychological Review 111 (4): 864–79. doi:10.1037/0033-295X.111.4.864.

Appendix A: Magnitude effect when the probability of forgetting large payoffs is smaller

Here we consider whether a cost of keeping track would still generate a magnitude effect when the probability of forgetting about large future payoffs is smaller than the probability of forgetting about smaller future payoffs. The magnitude effect requires that the discounted utility of a large payoff Ax ($A > 1$) be larger, in percentage terms relative to the undiscounted utility of the same payoff, than that of a smaller payoff x. Define $q \geq 0$ as the difference in the probability of forgetting quantity x versus Ax, such that the probability of forgetting larger amount Ax is smaller than the probability of forgetting smaller amount x. Our required condition is thus:

$$\frac{\delta\left(1-p\right)u\left(Ax\right)+\delta pu\left(Ax-C_{Ax}^{1}\right)}{u(Ax)} > \frac{\delta\left(1-p-q\right)u\left(x\right)+\delta\left(p+q\right)u\left(x-C_{x}^{1}\right)}{u(x)} \tag{24}$$

Proof. By concavity of $u(\cdot)$, and given that $A > 1$,

$$u(Ax + C_x^1) - u(Ax) < u(x + C_x^1) - u(x)$$

Because $u\left(\cdot\right)$ is monotonically increasing, $u(Ax) > u(x)$, and therefore:

$$\frac{u(Ax + C_x^1) - u(Ax)}{u(Ax)} < \frac{u(x + C_x^1) - u(x)}{u(x)}$$

Adding one on both sides and decreasing the arguments by C_x^1:

$$\frac{u(Ax)}{u(Ax - C_x^1)} < \frac{u(x)}{u(x - C_x^1)}$$

Inverting the fractions, we obtain:

$$\frac{pu(Ax - C_x^1)}{u(Ax)} > \frac{pu(x - C_x^1)}{u(x)}$$

We can add the term $\left(\frac{u\left(x-C_x^1\right)}{u(x)} - 1\right)q$ to the RHS, given that the term is negative by the strict monotonicity of $u\left(.\right)$.

$$\frac{pu\left(Ax - C_{Ax}^1\right)}{u(Ax)} > \frac{pu\left(x - C_x^1\right)}{u(x)} + \left(\frac{u\left(x - C_x^1\right)}{u(x)} - 1\right)q$$

Combining like terms and multiplying through by δ:

$$\frac{\delta pu\left(Ax - C_{Ax}^1\right)}{u(Ax)} > \frac{-\delta qu(x) + \delta\left(p+q\right)u\left(x - C_x^1\right)}{u(x)}$$

Adding $\delta(1-p)$ to both sides and combining each side into a single fraction, we obtain the desired result. $\square$

	Exponential discounting	Quasi-hyperbolic discounting	Cost of keeping track (Lump-sum cost, linear utility)	Cost of keeping track (General formulation)
Immediate utility (gains)	$u_0^+ = u(x_0)$	$u_0^+ = u(x_0)$	$u_0^+ = x_0$	$u_0^+ = u(x_0)$
Immediate utility (losses)	$u_0^- = u(-x_0)$	$u_0^- = u(-x_0)$	$u_0^- = -x_0$	$u_0^- = u(-x_0)$
Delayed utility (gains)	$u_1^+ = \delta u(x_1)$	$u_1^+ = \beta\delta u(x_1)$	$u_1^+ = \delta(x_1 - c)$	$u_1^+ = \delta(1-p)u(x_1) + \delta p u(x_1 - C_{x_1}^1)$
Delayed utility (losses)	$u_1^- = \delta u(-x_1)$	$u_1^- = \beta\delta u(-x_1)$	$u_1^- = \delta(-x_1 - c)$	$u_1^- = \delta(1-p)u(-x_1) + \delta p u(-x_1 - C_{x_1}^1)$
Results				
1. More discounting of gains	No	Yes	Yes	Yes
2. Less discounting of losses	No	No	Yes	Yes
3. Pre-crastination	No	No[a]	Yes	Yes
4. Sign effect	No	No	Yes	Yes
5. Gain-loss asymmetry	No	No	Yes	Yes
6. Magnitude effect (gains)	Yes[b]	Yes[c]	Yes	Yes
7. Reversed magnitude effect (losses)	No	No	Yes	Yes
8. Decreasing impatience	No	Yes	Yes	Yes (α small)
9. Andreoni-Sprenger 1	No	No	Yes	c: $\bar{x} - c$ small m: $\bar{x} - m$ small γ, α: $x_1 - x_0$, p large, $x_0 - \bar{x}$ small
10. Andreoni-Sprenger 2	No	No	Yes	c: Yes m: p small γ, α: No

Table 1: Summary of results and comparison of cost of keeping track model to the standard discounting and quasi-hyperbolic disocunting models.

[a] O'Donoghue & Rabin (1999) show "pre-properation" when agents are sophisticated and quasi-hyperbolic. However, this only occurs when there is a cost of acting in the future, as in the present paper.

[b] With concave utility

[c] With concave utility

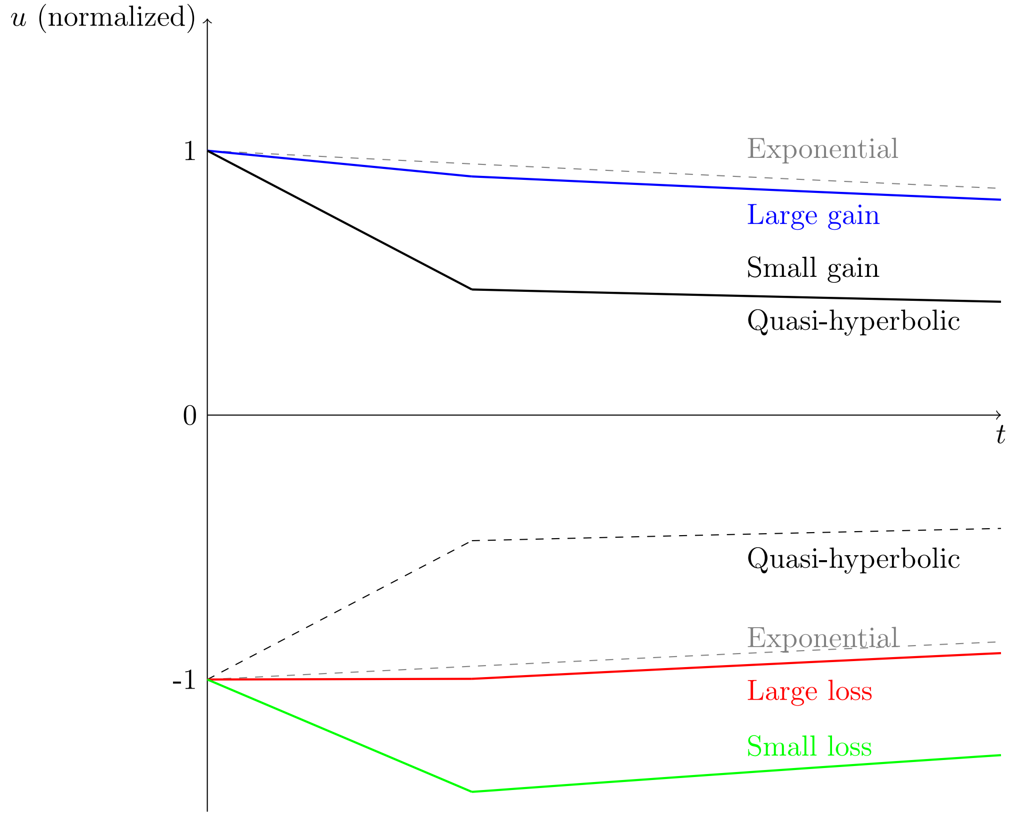
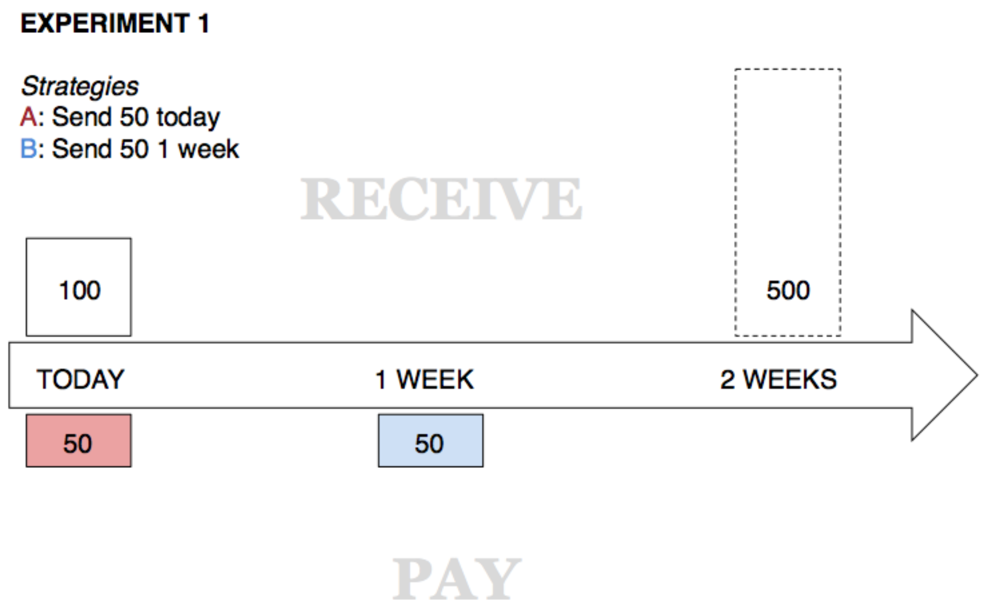


Figure 1: A lump-sum cost of keeping track of future transactions produces decreasing impatience, a magnitude effect, a sign effect, and pre-crastination. In this example, gains and losses of 10 and 100 are discounted with $\delta = 0.95$ for three periods. I plot the normalized discounted utility separately for exponential discounting (dashed gray lines), quasi-hyperbolic discounting with $\beta = 0.5$ (dashed black lines), and exponential discounting with a lump-sum cost of keeping track of $c = 5$. Note that a positive cost of keeping track leads to increased discounting for gains compared to exponential discounting: the solid blue and black lines are below the dashed gray line in the gains domain (Claim 1). Analogously, a positive cost of keeping track leads to *less* discounting of losses compared to exponential discounting: the red and green lines are below the dashed gray line in the loss domain (Claim 2). As a consequence, gains are thus discounted more than losses, leading to the sign effect in discounting (Claim 4), and a gain-loss asymmetry for future outcomes, similar to that observed in loss aversion (Claim 5). If the cost of keeping track is large enough, agents "pre-crastinate", i.e. they prefer to incur losses sooner rather than later; this is evident in the green line, which shows higher utility for incurring the loss of 5 immediately than for incurring it in any of the depitcted future periods (Claim 3). Because the cost of keeping track is a lump-sum, it is proportionally smaller for large outcomes, leading to less discounting of large than small amounts; the magnitude effect (Claim 6): the blue line for large gains lies above the black line for small gains. (Note that the converse is true in the loss domain, where small losses are discounted less than large losses, evident in the fact that the green line for small losses lies below the red line for large losses. This reversal of the magnitude effect in the loss domain has recently been documented empirically (Hardisty, 2011).) Finally, because the cost of keeping track is constant subtracted from all future outcomes, it creates a kink similar to that observed in quasi-hyperbolic discounting; as a result, agents exhibit decreasing impatience and dynamic inconsistency (8).



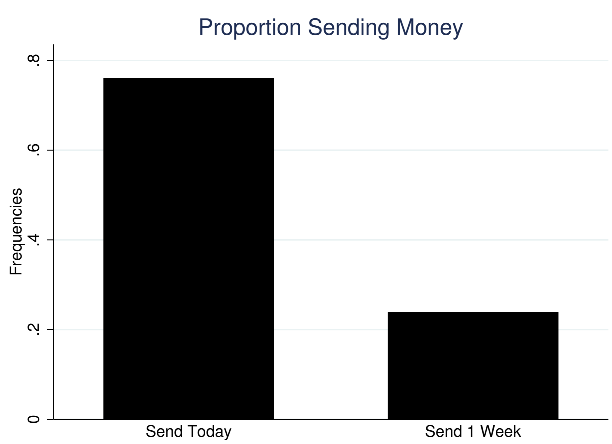


Figure 2: Timeline and results of Experiment 1. Respondents first received KES 100 through MPesa, and then had a choice between sending KES 50 back to the experimenters either on the same day or one week later. If they sent back KES 50 on the day they had chosen, they would receive a transfer of KES 500 two weeks after the initial call. The results show that a majority of respondents preferred to send back KES 50 on the same day rather than a week later.

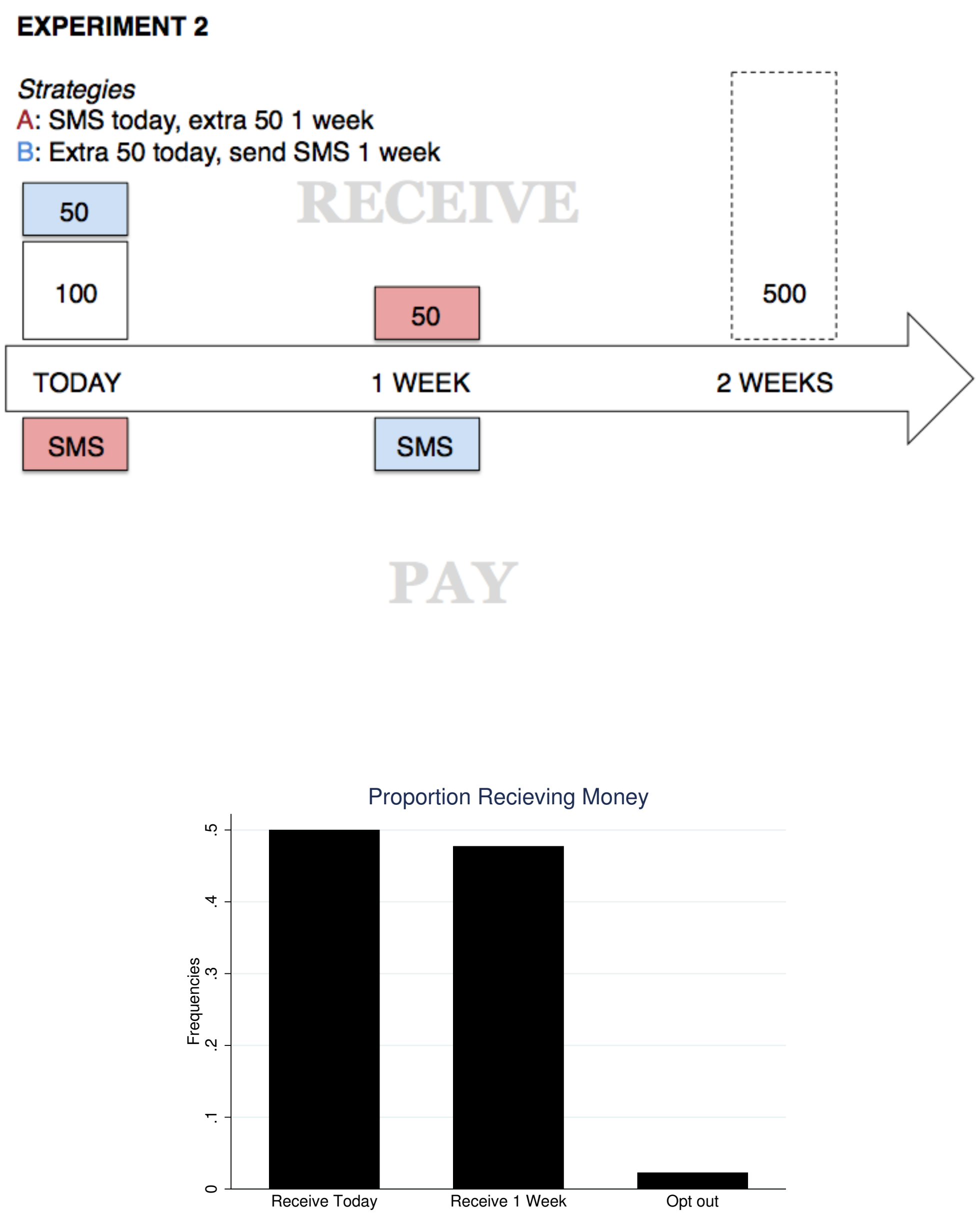
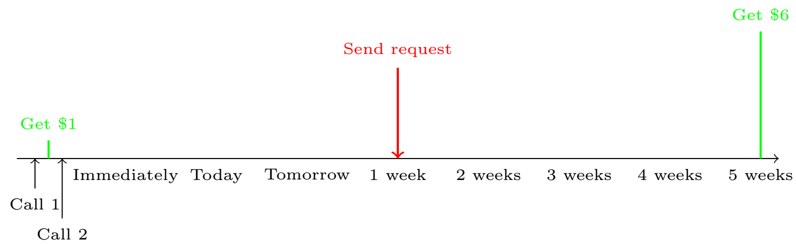
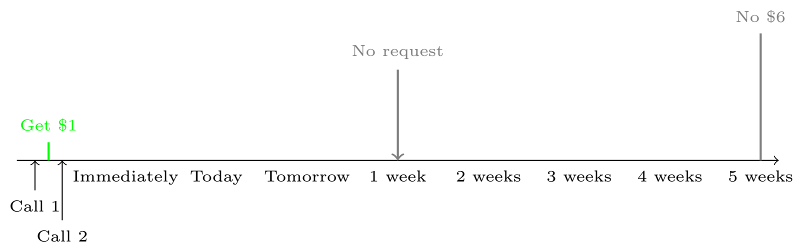


Figure 3: Timeline and results of Experiment 2. Respondents first received KES 100 through MPesa, and then had a choice between sending an SMS back to the experimenters either on the same day or one week later. If they sent the SMS on the day they had chosen, they would receive a transfer of KES 500 two weeks after the initial call. In addition, choosing to send the SMS on the first day additionally entailed receiving a payment of KES 50 after one week, while choosing to send the SMS a week later entailed receiving an additional payment of KES 50 on the same day. The results show that a majority of respondents preferred to send the SMS on the same day rather than a week later, and receive KES 50 a week later rather than on the same day.

Option 1:



Option 2:



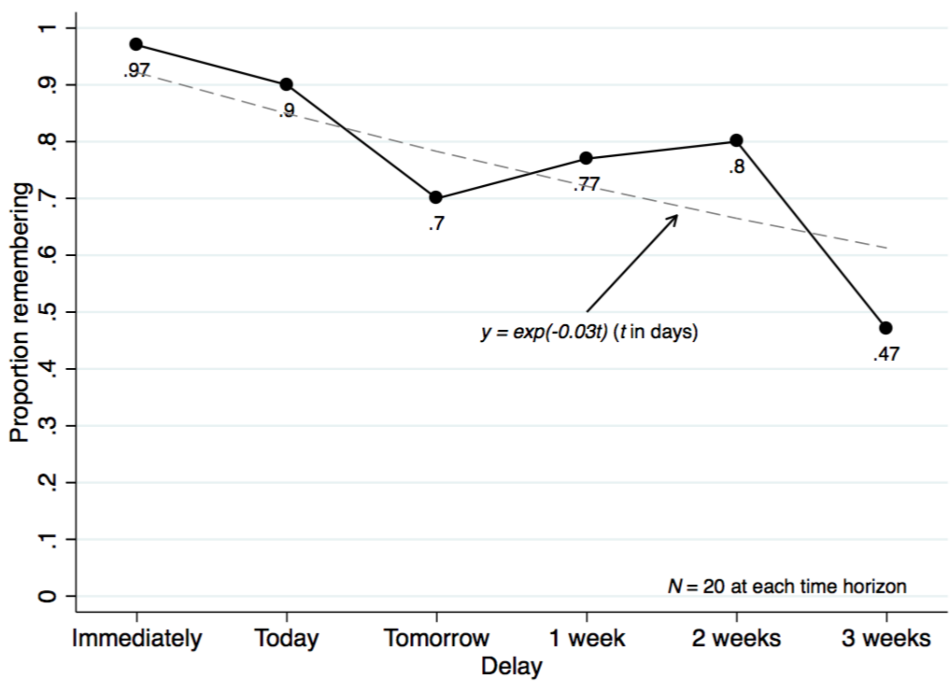


Figure 4: Timeline and results of Experiment 3. Respondents first received KES 100 through MPesa, and then had to send a request consisting either of a phonecall, an SMS, or a (free) "call me" request back back to the experimenters in order to receive a transfer of KES 500 five weeks after the initial call. In each of seven conditions, the request had to be sent at a different timepoint to qualify the respondent for the KES 500 payment: *immediately* (within 5 minutes after the initial call); *on the same day* as the initial call; *on the day after* the initial call; or *exactly 1, 2, 3, or 4 weeks after* the initial call. If they sent the SMS at the time assigned to them, participants received a transfer of KES 500 five weeks after the initial call. The results show that the proportion of participants who send the request at the correct timepoint decreases over time: while 97% (29/30) of the participants who had to send the request immediately after the initial call successfully sent the request, this success rate dropped to 90% (27/30) when the request had to be sent later that day, and to 70% (21/30) when the request had to be sent on the following day.