The Cost of Keeping Track

Johannes Haushofer[*]

First draft: August 9, 2014
This version: June 30, 2015

Abstract

I show that a lump-sum cost of keeping track of future transactions predicts several known departures from the standard discounting model: decreasing impatience with dynamic inconsistency; a magnitude effect; a reversal of this magnitude effect in the loss domain; a sign effect; and an Andreoni-Sprenger (2012) type reduction of discounting and decreasing impatience when money is added to existing payoffs. Agents of this type "pre-crastinate" on losses and are willing to pay for reminders. These results speak to failures of technology adoption in developing countries, and empirical tests conducted in Nairobi, Kenya confirm the model's predictions.

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

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

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

---

[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.

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, Loewenstein, and O'Donoghue 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 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, Gong, and Potts 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 and Rabin 1999 show that the well-known quasi-hyperbolic discounting model (Laibson 1997) predicts "pre-properation" under certain conditions. However, this occurs only when (opportunity) 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; Benzion, Rapoport, and Yagil 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 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, Rapoport, and Yagil 1989). It should be noted that a magnitude effect can also be produced with concavity of the utility function (Loewenstein and 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 it. The present model suggests one plausible source for the magnitude effect.

Seventh, Hardisty, Appelt, and Weber 2013 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 discount-

---

[2]To my knowledge, this hypothesis has not been empirically tested.

ing 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 *present bias, myopia,* or *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 all future outcomes, but not present outcomes. As such, it is similar in spirit to the quasi-hyperbolic model (Laibson 1997), 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 is added to existing payoffs. Andreoni and 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: in Andreoni & Sprenger's experiment, by the time subjects make the allocation to the sooner vs. later timepoint, they have

---

[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).

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 hyperbolicity (decreasing impatience) 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 (Laibson 1997). In the gains domain, the two models produce some similar results: if $\delta = \frac{1}{1+r}$ (where δ is the exponential discount factor in the quasi-hyperbolic model and r is the interest rate), the quasi-hyperbolic model predicts discounting at a rate higher than 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 under the quasi-hyperbolic model (as well as the standard exponential model). However, neither the quasi-hyperbolic nor the exponential model predict the Andreoni-Sprenger results of reduced discounting and decreasing impatience when money is added to future payoffs.

The most salient differences between the present model and quasi-hyperbolic discounting are in the loss domain, where a cost of keeping track model predicts a number of empirical regularities which other models, including the quasi-hyperbolic model, do not capture: less discounting of losses compared to the interest rate; precrastination; the sign effect; a gain-loss asymmetry; and a reversed magnitude effect. 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 I.

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, Niederle, and Sprenger 2015; Sadoff, Samek, and Sprenger 2015), while not finding evidence of inconsistency for tradeoffs between money (Augenblick, Niederle, and Sprenger 2015). However, note that the absence of dynamic inconsistency in the money domain observed by Augenblick et al. occurs 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. It should be noted that other augmentations of the standard discounting model do somewhat better at accounting for these stylized facts. In particular, Loewenstein and Prelec 1992 show that combining hyperbolic discounting with a value function that exhibits loss

aversion, a reference point, and concavity can predict a 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 delay-speedup asymmetry in discounting that has been described by Loewenstein 1988. However, this 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 present paper is also related to the work of Benhabib, Bisin, and Schotter 2010, who fit participants' stated indifference points for different amounts of money available sooner vs. later using a flexible discount function that incorporates exponential and hyperbolic discounting as well as a fixed cost. Benhabib et al. find that the data are best described by a fixed cost of $4 for future transactions. This finding supports the main assumption of the present paper; however, the analysis presented here goes beyond that of Benhabib et al. in a) making explicit which empirical regularities are predicted by a fixed cost, especially in the loss domain, and b) providing an explicit account for the origin of this fixed cost.

Finally, this paper is related to the work of Keith Ericson (2014), who analyzes the behavior of agents with hyperbolic preferences and imperfect memory. Ericson's analysis focuses on how present bias and imperfect memory interact to affect time preferences, and how such agents respond to reminders. The main contrast to the present paper is that I use forgetting to produce present bias, rather than assuming both; put differently, I ask not how forgetting and present bias interact, but whether forgetting can look like present bias.

The basic exposition of the model assumes that agents do not have a reminder technology at their disposal. In Section II, 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 does not have a reminder technology at her disposal and 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 II 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.

The particular form of the cost of keeping track matters for some of these results, but not for others. In Appendix A, 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).

Appendix A also 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 declines exponentially, 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 decreasing impatience does not hold when the cost of keeping track is per-period and proportional. 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.

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 may 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: 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). Thus, 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 and Rabin 1999 obtain "preproperation").

The remainder of the paper is organized as follows. Section I presents the model and derives predictions for intertemporal choice. Section II 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 III presents the design and results of several experiments conducted in Kenya which test predictions of the model. Section IV describes a few applications of the model in the developing world. Section V concludes.

I   A Simple Model for the Cost of Keeping Track

We begin by considering agents in an infinite horizon model for whom "tasks" arise at the beginning of period 0. Tasks consist either of payments to be made, or payments to be received, and are indivisible (this is one of the central differences between this model and a standard intertemporal choice model). 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 can be completed in period t' or t'', with t' and t'' exogenously given and $t'' \geq t'$. The agent either decides to act on the task in period t', or to act on it in period t''. Each task

is defined by the payoff of acting on it sooner, $x_{t'}$, and the payoff of acting on it later, $x_{t''}$, with $x_{t''} \geq x_{t'}$. These payoffs accrue in the period in which the agent acts on the task, although note that this is not required; the results 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 any future period, the transaction fails with a non-zero probability (because the agent forgets about the task, or because external factors get in the way), in which case the agent incurs a cost c. I model the forgetting process as exponential, with a per-period probability p of forgetting.[4] Agents maximize discounted expected utility, which is given by:

$$
\begin{aligned}
U_t &= \sum_{\tau=0}^{\infty} \delta^\tau \mathbb{E}\left[u(x_{t+\tau})\right] \\
&= \sum_{\tau=0}^{\infty} \delta^\tau \left[(1-p)^\tau u(x_{t+\tau}) + \left[1 - (1-p)^\tau\right] u(x_{t+\tau} - C_{x_{t+\tau}}^\tau)\right]
\end{aligned}
$$

subject to

$$
C_{x_{t+\tau}}^\tau = \begin{cases} 0, & x_{t+\tau} = 0 \\ c, & x_{t+\tau} \neq 0 \end{cases}
$$

$$
\forall t : \ x_t > 0 \iff x_{-t} = 0
$$

Here, $C_{x_{t+\tau}}^\tau$ is the cost of keeping track of a transaction for $x_{t+\tau}$ over τ periods. I begin by modeling the cost as a lump-sum cost c; Appendix A extends the framwork to other formulations. The first constraint says that agents do not incur a cost of keeping track when there is no transaction to keep track of. The second constraint is an indivisibility constraint: the agent completes the task entirely in one period. In this respect, the present setup is similar to that of O'Donoghue & Rabin (1999), and differs from standard intertemporal choice models.

To illustrate the results, we assume, without loss of generality, that $p = 1$, i.e. transactions always fail for future periods. Note that because $(1 - p)^\tau = 1$ for $\tau = 0$, transactions never fail in the current period. In addition, in what follows, I use linear utility; this choice is motivated by the fact that for the relatively small

---

[4]It will be easy to see that the results do not change when the probability is constant across time.

magnitude of the transactions which this model concerns, linear utility is a reasonable approximation. Again Appendix A 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 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. I set $t' = 0$ and $t'' = 1$ to illustrate the results.

Denoting by u_t^+ the utility of acting on a gain in period t, the utility of acting in period 0 is:

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

Analogously, the utility of acting in period 1 is:

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

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. 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:

$$(4) \qquad\qquad\qquad u_0^- = -x_0$$

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

The condition for acting in period 1 is again given by $u_1^- > u_0^-$, which simplifies to:

$$(6) \qquad\qquad\qquad x_1 < \frac{x_0}{\delta} - c$$

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.

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

Proof. 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 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.

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 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 *disutility* of future losses *increases* 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 cost of forgetting, which corresponds to more discounting for gains and less discounting for losses. □

Proposition 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.*

Proof. 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 painful 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, Gong, and Potts 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.[5] 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.

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 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.

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

---

[5]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.

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 $| u_1^- | > u_1^+$. $\qquad \square$

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.

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. $\qquad \square$

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

$$u_0 = A x_0$$

$$u_1 = \delta(A x_1 - c)$$

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. $\qquad \square$

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

Proof. 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, Appelt, and Weber 2013. $\square$

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

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:

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

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

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. $\square$

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.*

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

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

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

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

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

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

$$(7) \qquad\qquad x_1 > \frac{x_0}{\delta}$$

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. $\square$

In their study on estimating time preferences from convex budgets, Andreoni and 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 and Sprenger 2012 find much lower discount rates than most other studies on discounting. This finding is reflected in the result above.

Proposition 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.*

Proof. 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 Proposition 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.

In Appendix I, I study which of the findings outlined in this section hold with a general formulation for the cost of keeping track, a more general formulation for the utility function, and an exponential forgetting function. I find that Propositions 1 through 7 always hold, while propositions 8-10 hold under parameter-specific conditions, which are summarized in Table 1.
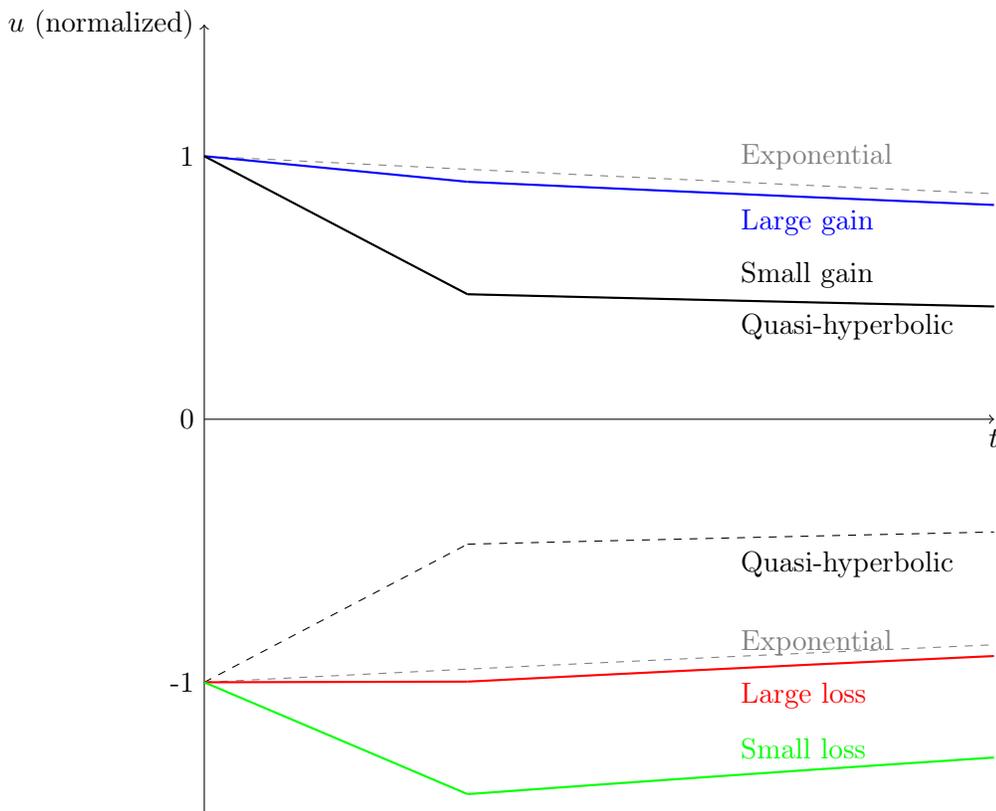
Figure 1: A lump-sum cost of keeping track of future transactions produces decreasing impatience, a magnitude effect, a sign effect, and pre-crastination. In this example, gains and losses of 10 and 100 are discounted with $\delta = 0.95$ for three periods. I plot the normalized discounted utility separately 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 (Proposition 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 (Proposition 2). As a consequence, gains are thus discounted more than losses, leading to the sign effect in discounting (Proposition 4), and a gain-loss asymmetry for future outcomes, similar to that observed in loss aversion (Proposition 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 (Proposition 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 (Proposition 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.) 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).

	Exponential discounting	Quasi-hyperbolic discounting	Cost of keeping track (Lump-sum cost, linear utility)	Cost of keeping track (General formulation)
Immediate utility (gains)	$u_0^+ = u(x_0)$	$u_0^+ = u(x_0)$	$u_0^+ = x_0$	$u_0^+ = u(x_0)$
Delayed utility (gains)	$u_1^+ = \delta u(x_1)$	$u_1^+ = \beta\delta u(x_1)$	$u_1^+ = \delta(x_1 - c)$	$u_1^+ = \delta(1-p)u(x_1) + \delta pu(x_1 - C_{x_1}^1)$
Immediate utility (losses)	$u_0^- = u(-x_0)$	$u_0^- = u(-x_0)$	$u_0^- = -x_0$	$u_0^- = u(-x_0)$
Delayed utility (losses)	$u_1^- = \delta u(-x_1)$	$u_1^- = \beta\delta u(-x_1)$	$u_1^- = \delta(-x_1 - c)$	$u_1^- = \delta(1-p)u(-x_1) + \delta pu(-x_1 - C_{x_1}^1)$
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 (lump-sum cost); no (proportional cost)
9. Andreoni-Sprenger 1	No	No	Yes	Yes (with parameter restrictions)
10. Andreoni-Sprenger 2	No	No	Yes	Yes (with parameter restrictions)

Table 1: Summary of results and comparison of cost of keeping track model to the standard discounting and quasi-hyperbolic discounting models. The general formulation of the cost of keeping track model referenced in column 4 is presented in Appendix A.

[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

II Extensions

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

A 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. 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 spend r to buy a reminder at $t = 0$ which sets the cost of keeping track of the delayed payoff to zero[6]. 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:

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

$$r < \delta c$$

---

[6]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.

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 propositions 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.

A.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 $t = 1$:

24

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

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

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. In this case, the condition for choosing the later payoff at $t = 0$ is
$$x_1 > \frac{x_0}{\delta} - \frac{1 - \delta}{\delta}c.$$

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

(8)
$$x_1 > \frac{x_0}{\delta}.$$

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

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

26

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.

A.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:

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

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

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 de-layed 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 Proposition 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 wants 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.

B 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 is 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*.[7] 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) 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.

III   Empirical findings

In the following, I briefly discuss the findings of several field experiments conducted in Kenya bewteen October 2014 – May 2015. Kenya offers a convenient setting to test behavior over time because a) the Busara Center for Behavioral Economics has established a subject pool in the Nairobi informal settlements with a high level of trust in the studies conducted there and the associated payments; c) the high penetration of mobile phones among the subject pool makes it possible to ask people to complete tasks (such as sending a text message) either immediately or in the future while keeping the cost constant; b) the mobile money system *MPesa* offers a convenient payment mechanism that allows to keep transaction costs equal between immediate and delayed transactions. The properties of MPesa have been written about extensively elsewhere (Jack and Suri 2014)); briefly, MPesa consists of an account associated with a SIM card, protected by a PIN, and transactions can be completed by text message. Clients can deposit and withdraw money from

---

[7]I follow Ericson 2014; O'Donoghue and Rabin 2006; Heidhues and Kőszegi 2010 and Gruber and Kőszegi 2004 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.

the account at any of more than 10,000 agents throughout the country, with numerous agents both in the informal settlements where our participants live, and in the vicinity of the lab.

The goal of the experiments was to test whether a) people fail to act on tasks they need to complete in the future (Experiment 1), b) they choose to pre-crastinate on losses and tasks, possibly as a result of being aware of their own probability of failing to act on future tasks (Experiments 2A and 2B).

A   Experiment 1: 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?

A total of 392 respondents in Nairobi were called by phone to announce the study, and received an initial transfer of $3 (KES 100; purchasing power parity factor 38.38[8]) through the mobile money system *MPesa*. After this transfer had been made, they received another call in which they were were given a choice between doing nothing, or sending a "request" to the experimenters at a specified future timepoint and receiving a transfer of $13 (KES 500) five weeks after the initial call. All 392 participants chose to send the request, suggesting that the transaction cost of doing so did not exceed the utility of the future transfer.

The "request" to the experimenters 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; together with a time cost of only a few seconds, this fact brought the cost of the request close to zero. The 392 respondents were randomized into one of eight experimental conditions, with an average of 49 respondents per condition (range: 46-52). 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 10 minutes after the end of the call with the experimenter. In conditions 2 and 3, participants had to send the request *later on the same day* (between noon and 5pm; initial calls were made between 9am-noon) or *on the next day* (between 9am-5pm), respectively. Finally, in conditions 4–7, participants had to send the request *exactly* on the day 1, 2, 3, 4, or 5 weeks after the initial call. Participants were informed that if they sent the request at the correct

---

[8]http://data.worldbank.org/indicator/PA.NUS.PPP

timepoint, they would receive the $13 transfer five weeks later; failing to send the request would lead to not receiving the $13. The main outcome of interest was simply whether participants in the different conditions managed to send the request on the correct day; this allows us to plot the "forgetting fucntion", i.e. an empirical estimate of how the probability of forgetting a task evolves over time[9].

Results   The results of the experiment are shown in Figure 2. Overall, 72 percent of participants successfully sent the request at the correct timepoint. As expected, the proportion of participants who send the request at the correct timepoint decreases over time: while 94 percent of the participants who had to send the request immediately after the initial call successfully sent the request, this success rate dropped to 75 percent when the request had to be sent the next day, and to 65 percent when the request had to be sent five weeks later. Thus, participants are much less likely to successfully complete 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 $13).

---

[9]Of course participants may fail to follow through for reasons other than forgetting; however, in follow-up calls with those who failed to follow through, the vast majority said that the reason for their failure was forgetting.
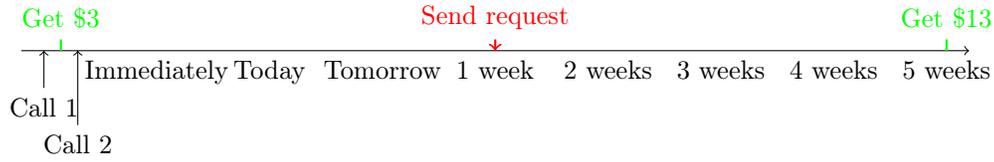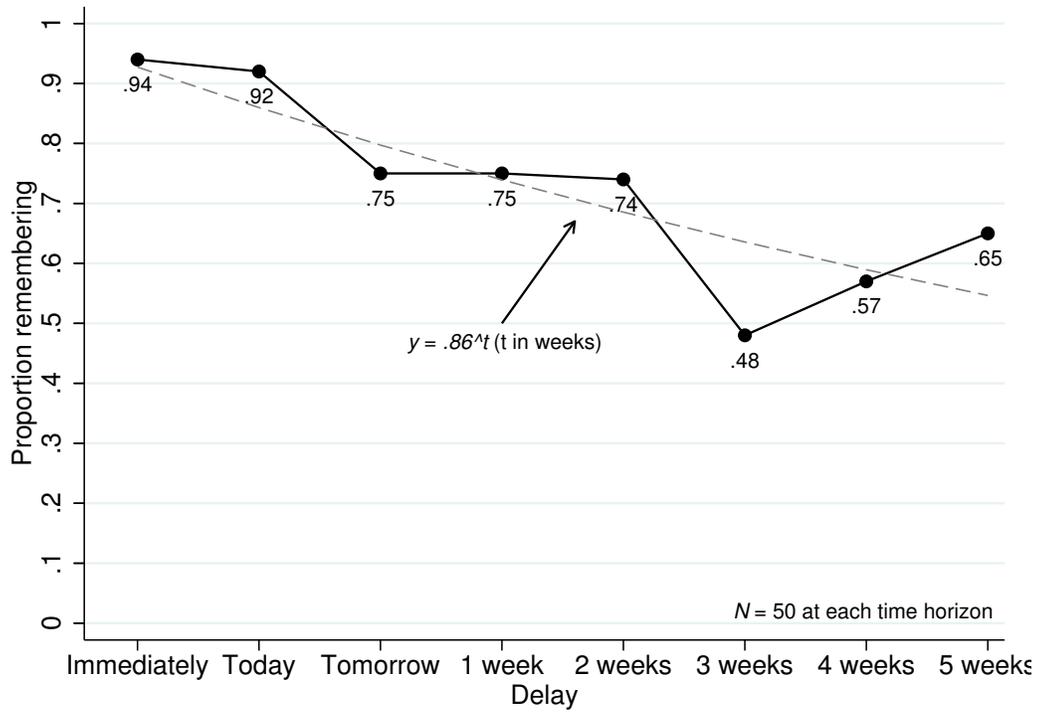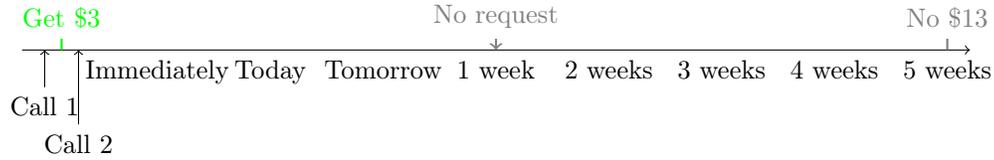
Figure 2: Timeline and results of Experiment 1. Respondents first received $3 through *MPesa*, and then had to send a request consisting either of a phonecall, an SMS, or a (free) "call me" request back back to the experimenters in order to receive a transfer of $13 five weeks after the initial call. In each of eight conditions, the request had to be sent at a different timepoint to qualify the respondent for the $13 payment: *immediately* (within 10 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, 4, or 5 weeks after* the initial call. If they sent the request at the time assigned to them, participants received a transfer of $13 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.

Option 1:

Get $3   Send $1.50 or request   Get $13

Call 1   Today   1 week   2 weeks
Call 2

Option 2:

Get $3   Send $1.50 or request   Get $13

Call 1   Today   1 week   2 weeks
Call 2

Option 3:

Get $3   Do not send $1.50 or request   Do not get $13

Call 1   Today   1 week   2 weeks
Call 2

2A Send money   **2B** Send request

Frequency

Option 1 (Send money today)	Option 2 (Send money in 1 week)	Option 1 (Send request today)	Option 1 (Send request in 1 week)
35	11	40	10
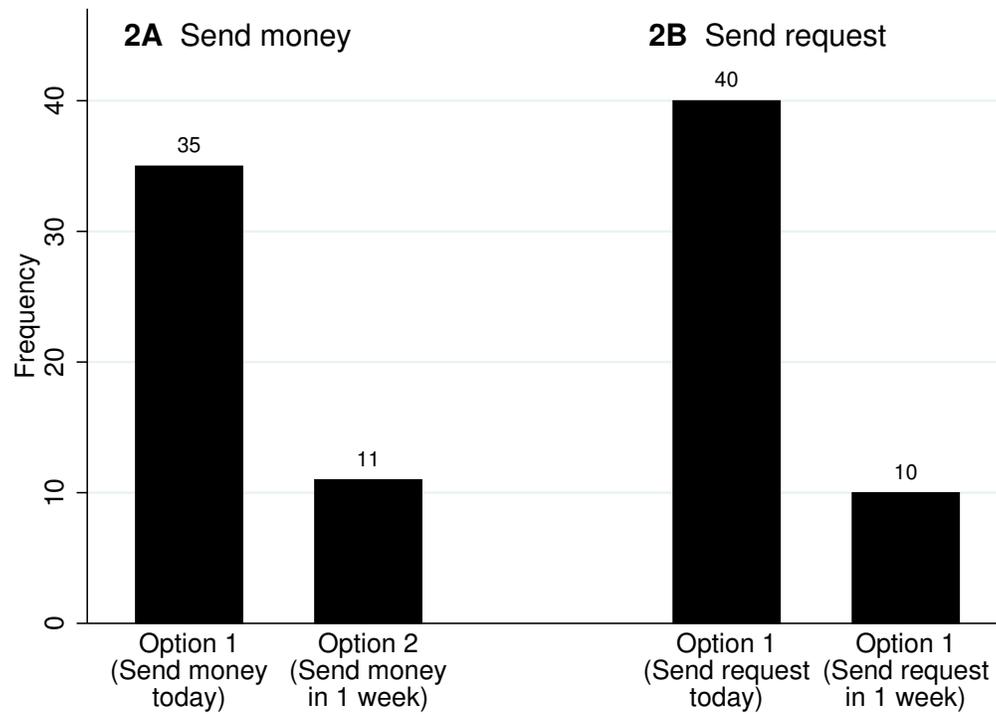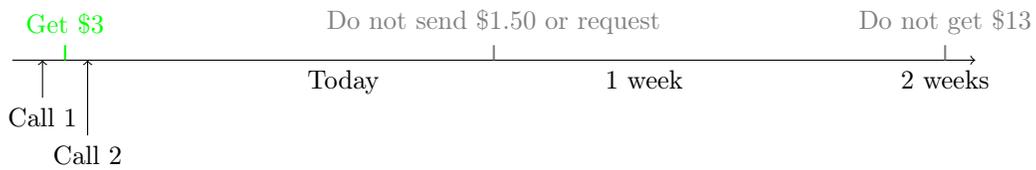
Figure 3: Timeline and results of Experiments 2A and 2B. Respondents first received $3 through *MPesa*, and then had a choice between sending $1.50 back to the experimenters either on the same day or one week later (2A), or a choice between sending a request to the experiments on the same day or one week later (2B). If they sent back $1.50 or the request on the day they had chosen, they received a transfer of $13 two weeks after the initial call. The results show that a majority of respondents preferred to send the money and the request on the same day rather than a week later.

B   Experiment 2A: Pre-crastination on losses

Design   This experiment aimed to test pre-crastination, i.e. whether respondents prefer to incur losses sooner rather than later. The timeline for the experiment is shown in Figure 3. We again made phone calls to 46 randomly selected respondents and sent them an initial transfer of $3 through *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 $1.50 back to the experimenters on the same day; in return, they would receive a transfer of $13 two weeks later. The second option required respondents to transfer $1.50 back to the experiments exactly one week later; in return, they would recieve a transfer of $13 two weeks later. The third option offered respondents simply to keep the initial $3, not send any money back to the experimenters, and not receive the $13.

No respondent chose the third option, suggesting again that sending the request was not too costly. Thus, respondents were faced with a choice between sending $1.50 to the experimenters on the same day vs. a week later; if they did this, they would receive $13 two weeks later. Because the payment had to arrive on the correct day in either case, our the prediction was that sending $1.50 in one week would be subject to an additional cost of keeping track, and therefore would be less attractive than sending back $1.50 on the same day.

Results   Figure 3 (left panel) shows the results of the experiment. Not surprisingly, no participant chose not to send back $1.50 and forgo the $13 in two weeks. However, in deciding between whether to send back $1.50 on the same day or a week later, a majority (35/46 = 76 percent) of respondents preferred to send back $1.50 on the same day instead of a week later. This result suggests that sending back $1.50 in a week is less desirable than sending back $1.50 on the same day, possibly because subjects are worried that they may forget about the transaction and thereby "lose" the $13 in two weeks. Recall that the standard model would predict that agents should prefer to send $1.50 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 $1.50 in one week would result in them procrastinating on sending the $1.50 in one week should be even less likely to compensate for this by sending the money 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 $3 in the first week and be unable to send $1.50 after a week, they might exhibit the behavior we observe. Experiment 2b aimed to control for this possibility.

C   Experiment 2B: Pre-crastination on non-monetary outcomes

Design   The results of Experiment 2A 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 instead of sending money, participants had to send a "request" as described in Experiment 1. Thus, the action to be performed 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. The timeline of the experiment is shown in Figure 3. Fifty respondents of the subject pool of Busara were again sent $3 on the first day, and were then offered the opportunity to send a request to the experimenters on the same day or a week later. In return for sending the request, they would again receive $13 two weeks after the start of the study. The standard model would again predict that respondents should prefer to perform this task as late as possible.

Results   Figure 3 (right panel) shows the results of the experiment. Again not surprisingly, no participant preferred to forgo the $13 two weeks later and chose not to send a request at all. In deciding between sending the request on the same day vs. a week later, again a majority of participants, i.e. 40/50 (80 percent), preferred to send the request on the same day. Thus, participants appear to be aware that they might forget to send the request a week later and thereby forgo the $13 two weeks later. This 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.

IV   Applications

The framework described above unifies a number of stylized facts 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.

A   Chlorinating water at the source vs. the home

In many developing countries, access to clean water is difficult. Households usually 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 has little effect on chlorination levels. However, a slightly different intervention is 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 here: when an individual 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 choice: chlorination was not available at the source, and "later" was the only option. It is likely that 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.

B   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.

C   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.

V   Conclusion

This paper has argued that a number of features of empirically observed discounting behavior 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 and 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.

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

Andreoni, James, and Charles Sprenger. 2012. "Estimating Time Preferences from Convex Budgets." *American Economic Review* 102 (7): 3333–56.

Augenblick, Ned, Muriel Niederle, and Charles Sprenger. 2015. "Working Over Time: Dynamic Inconsistency in Real Effort Tasks*." *The Quarterly Journal of Economics*, May, qjv020.

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.

Benhabib, Jess, Alberto Bisin, and Andrew Schotter. 2010. "Present-bias, quasi-hyperbolic discounting, and fixed costs." *Games and Economic Behavior* 69 (2): 205–223.

Benzion, Uri, Amnon Rapoport, and Joseph Yagil. 1989. "Discount Rates Inferred from Decisions: An Experimental Study." *Management Science* 35 (3): 270–284 (March).

Ericson, Keith M. Marzilli. 2014, August. "On the Interaction of Memory and Procrastination: Implications for Reminders." Working Paper 20381, National Bureau of Economic Research.

Frederick, Shane, George Loewenstein, and Ted O'Donoghue. 2002. "Time discounting and time preference: A critical review." *Journal of economic literature*, pp. 351–401.

Gruber, Jonathan, and Botond Kőszegi. 2004. "Tax incidence when individuals are time-inconsistent: the case of cigarette excise taxes." *Journal of Public Economics* 88 (9-10): 1959–1987.

Hardisty, David J., Kirstin C. Appelt, and Elke U. Weber. 2013. "Good or bad, we want it now: Fixed-cost present bias for gains and losses explains magnitude asymmetries in intertemporal choice." *Journal of Behavioral Decision Making* 26 (4): 348–361.

Heidhues, Paul, and Botond Kőszegi. 2010. "Exploiting Naïveté about Self-Control in the Credit Market." *American Economic Review* 100 (5): 2279–2303.

Jack, William, and Tavneet Suri. 2014. "Risk Sharing and Transactions Costs: Evidence from Kenya's Mobile Money Revolution." *American Economic Review* 104 (1): 183–223.

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*.

Kremer, Michael, Edward Miguel, Sendhil Mullainathan, Clair Null, and Alix P. Zwane. 2009. "Making water safe: Price, persuasion, peers, promoters, or product design." *Unpublished manuscript*.

Laibson, David. 1997. "Golden Eggs and Hyperbolic Discounting." *The Quarterly Journal of Economics* 112 (2): 443–478 (May).

Loewenstein, George. 1987. "Anticipation and the valuation of delayed consumption." *The Economic Journal*, pp. 666–684.

Loewenstein, George, and Drazen Prelec. 1992. "Anomalies in Intertemporal Choice: Evidence and an Interpretation." *The Quarterly Journal of Economics* 107 (2): 573–97.

Loewenstein, George F. 1988. "Frames of Mind in Intertemporal Choice." *Manage. Sci.* 34 (2): 200–214 (February).

Noor, Jawwad. 2011. "Intertemporal choice and the magnitude effect." *Games and Economic Behavior* 72 (1): 255–270 (May).

O'Donoghue, Ted, and Matthew Rabin. 1999. "Doing It Now or Later." *American Economic Review* 89 (1): 103–124 (March).

———. 2006. "Optimal sin taxes." *Journal of Public Economics* 90 (10–11): 1825–1849 (November).

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.

Sadoff, Sally, Anya Savikhin Samek, and Charles Sprenger. 2015, January. "Dynamic Inconsistency in Food Choice: Experimental Evidence from a Food Desert." SSRN Scholarly Paper ID 2572821, Social Science Research Network, Rochester, NY.

Thaler, Richard. 1981. "Some empirical evidence on dynamic inconsistency." *Economics Letters* 8 (3): 201–207.

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): 864–879 (October).

Appendix

A The Model: General formulation

The discussion in Section I 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.

A 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 ($m_t = N$).

B   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 Section I. 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 $1 - (1 - p)^t$.[10]

---

[10]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

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 $C_{x_1}^t$ of forgetting to keep track of the payment over t periods, with $C_{x_1}^t = u(x_1) - u(x_D) \geq 0$ for gains and $C_{x_1}^t = u(x_D) - u(x_1) \geq 0$ 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 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.

C Cost structure

In Section I, 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_t over t periods as follows:

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

with

$$c \geq 0$$
$$m \geq 0$$
$$\gamma \geq 0$$
$$\alpha \geq 0$$

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

---

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 Laibson 1997.

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".[11] 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).[12] 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 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).

D   Utility

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

---

[11]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.

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

above, the agent maximizes lifetime utility given by

$$U_t = \sum_{\tau=0}^{\infty} \delta^\tau \mathbb{E}\left[u(x_{t+\tau})\right]$$

$$= \sum_{\tau=0}^{\infty} \delta^\tau \left[(1-p)^\tau u(x_{t+\tau}) + \left[1 - (1-p)^\tau\right] u(x_{t+\tau} - C^\tau_{x_{t+\tau}})\right]$$

subject to

$$C^\tau_{x_{t+\tau}} = \begin{cases} 0, & x_{t+\tau} = 0 \\ c, & x_{t+\tau} \neq 0 \end{cases}$$

$$\forall t : \; x_t > 0 \iff x_{-t} = 0$$

E   Maximization

As above, In most of what follows, we again 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:

(9)
$$u_0^+ = u(x_0)$$

The utility of acting in period 1 is:

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

The condition for preferring to act in period 1 is:

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

The utility-maximizing 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:

$$(12) \qquad\qquad\qquad u_0^- = u(-x_0) = -u(x_0)$$

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

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:

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

The utility-maximizing strategy of the agent is therefore

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

F   Results

We can now formulate more general versions of the propositions in Section I.

Proposition A.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 A.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 $\frac{\partial u_1^-}{\partial C_{x_1}^1} < 0$. Thus, the disutility of future losses increases in $C_{x_1}^1$, which implies that future losses are discounted less as the cost of keeping track increases. ☐

Proposition A.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., $C_{\bar{x}}^1 > 0$), agents may prefer to act in period 0: the condition for acting in period 0 implied by 12 and 13 is $-u(\bar{x}) > -\delta(1-p)u(\bar{x}) - \delta p u(\bar{x} + C_{\bar{x}}^1)$. This expression simplifies to:

$$\frac{u\left(\bar{x} + C_{\bar{x}}^1\right)}{u\left(\bar{x}\right)} > \frac{1 - \delta(1-p)}{\delta p}$$

Because $\frac{u\left(\bar{x} + C_{\bar{x}}^1\right)}{u(\bar{x})} > 1$ 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". As an example, assume $\delta = 0.90$ and $p = 0.2$. We then obtain $\frac{u\left(\bar{x} + C_{\bar{x}}^1\right)}{u(\bar{x})} > 1.56$. Thus, if the combined disutility of the loss $\bar{x}$ and the cost of keeping track is more than 1.56 times as large as the simple disutility of the loss $\bar{x}$, agents prefer to precrastinate. A graphical representation is shown in Figure A.1. ☐

Proposition A.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

$$| u_1^- | = | -\delta (1-p) u (x_1) - \delta p u \left(x_1 + C_{x_1}^1\right) | = \delta (1-p) u (x_1) + \delta p u \left(x_1 + C_{x_1}^1\right)$$

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

Proposition A.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 A.4. □

Proposition A.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:

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

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 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:

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

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

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 A.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:

$$(16) \qquad \frac{-\delta(1-p)u(Ax) - \delta pu\left(Ax + C^1_{Ax}\right)}{-u(Ax)} < \frac{-\delta(1-p)u(x) - \delta pu\left(x + C^1_x\right)}{-u(x)}$$

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

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

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

$$u(Ax + C^1_x) - u(Ax) < u(x + C^1_x) - u(x)$$

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

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

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

$$\frac{\delta pu(Ax + C^1_x)}{u(Ax)} < \frac{\delta pu(x + C^1_x)}{u(x)}$$

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

Proposition A.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:

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

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

The condition for acting later can be written and simplified as follows:

(17)
$$u_2(x_1) > u_1(x_0)$$

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

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

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

To obtain decreasing impatience, we this condition must be 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. Recall that the original condition 11 is:

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

First note that we can set $p = 1$ without loss of generality, since any value $0 < p < 1$ simply moves the conditions towards exponential discounting, until they reduce to exponential discounting when $p = 0$. With this simplifaction, the condition when

both outcomes are in the future simplifies to:

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

Analogously, the condition when one outcome is immediate reduces to:

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

To obtain decreasing impatience, we need to show that the first of these conditions can be fulfilled while the second is not, i.e. the agent chooses the later outcome when both outcomes are in the future, but the sooner outcome when it is immediate.

We consider in turn the cases where one of the cost parameters is non-zero and all others are zero.

Case 1: $c > 0$, $m = \gamma = \alpha = 0$

Beginning with c, we set $m = \gamma = \alpha = 0$ and $c > 0$. To obtain decreasing impatience, we require that the agent chooses to wait when both outcomes are in the future:

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

At the same time, however, when one outcome is immediate, the agent chooses that outcome:

$$u(x_1 - c) < \frac{u(x_0)}{\delta}$$

First note that there exists a utility level ζ with the property $u(x_0 - c) < \zeta < u(x_0)$. Because both inequalities do not depend on the value of x_1, and $u(x_1 - c)$ can assume all values in $\Re$ due to the monotonicity of $u(\cdot)$, there exists x_1 such that $u(x_1 - c) = \zeta$, and therefore we obtain decreasing impatience.

Case 2: $m > 0$, $c = \gamma = \alpha = 0$

In this case, the condition when both outcomes are in the future simplifies to:

$$u(x_0 - m) < \delta u(x_1 - 2m)$$

The condition when one outcome is immediate is:

$$u(x_0) > \delta u(x_1 - m)$$

First note that for a given x_1, δ, and $m < x_1$, there is always a small-soon payoff x_0 such that the second condition is fulfilled (while $x_0 \leq x_1$ is preserved). Invoking Jensen's inequality, this implies that

$$\delta \left[u(x_1 - m) - u(x_1 - 2m) \right] < u(x_0) - u(x_0 - m)$$

We now designate as Δ the utility difference between the RHS and LHS terms:

$$\Delta = u(x_0) - u(x_0 - m) - \delta \left[u(x_1 - m) - u(x_1 - 2m) \right] > 0$$

Because $\Delta > 0$, there exists a scalar $0 < \xi < 1$ such that $\delta u(x_1 - m) = u(x_0) - \xi \Delta$. This implies

$$\delta u(x_1 - 2m) = u(x_0 - m) + \xi \Delta$$

which in turn, because $\xi \Delta > 0$, implies the first condition, i.e.

$$u(x_0 - m) < \delta u(x_1 - 2m).$$

Thus, decreasing impatience and dynamic inconsistency are possible with a per-period lump-sum cost.

For completeness, we briefly consider the linear utility case. When both outcomes are in the future, the condition to choose the late outcome is

$$x_0 < \delta(x_1 - 2m) + m$$

When one outcome is immediate, to condition to choose the immediate outcome is

$$x_0 > \delta(x_1 - m)$$

We again show that there exists x_0 for any x_1, δ, and $m < x_1$ such that both conditions are fulfilled. Substituting the first into the second condition, we obtain

$$\delta(x_1 - m) < x_0 < \delta(x_1 - 2m) + m,$$

which simplifies to

$$(1 - \delta)m > 0.$$

This condition is always fulfilled, implying that decreasing impatience and dynamic inconsistency are possible with a per-period lump-sum cost under linear utility.

Case 3: $\gamma > 0$, $c = m = \alpha = 0$

In this case, the condition when both outcomes are in the future simplifies to:

(19)
$$\delta u \left[(1 - \gamma)x_1\right] > u \left[(1 - \gamma)x_0\right]$$

The condition when one outcome is immediate is:

$$\delta u \left[(1 - \gamma)x_1\right] < u(x_0)$$

These conditions simplify to

$$u(x_0) > u \left[(1 - \gamma)x_0\right]$$

This condition is always fulfilled; thus, decreasing impatience and dynamic inconsistency are possible when the cost of keeping track is a one-time proportional cost. Also note that in the linear utility case, the condition simplifies to $\gamma > 0$, which is true by assumption.

Case 4: $\alpha > 0$, $c = m = \gamma = 0$

In this case, the condition when both outcomes are in the future simplifies to:

$$\delta u \left[(1 - 2\alpha - \alpha^2)x_1\right] > u \left[(1 - \alpha)x_0\right]$$

Analogously, the condition when one outcome is immediate reduces to:

$$\delta u \left[(1 - \alpha)x_1\right] < u(x_0)$$

Because of the generality of the utility function, I show by example that both conditions can be met. For instance, consider the CRRA utility function $u(x) = \frac{x^{1-\theta}-1}{1-\theta}$ with $\theta = 0.95$ (risk aversion), $\alpha = 0.1$, $\delta = 0.7$, $x_1 = 100$ and $x_0 = 26$. When both

outcomes are in the future, the utility of the larger outcome is $u(x_1)_2 = 3.4184$, while that of the smaller outome is $u(x_0)_1 = 3.4148$. Thus, the agent will choose x_1. However, when the sooner of these outcomes is available immediately, its utility is $u(x_0)_0 = 3.5385$, while that of the larger outcome is now $u(x_1)_1 = 3.5324$; thus, the agent now chooses the small-soon outcome. However, it should be noted that the conditions under which decreasing impatience can be observed when the cost of keeping track is a lump-sum proportional cost are highly restricted. Figure A.5 illustrates this point for a principal of $x_1 = 100$, CRRA utility, and $\alpha = 0.1$ (top panels) and $\alpha = 0.2$ (bottom panels). Both conditions can be fulfilled, but only by a restricted range of values of x_0 and θ.

In addition, linear utility does not yield decreasing impatience and dynamic inconsistency with a proportion per-period cost of keeping track. In the linear case, the condition where both outcomes are in the future simplifies to:

$$\frac{\delta(1 - 2\alpha - \alpha^2)x_1}{1 - \alpha} > x_0$$

The condition where one outcome is immediate simplifies to:

$$\delta(1 - \alpha)x_1 < x_0$$

Substituting, we obtain:

$$\frac{\delta(1 - 2\alpha - \alpha^2)x_1}{1 - \alpha} > x_0 > \delta(1 - \alpha)x_1$$

This condition reduces to $-1 > 1$, which is never fulfilled. Thus, with linear utility, a proportional per-period cost of keeping track will not result in decreasing impatience and dynamic inconsistency.

Overall, we establish decreasing impatience, and the associated dynamic inconsistency, when $c > 0$, $m > 0$, and $\gamma > 0$, but only under very restricted conditions when $\alpha > 0$ (with all other cost parameters zero). $\qquad\square$

Proposition A.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

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

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

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

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:

(20)
$$u(\bar{x}) + \delta\left(1 - p\right) u\left(\bar{x} + x_1\right) + \delta p u(\bar{x} + x_1 - C^1_{\bar{x}+x_1}) > u(\bar{x} + x_0) + \delta\left(1 - p\right) u\left(\bar{x}\right) + \delta p u(\bar{x} - C^1_{\bar{x}})$$

Rearranging further, we obtain:

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

We establish our result if this new condition is less strict that the original condition below:
$$(1-p) u\left(x_1\right) + pu\left(x_1 - c - m - \gamma x_1 - \alpha x_1\right) > \frac{u\left(x_0\right)}{\delta}$$

As before we can set $p = 1$ without loss of generality. We distinguish four cases:

Case 1: $c > 0$, $m = \gamma = \alpha = 0$

In this case, the conditions simplify as follows:

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

$$\delta u(x_1 - c) > u(x_0)$$

To establish the result, we require that:

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

We show again by example that this condition can be met. Consider the CRRA utility function $u(x) = \frac{x^{1-\theta}-1}{1-\theta}$ with $\theta = 0.90$ (risk aversion), $c = 2$, $\delta = 0.9$, $x_1 = 100$

and $x_0 = 80$. The agent receives a fixed payoff $\bar{x} = 5$ in each period and chooses whether to add x_0 at $t = 0$ or x_1 at $t = 1$. The agent's lifetime utility when she chooses to add x_0 at $t = 0$ is $u(\bar{x} + x_0) + \delta u(\bar{x} - c) = 6.64$. Adding x_1 at $t = 1$ generates a lifetime utility of $u(\bar{x}) + \delta u(\bar{x} + x_1 - c) = 7.05$. Without a fixed payoff $\bar{x}$ at both timepoints, the utilities would be $u(x_0) = 5.50$ and $\delta u(x_1 - c) = 5.24$. Because in this case $\delta \left[u(\bar{x}) + \delta u(\bar{x} + x_1 - c) \right] - \left[u(\bar{x} + x_0) + \delta u(\bar{x} - c) \right] > \delta u(x_1 - c) - u(x_0)$, the condition on choosing the delayed outcome is easier to fulfill, and therefore agents exhibit less discounting, when there is a fixed initial payoff compared to when there is none.

Figure A.3 illustrates this point for a principal of $x_1 = 100$, CRRA utility, and $c = 2$ (top panels). As shown in the panels on the right, agents discount less when money is added to existing payoffs across a relatively broad range of values of x_0 and θ.

Case 2: $m > 0$, $c = \gamma = \alpha = 0$

In this case, the conditions are symmetric to those of the $c > 0$ case.

Case 3: $\gamma > 0$, $c = m = \alpha = 0$

In this case, the conditions simplify to:

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

$$\delta u\left[(1 - \gamma) x_1 \right] > u(x_0)$$

To establish the result, we require that:

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

Using again CRRA utility with $\theta = 0.90$ (risk aversion), $\gamma = 0.1$, $\delta = 0.9$, $x_1 = 100$, $x_0 = 80$, and a fixed payoff $\bar{x} = 5$ in each period, we find that the agent's lifetime utility when she chooses to add x_0 at $t = 0$ is $u(\bar{x} + x_0) + \delta u\left[(1 - \gamma)\bar{x} \right] = 7.05$. Adding x_1 at $t = 1$ generates a lifetime utility of $u(\bar{x}) + \delta u\left[(1 - \gamma)(\bar{x} + x_1) \right] = 6.93$. Without a fixed payoff $\bar{x}$ at both timepoints, the utilities would be $u(x_0) = 5.50$ and $\delta u\left[(1 - \gamma)x_1 \right] = 5.11$. Because again $\delta \left[u\left[(1 - \gamma)(\bar{x} + x_1) \right] - u\left[(1 - \gamma)\bar{x} \right] \right] - \left[u(\bar{x} + x_0) - u(\bar{x}) \right] > \delta u\left[(1 - \gamma) x_1 \right] - u(x_0)$, the condition on choosing the delayed outcome is easier to fulfill, and therefore agents exhibit less discounting, when there is a fixed initial payoff compared to when there is none.

Figure A.3 illustrates this point for a principal of $x_1 = 100$, CRRA utility, and $\gamma = 0.1$ (bottom panels). As shown in the panels on the right, agents discount less

when money is added to existing payoffs across a relatively broad range of values of x_0 and θ.

Case 4: $\alpha > 0$, $c = m = \gamma = 0$

In this case, the conditions are symmetric to those of the $\gamma > 0$ case.

Thus, with a cost of keeping track, agents discount less when money is added to existing payoffs under a broad range of parameters, □

Proposition A.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:

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

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

(22)
$$
\begin{aligned}
&\delta \left(1 - p\right) u \left(\bar{x}\right) + \delta p u(\bar{x} - C^1_{\bar{x}}) + \delta^2 \left(1 - p\right)^2 u \left(\bar{x} + x_1\right) + \delta^2 \left[1 - (1 - p)^2\right] u \left(\bar{x} + x_1 - C^2_{\bar{x}+x_1}\right) \\
&> \delta \left(1 - p\right) u \left(\bar{x} + x_0\right) + \delta p u \left(\bar{x} + x_0 - C^1_{\bar{x}+x_0}\right) + \delta^2 \left(1 - p\right)^2 u \left(\bar{x}\right) + \delta^2 \left[1 - (1 - p)^2\right] u \left(\bar{x} - C^2_{\bar{x}}\right)
\end{aligned}
$$

To simplify, we again set $p = 1$ withou loss of generality. The conditions then become:

$$
(23) \quad
\begin{aligned}
&u(\bar{x}) + \delta u \left(\bar{x} + x_1 - c - m - \gamma \left(\bar{x} + x_1\right) - \alpha \left(\bar{x} + x_1\right)\right) \\
&> u \left(\bar{x} + x_0\right) + \delta u \left(\bar{x} - c - m - \gamma \bar{x} - \alpha \bar{x}\right)
\end{aligned}
$$

$$
(24) \quad
\begin{aligned}
&\delta u(\bar{x} - c - m - \gamma \bar{x} - \alpha \bar{x}) + \\
&\quad + \delta^2 u \left(\bar{x} + x_1 - c - 2m - \gamma \left(\bar{x} + x_1\right) - \left((1 + \alpha)^2 \left(\bar{x} + x_1\right) - \left(\bar{x} + x_1\right)\right)\right) \\
&> \delta u \left(\bar{x} + x_0 - c - m - \gamma \left(\bar{x} + x_0\right) - \alpha \left(\bar{x} + x_0\right)\right) \\
&\quad + \delta^2 u \left(\bar{x} - c - 2m - \gamma \bar{x} - \left((1 + \alpha)^2 \bar{x} - \bar{x}\right)\right)
\end{aligned}
$$

Without a fixed initial payoffs and $p = 1$, the conditions are as follows:

$$(25) \qquad \begin{aligned} &\delta u \left(x_1 - c - m - \gamma x_1 - \alpha \left(x_1 \right) \right) \\ &> u \left(x_0 \right) \end{aligned}$$

$$(26) \qquad \begin{aligned} &\delta^2 u \left(x_1 - c - 2m - \gamma x_1 - \left((1+\alpha)^2 x_1 - x_1 \right) \right) \\ &> \delta u \left(x_0 - c - m - \gamma x_0 - \alpha x_0 \right) \end{aligned}$$

Our desired result requires that condition 26 be easier to meet than condition 25, such that acting later is more likely when both periods are in the future and there is no fixed payoff at either timepoint; importantly, however, this effect should be reduced with a fixed initial payoff at both timepoints, i.e. the difference between conditions 24 and 23 should be smaller than that between conditions 26 and 25. Specifically, we require:

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

We again proceed case by case.

Case 1: $c > 0$, $m = \gamma = \alpha = 0$

In this case, our condition becomes:

$$\begin{aligned} \delta^2 u \left(x_1 - c \right) - \delta u \left(x_0 - c \right) - \left[\delta u \left(x_1 - c \right) - u(x_0) \right] > \\ \delta u(\bar{x} - c) + \delta^2 u(\bar{x} + x_1 - c) - \left[\delta u \left(\bar{x} + x_0 - c \right) + \delta^2 u \left(\bar{x} - c \right) \right] \\ - \left[u(\bar{x}) + \delta u(\bar{x} + x_1 - c) - \left(u(\bar{x} + x_0) + \delta u \left(\bar{x} - c \right) \right) \right] \end{aligned}$$

Simplifying:

$$\delta^2 u\left(x_1 - c\right) - \delta u\left(x_0 - c\right) - \delta u\left(x_1 - c\right) + u(x_0) >$$
$$\delta u(\bar{x} - c) + \delta^2 u(\bar{x} + x_1 - c) - \delta u\left(\bar{x} + x_0 - c\right)$$
$$-\delta^2 u\left(\bar{x} - c\right) - u(\bar{x}) - \delta u(\bar{x} + x_1 - c) + u(\bar{x} + x_0) + \delta u\left(\bar{x} - c\right)$$

We show again by example that this condition can be met. Let utility be CRRA with $\theta = 0.90$ (risk aversion), $c = 2$, $\delta = 0.9$, $x_1 = 100$, and $x_0 = 80$. With a fixed payoff $\bar{x} = 5$ in each period, the incentive to choose x_1 two periods in the future over x_0 one period in the future is $[U(0, \bar{x}, \bar{x} + x_1) - U(0, \bar{x} + x_0, \bar{x})] = -0.12057$, and the incentive to choose x_1 one period in the future over x_0 in the present is $[U(\bar{x}, \bar{x} + x_1, 0) - U(\bar{x} + x_0, \bar{x}, 0)] = 0.41391$. Thus, the agent discounts *more* in the future compared to the present; the difference in the two differences is $[U(0, \bar{x}, \bar{x} + x_1) - U(0, \bar{x} + x_0, \bar{x})] - [U(\bar{x}, \bar{x} + x_1, 0) - U(\bar{x} + x_0, \bar{x}, 0)] = -0.53448$. In contrast, without a fixed initial payoff in each period, the incentive to choose x_1 two periods in the future over x_0 one period in the future is $[U(0, 0, x_1) - U(0, x_0, 0)] = -0.20227$, and the incentive to choose x_1 one period in the future over x_0 in the present is $[U(0, x_1, 0) - U(x_0, 0, 0)] = -0.26394$. Thus, the agent discounts less in the future compared to the present; the difference in the two differences is $[U(0, 0, x_1) - U(0, x_0, 0)] - [U(0, x_1, 0) - U(x_0, 0, 0)] = 0.061666$. Together, these results imply that there is less decreasing impatience ("hyperbolicity") when money is added to existing payoffs under these parameter values – in fact, in this case there is increasing impatience.

Figure A.4 illustrates this point for a principal of $x_1 = 100$, CRRA utility, and $c = 2$ (first row). As shown in the panels on the right, agents show less decreasing impatience when money is added to existing payoffs across a relatively broad range of values of x_0 and θ.

Case 2: $m > 0$, $c = \gamma = \alpha = 0$

In this case, our condition becomes:

$$\delta u\left(x_1 - 2m\right) - u\left(x_0 - m\right) - \left[\delta u\left(x_1 - m\right) - u(x_0)\right]$$
$$> u(\bar{x} - m) + \delta u\left(\bar{x} + x_1 - 2m\right)$$
$$- \left[u\left(\bar{x} + x_0 - m\right) + \delta u\left(\bar{x} - 2m\right)\right]$$
$$- \left[u(\bar{x}) + \delta u\left(\bar{x} + x_1 - m\right) - \left(u\left(\bar{x} + x_0\right) + \delta u\left(\bar{x} - m\right)\right)\right]$$

We show again by example that this condition can be met. Let utility be CRRA with $\theta = 0.10$ (risk aversion), $m = 2$, $\delta = 0.9$, $x_1 = 100$, and $x_0 = 80$. With a fixed payoff

$\bar{x} = 5$ in each period, the incentive to choose x_1 two periods in the future over x_0 one period in the future is $[U(0, \bar{x}, \bar{x} + x_1) - U(0, \bar{x} + x_0, \bar{x})] = 5.73$, and the incentive to choose x_1 one period in the future over x_0 in the present is $[U(\bar{x}, \bar{x} + x_1, 0) - U(\bar{x} + x_0, \bar{x}, 0)] = 6.27$. Thus, the agent discounts *more* in the future compared to the present; the difference in the two differences is $[U(0, \bar{x}, \bar{x} + x_1) - U(0, \bar{x} + x_0, \bar{x})] - [U(\bar{x}, \bar{x} + x_1, 0) - U(\bar{x} + x_0, \bar{x}, 0)] = -0.54$. In contrast, without a fixed initial payoff in each period, the incentive to choose x_1 two periods in the future over x_0 one period in the future is $[U(0, 0, x_1) - U(0, x_0, 0)] = 4.38$, and the incentive to choose x_1 one period in the future over x_0 in the present is $[U(0, x_1, 0) - U(x_0, 0, 0)] = 4.72$. Thus, the agent discounts more in the future compared to the present, but to a lesser extent than in the case with fixed initial payoffs; the difference in the two differences is $[U(0, 0, x_1) - U(0, x_0, 0)] - [U(0, x_1, 0) - U(x_0, 0, 0)] = -0.33$. Together, these results imply that there is less decreasing impatience ("hyperbolicity") when money is added to existing payoffs under these parameter values.

Figure A.4 illustrates this point for a principal of $x_1 = 100$, CRRA utility, and $m = 2$ (second row). As shown in the panels on the right, agents show less decreasing impatience when money is added to existing payoffs; however, this effect is restricted to a relatively narrow range of values of x_0 and θ.

Case 3: $\gamma > 0$, $c = m = \alpha = 0$

In this case, our condition becomes:

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

We show again by example that this condition can be met. Let utility be CRRA with $\theta = 0.10$ (risk aversion), $\gamma = 0.1$, $\delta = 0.9$, $x_1 = 100$, and $x_0 = 80$. With a fixed payoff $\bar{x} = 5$ in each period, the incentive to choose x_1 two periods in the future over x_0 one period in the future is $[U(0, \bar{x}, \bar{x} + x_1) - U(0, \bar{x} + x_0, \bar{x})] = -0.08$, and the incentive to choose x_1 one period in the future over x_0 in the present is $[U(\bar{x}, \bar{x} + x_1, 0) - U(\bar{x} + x_0, \bar{x}, 0)] = 0.01$. Thus, the agent discounts *more* in the future compared to the present; the difference in the two differences is $[U(0, \bar{x}, \bar{x} + x_1) - U(0, \bar{x} + x_0, \bar{x})] - [U(\bar{x}, \bar{x} + x_1, 0) - U(\bar{x} + x_0, \bar{x}, 0)] = -0.09$. In contrast, without a fixed initial payoff in each period, the incentive to choose x_1 two periods in the future over x_0 one period in the future is $[U(0, 0, x_1) - U(0, x_0, 0)] = -0.19$, and the incentive to choose x_1

one period in the future over x_0 in the present is $[U(0, x_1, 0) - U(x_0, 0, 0)] = -0.38$. Thus, the agent discounts less in the future compared to the present; the difference in the two differences is $[U(0, 0, x_1) - U(0, x_0, 0)] - [U(0, x_1, 0) - U(x_0, 0, 0)] = 0.18$. Together, these results imply that there is less decreasing impatience ("hyperbolicity") when money is added to existing payoffs under these parameter values.

Figure A.4 illustrates this point for a principal of $x_1 = 100$, CRRA utility, and $\gamma = 0.1$ (third row). As shown in the panels on the right, agents show less decreasing impatience when money is added to existing payoffs across a relatively broad range of values of x_0 and θ.

Case 4: $\alpha > 0$, $c = m = \gamma = 0$

In this case, our condition becomes:

$$
\begin{aligned}
&\delta^2 u\left((1 - 2\alpha - \alpha^2)x_1\right) - \delta u\left((1 - \alpha)x_0\right) - \left[\delta u\left((1 - \alpha)x_1\right) - u(x_0)\right] \\
>&\delta u((1 - \alpha)\bar{x}) + \delta^2 u\left((1 - 2\alpha - \alpha^2(\bar{x} + x_1)\right) \\
&- \left[\delta u\left((1 - \alpha)(\bar{x} + x_0)\right) + \delta^2 u\left((1 - 2\alpha - \alpha^2)\bar{x}\right)\right] \\
&- \left[u(\bar{x}) + \delta u\left((1 - \alpha)(\bar{x} + x_1)\right) - \left(u\left(\bar{x} + x_0\right) + \delta u\left((1 - \alpha)\bar{x}\right)\right)\right]
\end{aligned}
$$

We show again by example that this condition can be met. Let utility be CRRA with $\theta = 0.10$ (risk aversion), $\alpha = 0.1$, $\delta = 0.9$, $x_1 = 100$, and $x_0 = 80$. With a fixed payoff $\bar{x} = 5$ in each period, the incentive to choose x_1 two periods in the future over x_0 one period in the future is $[U(0, \bar{x}, \bar{x} + x_1) - U(0, \bar{x} + x_0, \bar{x})] = -0.11$, and the incentive to choose x_1 one period in the future over x_0 in the present is $[U(\bar{x}, \bar{x} + x_1, 0) - U(\bar{x} + x_0, \bar{x}, 0)] = -0.12$. Thus, the agent discounts slightly *more* in the future compared to the present; the difference in the two differences is $[U(0, \bar{x}, \bar{x} + x_1) - U(0, \bar{x} + x_0, \bar{x})] - [U(\bar{x}, \bar{x} + x_1, 0) - U(\bar{x} + x_0, \bar{x}, 0)] = 0.005$. In contrast, without a fixed initial payoff in each period, the incentive to choose x_1 two periods in the future over x_0 one period in the future is $[U(0, 0, x_1) - U(0, x_0, 0)] = -0.36$, and the incentive to choose x_1 one period in the future over x_0 in the present is $[U(0, x_1, 0) - U(x_0, 0, 0)] = -0.38$. Thus, the agent discounts less in the future compared to the present; the difference in the two differences is $[U(0, 0, x_1) - U(0, x_0, 0)] - [U(0, x_1, 0) - U(x_0, 0, 0)] = 0.02$. Together, these results imply that there is less decreasing impatience ("hyperbolicity") when money is added to existing payoffs under these parameter values.

Figure A.4 illustrates this point for a principal of $x_1 = 100$, CRRA utility, and $\gamma = 0.1$ (fourth row). As shown in the panels on the right, agents show less decreasing impatience when money is added to existing payoffs across a relatively broad range of values of x_0 and θ. $\qquad\square$

G Summary of results

Above we studied which of the propositions described in Section I 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 propositions 1 through 7 always hold, while propositions 8-10 hold under parameter-specific conditions, which are summarized in Table 1.

B 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:
(27)
$$\frac{\delta (1 - p) u (Ax) + \delta p u \left(Ax - C_{Ax}^1\right)}{u(Ax)} > \frac{\delta (1 - p - q) u (x) + \delta (p + q) u \left(x - C_x^1\right)}{u(x)}$$

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

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

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 $\delta\left(1 - p\right)$ to both sides and combining each side into a single fraction, we obtain the desired result. $\qquad\square$
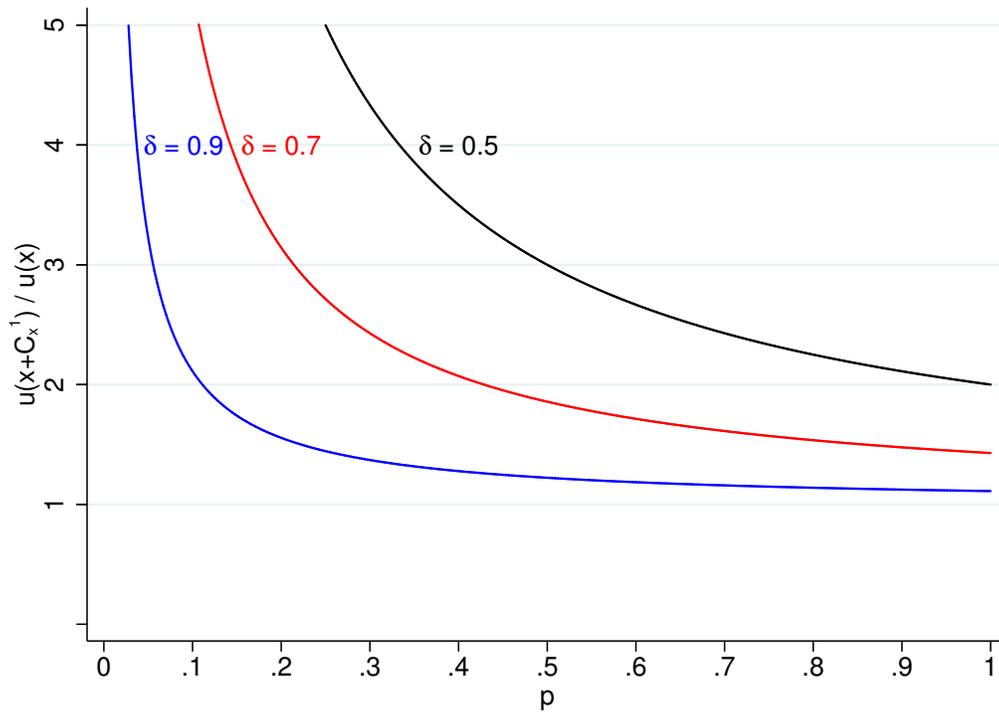
Figure A.1: Pre-crastination in the loss domain. The lines shows the values of $\frac{u\left(\bar{x}+C_{\bar{x}}^{1}\right)}{u(\bar{x})}$, i.e. ratio of the combined disutility of the loss $\bar{x}$ and the cost of keeping track, to the simple disutility of the loss $\bar{x}$, above which agents prefer to pre-crastinate for a given δ and p.
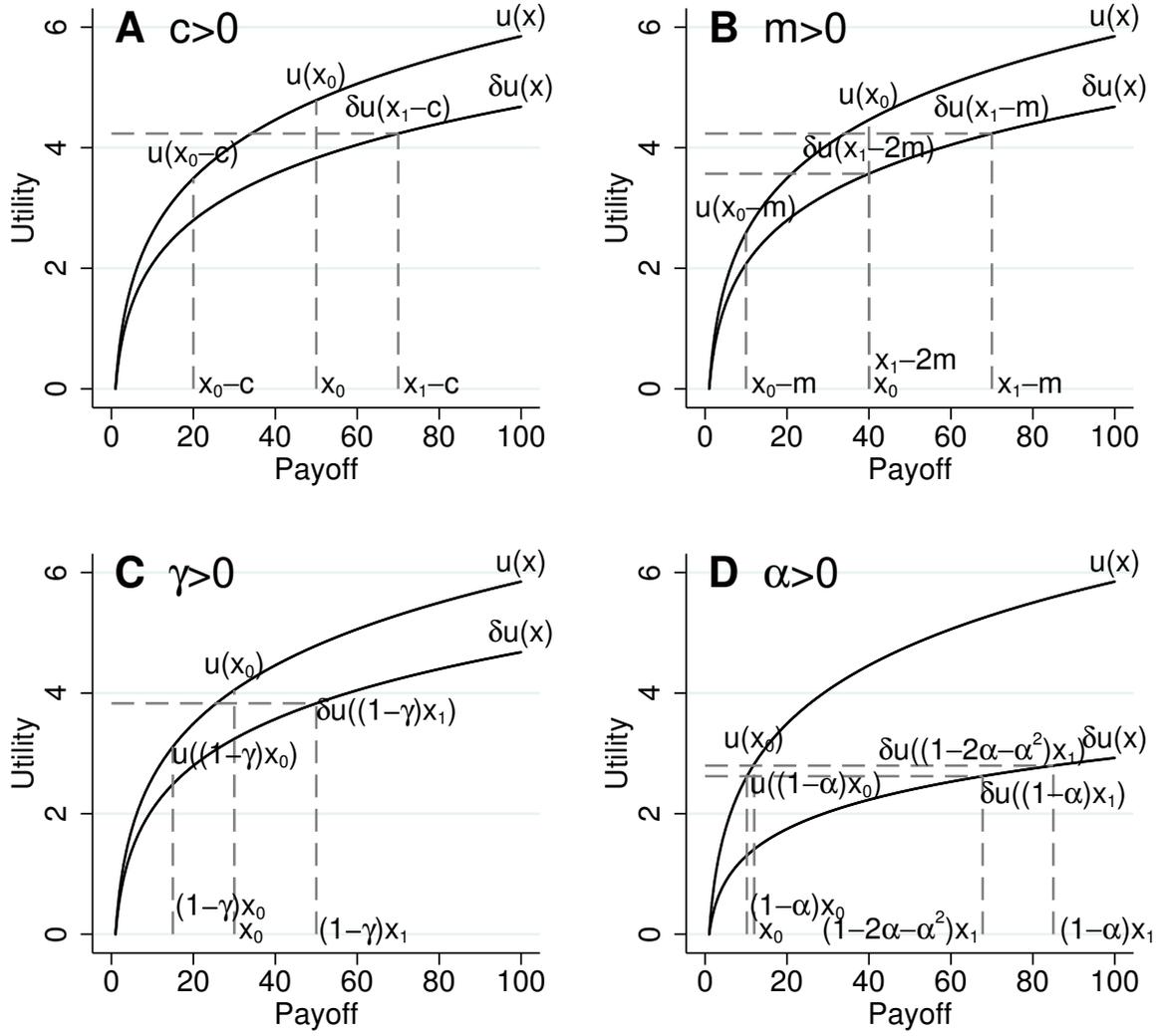
Figure A.2: Decreasing impatience and dynamic inconsistency with a cost of keeping track, illustrated on the utility function. In all panels, utility is CRRA, i.e. $u(x) = \frac{x^{1-\theta}-1}{1-\theta}$, with $\theta = 0.9$, and the prinicpal is $x_1 = 100$. The four panels show decreasing impatience with a one-time lump-sum cost c of keeping track (Panel A), a per-period lump cost m (Panel B), a proportional one-time cost γx_1 (Panel C), and a proportional per-period cost αx_1 (Panel D). In Panel A, $x_0 = 50$, $\delta = 0.8$, $c = 30$, and all other cost parameters are zero. In Panel B, $x_0 = 40$, $\delta = 0.8$, $m = 30$, and all other cost parameters are zero. In Panel C, $x_0 = 30$, $\delta = 0.8$, $\gamma = 0.5$, and all other cost parameters are zero. In Panel D, $x_0 = 12$, $\delta = 0.5$, $\alpha = 0.15$, and all other cost parameters are zero. In Panel A, the agent will choose the immediate payoff x_0 over a delayed payoff x_1 because $u(x_0) > \delta u(x_1 - c)$. However, when x_0 is not immediate, the agent will choose x_1 over x_0 when x_0 because $\delta u(x_1 - c) > u(x_0 - c)$. Analogously, in Panel B, we have $u(x_0) > \delta u(x_1 - m)$, i.e. the agent will choose the immediate over the delayed outcome, but when both outcomes are in the future, the agent will choose to wait because $\delta u(x_1 - 2m) > u(x_0 - m)$. Similarly, in Panel C, we have $u(x_0) > \delta u\left[(1-\gamma)x_1\right] > u\left[(1-\gamma)x_0\right]$ and therefore agents prefer the small-soon outcome when it is immediate, but not when it is delayed. The same is true in Panel D, where $u(x_0) > \delta u\left[(1-\alpha)x_1\right]$ but $\delta u\left[(1 - 2\alpha - \alpha^2)x_1\right] > u\left[(1-\alpha)x_0\right]$. However, in this last case, the conditions under which decreasing impatience is possible are very limited (cf. Figure A.5).
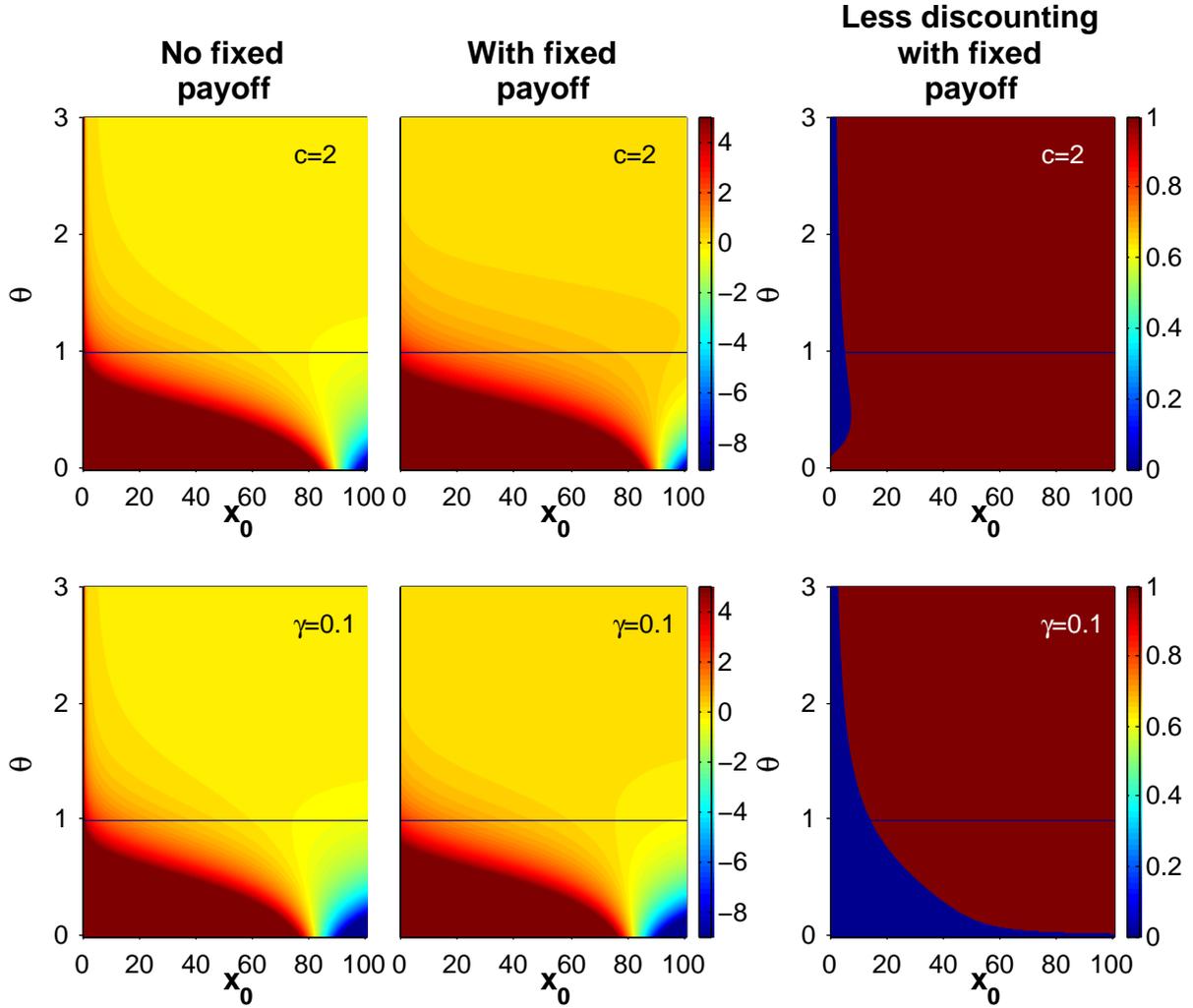
Figure A.3: Less discounting when money is added to existing payoffs with a one-time or per-period lump-sum cost of keeping track (top panels), and a one-time or per-period proportional cost of keeping track (bottom panels). Utility is CRRA, i.e. $u(x) = \frac{x^{1-\theta}-1}{1-\theta}$, the prinicpal is $x_1 = 100$, and the cost of keeping track parameter is either $c = 2$ (top panels) or $\gamma = 0.1$ (bottom panels). The leftmost panels show, for different values of $x_0 \in [0, 100]$ and $\theta \in [0, 3]$, the difference between the utility of an immediate payoff relative to a payoff tomorrow without a fixed payoff, i.e. $U(0, x_1) - U(x_0, 0)$. The middle panels show, for different values of $x_0 \in [0, 100]$ and $\theta \in [0, 3]$, the same difference *with* a fixed payoff, i.e. $U(\bar{x}, \bar{x} + x_1) - U(\bar{x} + x_0, \bar{x})$. The rightmost panels show the region where agents discount less when a fixed payoff is present compared to when there is none, i.e. $U(\bar{x}, \bar{x} + x_1) - U(\bar{x} + x_0, \bar{x}) > U(0, x_1) - U(x_0, 0)$.
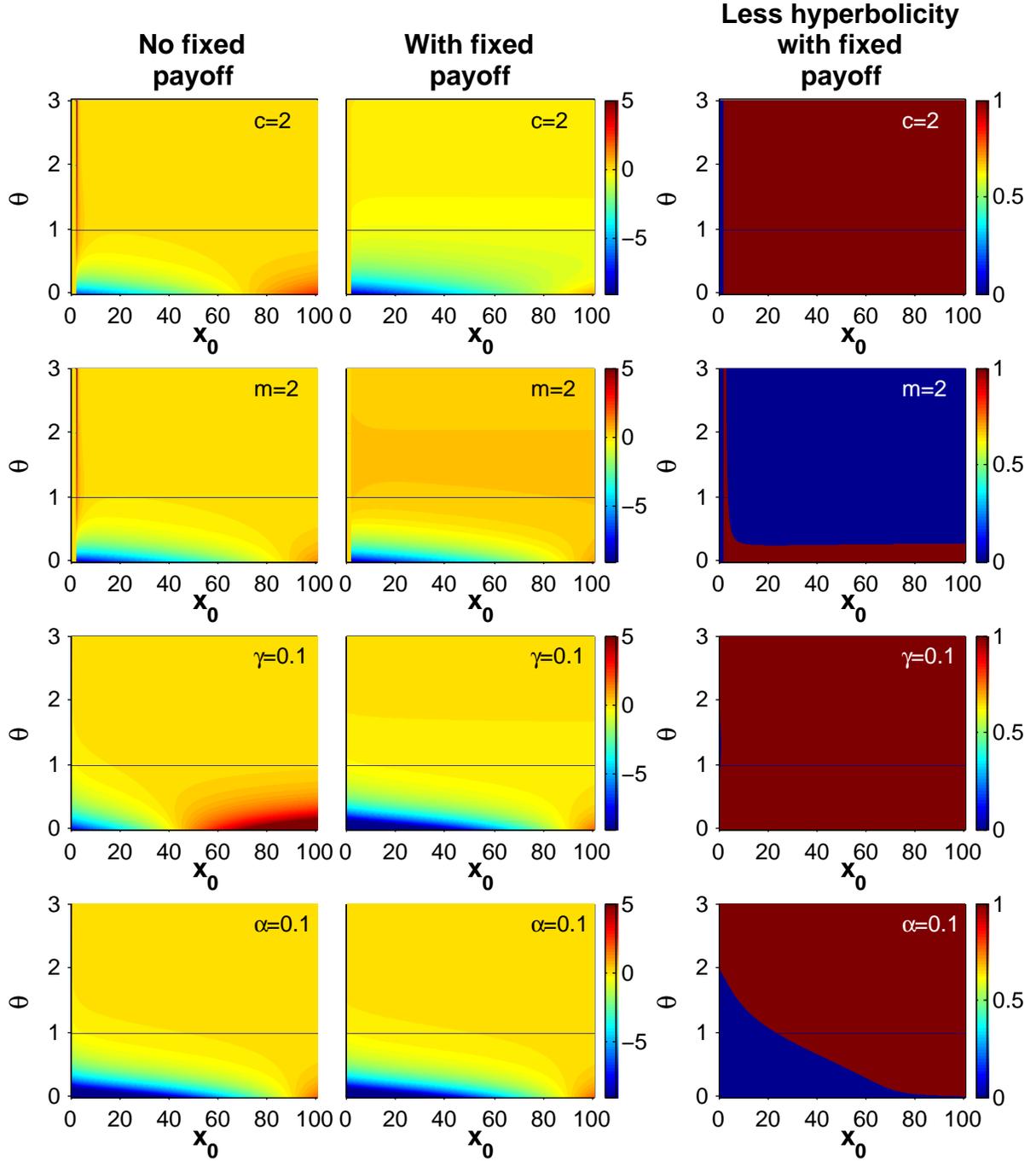
Figure A.4: Lower decreasing impatience ("hyperbolicity") when money is added to existing payoffs. Utility is CRRA, i.e. $u(x) = \frac{x^{1-\theta}-1}{1-\theta}$, the prinicpal is $x_1 = 100$, and the cost of keeping track paramater is either $c = 2$ (one-time lump-sum cost; first row), $m = 2$ (per-period lump-sum cost; second row), $\gamma = 0.1$ (one-time proportional cost; third row), or $\alpha = 0.1$ (per-period proportional cost; fourth row), with all other cost parameters zero. The leftmost graphs show, for different values of $x_0 \in [0, 100]$ and $\theta \in [0, 3]$, the difference in discounting over the near vs. the distant future when there are no fixed initial payoffs at both timepoints, $[U(0, 0, x_1) - U(0, x_0, 0)] - [U(0, x_1, 0) - U(x_0, 0, 0)]$. The middle graphs show the same difference with fixed initial payoffs, $[U(0, \bar{x}, \bar{x} + x_1) - U(0, \bar{x} + x_0, \bar{x})] - [U(\bar{x}, \bar{x} + x_1, 0) - U(\bar{x} + x_0, \bar{x}, 0)]$. The right graphs show the region where decreasing impatience is greater without fixed initial payoffs, i.e. $[U(0, \bar{x}, \bar{x} + x_1) - U(0, \bar{x} + x_0, \bar{x})] - [U(\bar{x}, \bar{x} + x_1, 0) - U(\bar{x} + x_0, \bar{x}, 0)] > [U(0, 0, x_1) - U(0, x_0, 0)] - [U(0, x_1, 0) - U(x_0, 0, 0)]$.
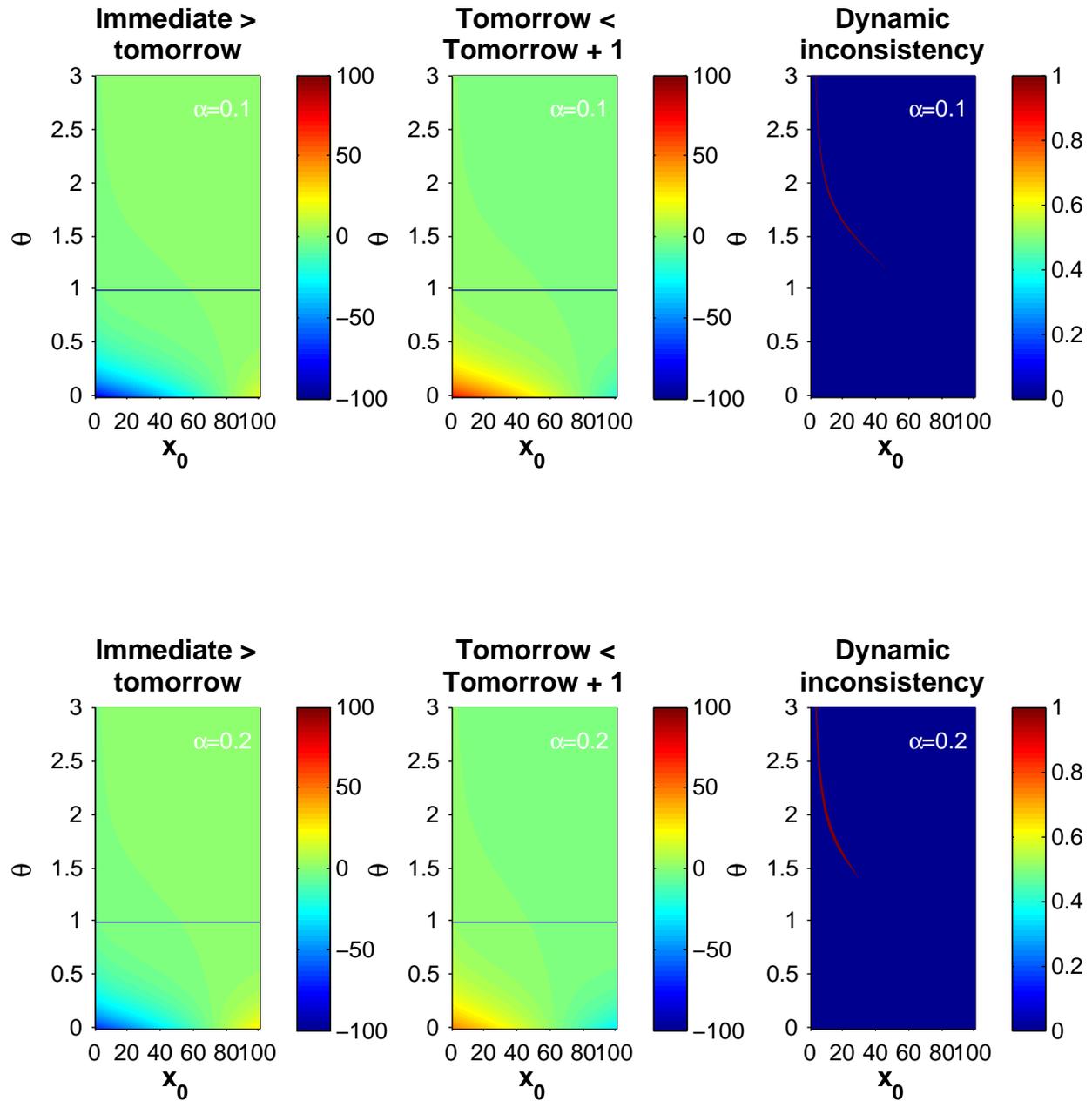
Figure A.5: Decreasing impatience and dynamic inconsistency with a proportional, per-period cost of keeping track. Utility is CRRA, i.e. $u(x) = \frac{x^{1-\theta}-1}{1-\theta}$, the prinicpal is $x_1 = 100$, and the cost of keeping track parameter is either $\alpha = 0.1$ (top panels) or $\alpha = 0.2$ (bottom panels). The graphs show, for different values of $x_0 \in [0, 100]$ and $\theta \in [0, 3]$, the difference between the utility of a delayed relative to an immediate payoff (left panels), the difference between a delayed and a less delayed payoff where neither payoff is immediate (middle panels), and the region where agents will prefer immediate payoffs over payoffs tomorrow, but will also prefer payoffs the day after tomorrow over payoffs tomorrow, i.e. dynamic inconsistency. It can be seen that dyanmic inconsistency is possible with this utility function and these parameters, but only under very restricted conditions.