Projection Bias in Effort Choices

Marc Kaufmann*

January 16, 2017

In progress. For latest version please see:
http://scholar.harvard.edu/mkaufmann

Abstract

Working becomes harder as we grow tired or bored. I model individuals who underestimate changes in marginal disutility – as implied by "projection bias" – when deciding whether or not to continue working. This bias leads to two mistakes. First, they are too pessimistic when they are tired and working is hard, and too optimistic when rested and work is easy. When within-day disutility is convex and individuals face a single task with all-or-nothing rewards (such as passing or failing a test), they initially underestimate the total disutility and start some overly ambitious tasks. As work becomes harder, they perceive the task as less worth completing and may quit. If the deadline for the task is far in the future, such individuals may repeatedly start working, yet quit earlier than anticipated. No matter how small the bias is, this can lead to large daily welfare losses. When tasks instead have concave rewards, including piece rates, such individuals work optimally if facing only a single task. But when working on multiple such tasks (for example, studying for two tests with continuous grades), they may mis-prioritize the tasks. In particular, they over-prioritize urgent tasks over important but non-urgent tasks, overestimating how much they will later work on the non-urgent tasks. Second, when tasks can be completed across multiple days, individuals smooth work too little over time. Because they underestimate how much the marginal disutility will increase on better days, they work too much on those days, and overreact to daily differences in opportunity costs, incentives, and productivity.

1 Introduction

Our tastes fluctuate, often rapidly: we grow tired and thirsty from running, and we savor food or crave coffee more the longer we go without. Our perceptions of our tastes are biased toward

*Harvard University, Department of Economics. Email: mkaufmann@fas.harvard.edu. I am grateful to my advisors Matthew Rabin, David Laibson, Gautam Rao, and Josh Schwartzstein. I also thank Ben Bushong, Krishna Dasaratha, Anastassia Fedyk, Tristan Gagnon-Bartsch, James Hodson, Annie Liang, Ben Lockwood, Neil Thakral, Linh T. Tô, and Gal Wettstein, as well as seminar participants at Harvard's economics department for helpful comments.

our current tastes: we misperceive future tastes as being closer to our current tastes than they will be – a tendency labeled *projection bias* by Loewenstein, O'Donoghue, and Rabin (2003).[1] This misperception can trigger undesirable and unintended habits and behaviors, such as buying too much when shopping on an empty stomach, or becoming addicted due to under-appreciating the future intensity of cravings. In this paper, I study effort choices where the distaste for work fluctuates, such as when students grow bored of studying or employees become tired of working. Due to projection bias, individuals mispredict future disutility of work, which can cause them to work on the wrong tasks, mis-prioritize between tasks, and inefficiently choose when to work on those tasks.

I describe the model in section 2. In the model, a person with projection bias either works or does not work at every moment of the day and is self-directed: she works if and only if at that moment she perceives it as worthwhile to continue working.[2] Whether she decides to work depends on her actual disutility, her future disutility as she perceives it, and the benefits she faces, each of which I describe in turn. Her instantaneous disutility changes over the course of the day: letting S denote the total time she has worked so far that day, instantaneous disutility is equal to $D'(S)$, where $D(\cdot)$ is the total *daily* disutility. Unless stated otherwise, the marginal disutility $D'(\cdot)$ increases – the person grows more tired and bored the longer she works. In this setting, projection bias implies that a person projects her current marginal disutility: she predicts that her marginal disutility after E hours of work on any day lies between her current marginal disutility, $D'(S)$, and her true marginal disutility, $D'(E)$. Finally, every task she faces has either decreasing (or constant) marginal benefits or all-or-nothing benefits, in which case she receives known, fixed benefits if she completes the task by the end of a given day.

A projection-biased person makes two mistakes. First, when marginal disutility is particularly low, she is too optimistic about how easy it will be to work, while she is too pessimistic when marginal disutility is particularly high. She therefore misperceives how unpleasant tasks are and mispredicts how much she will work later. She starts some overly costly all-or-nothing tasks, but as work becomes more unpleasant, she grows less optimistic and may give up – which, in this case, is correct. Second, she underestimates how much her marginal disutility

[1] Evidence for projection bias has been found for food (Read and Van Leeuwen (1998); Nordgren, Pligt, and Harreveld (2008)), drink (Van Boven and Loewenstein (2003)), sexual arousal (Loewenstein, Nagin, and Paternoster (1997); Ariely and Loewenstein (2006)), effortful tasks (Augenblick and Rabin (2016)), heroin substitute cravings (Badger et al. (2007)), the endowment effect (Loewenstein and Adler (1995)) and for predictions of gym attendance (Acland and Levy (2015)). Projection bias resembles immune neglect (Gilbert et al. (1998)) whereby people overestimate how long they will feel bad about negative events.

[2] I discuss irreversible choices in passing. Since a projection-biased person has no desire to commit, there is no reason why a particular plan should overrule her current plan – except in situations where a principal might demand that she commit.

increases the longer she works, she overreacts to differences in incentives, opportunity costs, and productivity across days. She ends up working more than she should on productive days, and less than she should on unproductive days. When working on an all-or-nothing task, at the start of productive days she underestimates how much she will work, while she overestimates it at the start of unproductive days.

I analyze the first mistake, that people fluctuate between being optimistic and pessimistic, in sections 3 and 4. In section 3, I consider different reward structures when all the tasks are due at the end of the first day. Because a person grows tired the longer she works, she underestimates how unpleasant the work will be later that day and overestimates how much longer she will work. With all-or-nothing rewards, a projection-biased person starts some overly ambitious tasks, only to abandon them later. With decreasing returns to effort, she works optimally. She works until the marginal benefits equal her marginal disutility, and then stops. But there is a catch: she works optimally only if she faces a *single* such task. If she has to allocate her time across multiple tasks, or distinct subtasks, she mis-prioritizes. Consider a student who has an exam the next day and who studies chapter 1 before switching to chapter 2. When the time comes when she should optimally switch, she overestimates how much longer she will study and therefore she keeps going on chapter 1. She ends up studying too much for chapter 1 – done when she is rested – and too little on chapter 2 – done when she is tired. This cannot happen if the student constantly switches to do the work with the highest marginal benefit. But even minor switching costs are enough to keep the student from doing so, since she doesn't realize her mistake. Thus, when multi-tasking, she works too much on early stages of tasks, such as planning or background research, and on time-sensitive tasks which are best done immediately, such as urgent requests from colleagues.

In section 4, I consider tasks that are due on some future day. The person is overly optimistic at the start of each day and potentially too pessimistic at the end. Consider a person working on an all-or-nothing task, say a student who has to study at least 100 hours to pass an exam. She may start each day planning to complete the task efficiently, yet stop studying earlier than anticipated, planning to drop the exam. This leads to one of two outcomes. Either she passes the exam despite working too little in the early days, but has to study harder later on. Or she repeatedly wastes time studying toward an exam that she eventually drops for good. Unless the benefits are high enough for the student that she actually passes the exam, higher benefits make her worse off, because they lead her to waste more time on more days for no change in the outcome. I show that even for a relatively unbiased student these repeated changes in plans can lead to a lot of wasted effort – up to the point where she does almost all the work necessary to pass the exam, yet just fails to do so.

The findings in sections 3 and 4 predict that people who work on tasks with decreasing returns to effort overestimate how well they will complete a task, consistent with research showing that people are often overly optimistic in predicting their own task completion. For instance, Buehler, Griffin, and Ross (1994) find that students believe that they will finish their bachelor's thesis earlier than they actually do. They explain this mistake – called the planning fallacy – in terms of students being overly optimistic about how many hours are necessary for the task or how many distractions they will face. Projection bias provides an alternative and complementary explanation, suggesting that people are overoptimistic about how much time they will spend working.[3] Even when they correctly estimate the work required, projection-biased students will spend too much time on early stages of tasks, such as planning and generating new ideas, as well as on time-sensitive tasks, such as problem sets, lectures, or sports. They don't realize that these tasks crowd out work on important long-term tasks, such as preparing for an exam. Moreover, when they are tired, they don't appreciate how much time they will fritter away on less important but time-sensitive tasks once they are rested and optimistic. Therefore they think that on future days – unlike in the past – they will spend more time on the long-term task and fail to realize how much work they should do today.

In section 5, I turn to the second mistake whereby projection bias leads people to smooth work too little over time. Consider a student who has two days to finish her term paper. A friend has offered to give feedback on her draft at the end of the first day, so that she should work more on the first day than on the second. Due to projection bias, the student underestimates the difference between the marginal disutility at the end of the first day and the marginal disutility at the end of the second day, and therefore she works too much on the first day. More generally, because she underestimates the differences in marginal disutility, she is prone to working too much when her marginal disutility is high and effectively overreacts to differences in incentives, opportunity costs, or productivity.

Moreover, the student mispredicts at the start of the day how much she will actually work that day. The more tired she is, the more unpleasant she perceives every additional hour of work, and the more valuable she finds a reduction in total work time. Thus she perceives the feedback her friend provides as more useful at the end of the day than she does at the start of the day, and therefore she wants to take advantage of it more than she initially thought. If instead the student were more productive on the second day, she would work less on the first day than she initially planned, in order to take advantage of her higher productivity on the second day. The situation is similar when the student is uncertain about how long she

[3] See Buehler, Griffin, and Peetz (2010) for a review of planning fallacy and of situations where people are overoptimistic about when they will complete a task.

needs to work and learns about this as time goes on. Work done early on is more likely to be unnecessary than work she does on a future day when she is better informed, thus leading her to delay too much work when uncertain.

Given how important it is whether a person is optimistic or pessimistic, in section 6 I study tasks where the marginal disutility is initially *decreasing*. Such tasks become easier initially, such as physical exercise and music practice that become easier after some warm-up, as well as mental tasks that require a high focus, such as programming or writing. Mirroring results in section 3, a person now *overestimates* the disutility initially and may fail to start a worthwhile task. And once she gets going, work is comparatively easy and she underestimates the disutility of resuming the task after a break. She therefore takes too many breaks, and either incurs higher disutility than expected once she resumes the task, or fails to resume the task.

My paper is most closely related to Loewenstein, O'Donoghue, and Rabin (2003) who formalize the model of projection bias and apply it to durable goods consumption, the endowment effect, and habit formation. My focus is on effort choices when people's plans repeatedly fluctuate so that their final behavior is the aggregate outcome from decisions made at different times that can be mutually inconsistent.[4] I also show that when people's tastes fluctuate repeatedly, the mistakes people make can lead to large welfare losses, thus highlighting that projection bias matters also outside domains where people experience large swings in taste – such as addiction – or in binding choices where even small proportional mistakes lead to large monetary losses – such as purchases of cars or homes. Thus the paper also relates to the growing literature identifying projection bias in the field (Conlin, O'Donoghue, and Vogelsang (2007); M. Levy (2009); Busse et al. (2015); Buchheim and Kolaska (forthcoming)), highlighting how and when projection matters in effort choices and thus facilitating empirical research in this economically important domain.

2 Projection Bias in Simple Effort Choices

In this section I show how both an unbiased and a biased person choose in a setting where their daily utility U is given by

$$U(E) := B(E) - D(E)$$

[4] A paper that also considers these distributed choices (as they call it) is Herrnstein and Prelec (1991), who analyze the implications of melioration.

where E is the possible number of hours worked. Here, $B(.)$ denotes the daily benefit and $D(.)$ the daily disutility. The person wants to maximize her utility. Usually, we think of this as a static problem, where the person chooses how much to work and then works that much. This abstracts away from the fact that people work over several hours and thus could decide something else later in the day than what they decided earlier on. Of course, with an unbiased person, in the absence of unanticipated information, the behavior is time-consistent, and modeling the dynamic nature of the problem adds nothing. Since projection-biased people need not have consistent plans, they may want to act differently after 3 hours of work, say, than they thought they would initially. For this reason, this section describes a dynamic setup that simply boils down to the maximization of U for an unbiased person, and extends it to allow for projection bias in a unique way when B and D are specified.

2.1 A Model of Dynamic Effort Choices Within a Day

The time in the day, $\tau \in [0, \infty)$ is continuous.[5] Each moment, the person either works or does not work, so that $e(\tau) \in \{0, 1\}$. She does not choose to work at a higher or lower intensity.[6] The instantaneous disutility from not working is 0, while the instantaneous disutility from working is $d(s(\tau))$. The latter depends on the current state, $s(\tau)$, which captures tiredness or boredom or anything else that affects how unpleasant effort is at time τ. I assume moreover that $s(\tau)$ depends only on the total amount of effort completed until time τ, which is therefore equal to τ. The interpretation is that the person works continuously without interruptions, and that once she stops working, she doesn't start again.

The final assumption is that people decide each moment whether or not they want to work that moment – they cannot choose at an earlier time whether they will work at a later time. I call such people *self-directed*:

Definition 1. *A person is* self-directed *if she does not make irreversible plans and in each moment work according to the plan she currently perceives as optimal.*

This situation is quite common, certainly at the day-to-day level. Most college students choose their classes, decide whether to attend lectures, and study when they want to. Whenever employees are not monitored around the clock – that is, whenever moral hazard is an issue – they have some lee-way about when and how much they work. This excludes situations

[5] One can also work with $\tau \in [0, \bar{\tau}]$ for finite $\bar{\tau}$.

[6] If I actually add something about intensity later, mention this here in a footnote. Otherwise get rid of this footnote.

where, for exogenous reasons, a person has to make an irreversible decision to work for the next 10 hours, for instance.[7]

So, how does the person actually decide whether or not to work each instant? Given that she is self-directed, the person continues working if she perceives that better than to stop *at that time*. Formally, let $E^*(\tau)$ denote the optimal amount of work a person plans to do at time τ. Then the person continues working at time τ, that is $e^*(\tau) = 1$, if and only if $E^*(\tau) > \tau$ – if she thinks she should work more than she has worked so far. Thus $E^*(\tau)$ solves

$$E^*_\tau = \arg\max_{E: E \geq \tau} B(E) - \int_\tau^E d(s(\tau'))d\tau'$$

where I implicitly assume that benefits in a day depend only on total effort exerted E. The total disutility from effort from working E hours in a day is

$$D(E) := \int_0^E d(s(\tau))d\tau = \int_0^E d(s(E'))dE' = \int_0^E (d \circ s)(E')dE' \tag{1}$$

Thus we have that $d \circ s = D'$, since D is the integral of $d \circ s$. Therefore the maximization problem the person solves becomes:[8]

$$E^*_\tau(\tau) = \arg\max_{E: E \geq \tau} B(E) - (D(E) - D(\tau))$$

When there is no uncertainty and the person is unbiased, then $E^*_\tau = E^*$ and we might as well ignore the dynamic nature of the problem. But, as we will now see, it is necessary to state it when people are projection-biased.

2.2 Projection Bias

Loewenstein, O'Donoghue, and Rabin (2003) define projection bias as follows. Let $d(e, s)$ be the instantaneous disutility a person experiences from exerting effort e when in state s. Suppose that a projection-biased person is currently in state s and she predicts how unpleasant effort would be if she were in state s'. Then she misperceives the disutility of exerting effort e' in state s' as lying between the actual disutility $d(e', s')$ and the disutility $d(e', s)$ that she

[7] In order to see what a person does who is self-directed, I will have to talk about the plans she makes at any time. These plans are effectively the choice she would make at that time if she had to make an irreversible decision. Thus, while I focus on the situation when a person is self-directed, in doing so I will also answer what choices a person would make when not self-directed.

[8] This assumes that the person hasn't already stopped working, in which case she won't restart again (by assumption) and there is nothing to decide.

would experience if she exerted the effort right now. The perceived disutility is denoted by $\tilde{d}(e', s'|s)$ and is given by

$$\tilde{d}(e', s'|s) := (1 - \alpha)d(e', s') + \alpha d(e', s) \tag{2}$$

where $\alpha \in [0, 1]$ is the degree of projection bias.[9] When $\alpha = 0$, the person has no projection bias and perceives future disutility correctly; when $\alpha = 1$ she has *full* projection bias and believes that future disutility is equal to current disutility of effort. Whenever there is ambiguity for $\alpha = 1$ – since such a person perceives her disutility as linear – I treat $\alpha = 1$ as the limit of α going to 1.

Thus, projection bias captures two features of people's perception of their future disutility. First, people understand that if they grow more tired that they will enjoy working less. They never think that they will want to work even more if only they were more tired. Second, people underestimate how much less they will want to work once they are more tired. Those studies that ask people to choose in at least two different states for at least two different future states reject both no projection bias and full projection bias. For instance Read and Van Leeuwen (1998) show that people choose to receive the more filling snack (a chocolate bar) over the less filling one (a fruit) when they will receive the snack at 4pm when they will be hungry rather than at 1pm after lunch. But they also choose to receive the more filling snack more when they are *currently* hungry. Fisher and Rangel (2013) in another experiment where people bid on food found that the over- and underbidding was symmetric: participants who were satiated bid less for food on a second day where they were hungry than on that day itself, and participants who were hungry bid more for food on the second day where they were satiated than on that day itself.

We can now map the definition of projection bias to the current framework. Effort e is either 0 or 1. The disutility of not working, of exerting effort $e = 0$, is 0 no matter what the state is. Therefore the perceived disutility from not working is also 0. Similarly, we have that the perceived disutility of working in a future state s'' when currently in state s', written $\tilde{d}(s''|s')$ is

[9] Specifically, this is what Loewenstein, O'Donoghue, and Rabin (2003) call *simple* projection bias. More general versions of projection bias could allow for α to depend both on the current and the future state which, for instance, would permit people to misperceive more when they are hungry than when they are sated, or vice versa. While there is little evidence on details of the structure of projection bias, Read and Van Leeuwen (1998) require people to choose both in craving and sated states for future craving and sated states. Assuming that people make the correct prediction when they are in the same state as they will be next week, they find that people project their current state whether they are sated or hungry. Moreover, Fisher and Rangel (2013) find that projection bias is symmetric in food choices.

$$\tilde{d}(s''|s') = (1-\alpha)d(s'') + \alpha d(s') \tag{3}$$

The goal is to rewrite this in terms of daily disutility D and its marginal D', and to replace the potential states s'' and s' by potential amounts of work a person may have completed at different times. In order to do this, let us define the following: $\tilde{d}_{|s'}(s'') := \tilde{d}(s''|s')$ and $\tilde{D}_{|E'}(E) := \int_0^E \tilde{d}_{|s(E')}(s(\tau))d\tau$, where $s(E')$ is the state a person is in after working E' hours. Then $\tilde{d}_{|s'}$ is the *perceived* instantaneous disutility function when the current state is s' and $\tilde{D}_{|E'}$ is the perceived daily disutility of working a given number of hours when the person has already worked E' hours so far that day. These are the instantaneous and daily disutility as the projection-biased person perceives them, and highlights the analogy to the unbiased case. Thus, just as $d \circ s = D'$, we have that $\tilde{d}_{|s'} \circ s = \tilde{D}'_{|E'}$ for all E' such that $s(E') = s'$. Finally, if s'' and s' are possible states, then let E'' and E' denote effort levels such that $s'' = s(E'')$ and $s' = s(E')$ – which must exist, since states by assumption only depend on total effort worked up to the current time.

From equation (3) we obtain the following:

$$\begin{aligned}
\tilde{d}(s''|s') = (1-\alpha)d(s'') + \alpha d(s') &\implies \tilde{d}_{|s'}(s(E'')) = (1-\alpha)d(s(E'')) + \alpha d(s(E')) && (4)\\
&\implies (\tilde{d}_{|s'} \circ s)(E'') = (1-\alpha)(d \circ s)(E'') + \alpha(d \circ s)(E') && (5)\\
&\implies \tilde{D}'_{|E'}(E'') = (1-\alpha)D'(E'') + \alpha D'(E') && (6)
\end{aligned}$$

We can integrate this to obtain

$$\tilde{D}_{E'}(E) = \int_0^E \tilde{D}'_{E'}(E'')dE'' = (1-\alpha)D(E) + \alpha D'(E') \cdot E$$

To make the notation consistent with the notation in Loewenstein, O'Donoghue, and Rabin (2003) and to highlight the difference between the role played by E'' and E' above, I write this as follows:

$$\tilde{D}'(E|s) = (1-\alpha)D'(E) + \alpha D'(s)$$

I call this special case *projecting marginal disutility* and use it throughout the paper.

Observation 1 (Projecting Marginal Disutility)**.** *Suppose the following holds in effort choices during a single period:*

1. The time in a period is continuous ($\tau \in [0, \infty)$);

2. each moment, a person either works or does not work ($e_\tau \in \{0, 1\}$);

3. once a person stops working, she doesn't start again, and each moment she is in state $s(\tau)$ that depends only on total work done until time τ

4. the instantaneous disutility from working is 0 from not working, $d(s(\tau))$ from working.

Then the total disutility of working E hours in a row is given by $D(E)$ with $D'(E) = (d \circ s)(E)$ and the perceived disutility $\tilde{D}$ depends on the total amount the person has completed so far that period, denoted by S. Specifically, we have

$$\tilde{D}(E|S) = (1 - \alpha)D(E) + \alpha D'(S) \cdot E \tag{7}$$

where $\alpha \in [0, 1]$ is the degree of projection bias.

Together with the following maximization problem, this completes the single-period setup in a way that allows us to forget about d and s:

$$\tilde{E}^*(S) = \underset{E:E \geq S}{\arg\max}\, B(E) - (\tilde{D}(E|S) - \tilde{D}(S|S))$$

As the next sections will show, $\tilde{E}^*(S)$ is in general not independent of S, so that plans are inconsistent. Before going there, it is important to draw a distinction between the disutility that is projected, and the benefits and opportunity costs.

2.3 Difference Between Utility and Opportunity Costs

Note that a projection-biased person does not project the benefits and opportunity costs. This is not an additional assumption, but follows from the definition of projection bias (and the empirical evidence justifying that definition). Projection bias acts on *utility*, not on choice sets or options available to a person. Projection bias makes a person misperceive the future utility of the options she will have, but it does not make a person misperceive that she will have different options. Economists are of course very familiar with this distinction, for instance from risk aversion. A risk-averse person who believes that a coin is fair perceives the chances of heads and tail as 50% each, just as a risk-neutral person does. But she may give different values to the outcomes of losing \$25,000 or winning \$25,000. The difference between projection bias and risk aversion is of course that risk aversion is about actual preferences, whereas projection bias is about how people perceive their preferences – but it is about how they perceive their preferences, not about how they perceive probabilities or add up money.

Thus a projection-biased student who is at a party right now and therefore has a high current opportunity cost of studying does not then think that her opportunity costs of studying will be high tomorrow at the same time too. What is possible though is that she might perceive those possibilities that make her future opportunity costs as more or less worthwhile depending on whether she feels more or less tired right now. For instance, one alternative to studying might be to go to the gym. It is possible that she might want to go less to the gym the more tired she is from working – not just in comparison to studying, but in comparison to all other remaining activities. In that case, studying would not only affect the perception of future studying, but also of future gym-going, which is bundled into opportunity costs. The reason for ignoring this effect is, first, simplicity; second, the fact that, as long as the effect of studying is substantially larger and more robust than the effect on gym-going, it is a good approximation; and third and finally, the fact that if multiple activities share a common state that changes from any of the activities, that we can model this the way any other complementarities are modeled in economics, and we can unbundle opportunity costs into those tasks that have high complementarities with the activity we are studying, and those that do not.

3 Single-Period Choices

Let us start with single-day decisions where people maximize their daily utility given by $U(E) = B(E) - D(E)$, with the dynamic interpretation described in section 2: people who have worked for S hours so far keep working if they perceive it optimal *at that time*. When the disutility D is convex and the benefits B are linear or concave, a projection-biased person works optimally – despite (in fact, because of) her plans changing. I then show that when the benefits are all-or-nothing, such that a person receives a known reward if she completes a minimum amount of work, people start overly ambitious tasks. They either end up completing the task despite it not being worthwhile, or they quit the task without receiving any benefits for their effort, which goes wasted. Finally, when the person works on multiple tasks each with decreasing returns to effort, then she spends too much time on the task done when she is rested, because she overestimates how long she will keep working on the second task.

3.1 Optimal Behavior with Convex Disutility and Linear Benefits

Consider Anna, a projection-biased student with $\alpha = 0.5$, who has an exam tomorrow. The benefits of every additional hour of studying are equal to 3, and from studying becomes more unpleasant the longer she studies. Specifically, Anna's daily disutility is quadratic in total time studied, thus $D(E) = \frac{E^2}{2}$ and $D'(E) = E$. After having studied for S hours, Anna plans to study until her currently perceived marginal disutility is equal to her marginal benefits (which are constant and equal to 3). I denote the time at which she plans to stop by $\tilde{E}^*(S)$, the total hours she plans to work after having worked for S hours. She perceives her marginal disutility after studying for E hours to lie between her current marginal disutility, $D'(S)$, and her actual marginal disutility after E hours of studying, $D'(E)$:

$$\underbrace{\tilde{D}'(E|S)}_{\text{Perceived } D'} = (1-\alpha)\overbrace{D'(E)}^{\text{Actual } D'} + \alpha \underbrace{D'(S)}_{\text{Current } D'} = \frac{1}{2}(D'(E) + D'(S))$$

At the start of the day, Anna hasn't studied at all and $S = 0$. So she thinks that her marginal disutility after E hours of studying will be $\tilde{D}'(E|0) = \frac{1}{2}D'(E)$. She plans to work for $\tilde{E}^*(0)$ hours, with $\tilde{D}'(\tilde{E}^*(0)|0) = 3 \iff \frac{1}{2}\tilde{E}^*(0) = 3 \iff \tilde{E}^*(0) = 6$. Anna plans to study for 6 hours and thus starts studying. After 2 hours of studying, the current marginal disutility is $D'(2) = 2$. Anna now plans to study for $\tilde{E}^*(2)$ hours in total, with $\tilde{D}'(\tilde{E}^*(2)|2) = 3$ – the first order condition as she perceives it now. This leads to $\tilde{E}^*(2) = 4$ hours. Finally, once she has completed 3 hours of studying, the current marginal disutility is $D'(3) = 3$, so that $\tilde{E}^* = 3$ and Anna stops studying.

The same logic applies when the returns to effort are decreasing rather than constant, so this example essentially proves proposition 1. (All proofs can be found in the appendix.)

Proposition 1. *$D(.)$ is strictly convex and $B(.)$ is linear or concave, $\alpha < 1$. Then a projection-biased person who is self-directed works optimally. Moreover, letting $\tilde{E}^*(S)$ be the optimal amount of work as perceived by the person after having worked for S hours and E^* the optimal amount, we have that $\tilde{E}^*(S) > E^*$ $\forall S < E^*$.*

Proposition 1 also highlights that Anna constantly overestimates how much she will work. Why? By assumption, the marginal disutility of effort increases, so that Anna – who projects her current marginal disutility – underestimates how high marginal disutility will be later that day, and therefore overestimates for long she will study.

Proposition 1 relies on the person being self-directed. If Anna had to make an irreversible (or hard-to-reverse) choice, then she would choose to work too much. This is not likely in the case of studying, but may be the case if Anna is grading exams for a course or working on a common project with a friend. In such situations, due to being overoptimistic, Anna will work too much, since she underestimates how unpleasant work later in the day will be.

Given that behavior is optimal when the disutility is convex and benefits are linear or concave, I now turn to settings where either of these assumptions does not hold. I first consider tasks with all-or-nothing benefits: benefits are received only if the person completes a minimum number of hours. Then, I highlight a major caveat to the optimality result in proposition 1. When a task consists of many small subtasks, each with concave benefits, a projection-biased person no longer exerts optimal effort, because she spends too much time on some subtasks at the expense of others.

3.2 Fixed-Hours Tasks

I now study the choice of a self-directed projection-biased person who can work on a single-day *all-or-nothing* task:[10]

Definition 2. *A single-day* all-or-nothing *task is a task that has benefits B only if the person completes at least E hours by a known deadline.*

Each instant, the person chooses whether to start or continue the task. She does so if and only if she *currently* thinks that completing the task is better than quitting the task. Suppose that Alice, a projection-biased high-school student with $\alpha = 0.5$, has a deadline to finish a college application tonight, which will take her 6 hours. Let's say that $D(6) = 18$ and $B = 12$, so that Alice should not complete the application. For an unbiased person, the decision is clearcut: since the disutility exceeds the benefits, the task is not worth doing. Whether Alice starts the application depends also on how unpleasant the task is at the start.

Does she start the application and, if so, does she finish it? She starts if the perceived disutility $\tilde{D}(6|0)$ is less than B. But $\tilde{D}(6|0) = (1-\alpha)D(6) + \alpha D'(0) \cdot 6 = 9 + \alpha D'(0) \cdot 6$. If $D(E) = \frac{E^2}{2}$, so that $D'(E) = E$, then $\tilde{D}(6|0) = 9 < 12 = B$ and Alice starts the application. After one hour, we have that the perceived disutility from completing the application is $\tilde{D}(6|1) - \tilde{D}(1|1) = \frac{1}{2}(D(6) - D(1)) + \frac{1}{2}D'(1) \cdot 5 = 11.5 < 12 = B$. Thus Alice continues studying. Now imagine what happened if Alice worked for another hour – which, as we will see, does not happen. Then the perceived disutility of completing the application would be $\tilde{D}(6|2) - \tilde{D}(2|2) = 13 > 12 = B$, and so she would not want to continue working. It is clear

[10] All-or-nothing tasks include tasks where a person makes a hard-to-reverse choice or commitment to another person to complete a certain task – thus even tasks that may not naturally be all-or-nothing (such as studying) may become so if an outsider sets incentives in that way.

that Alice will have stopped working before this point. This shows that, as she keeps working, Alice's perception of how unpleasant it is to complete the application increases – something that cannot happen for an unbiased person. As current effort becomes more unpleasant, the final 4 hours of work seem more unpleasant than they did at the start. Because the task now seems harder to complete than it did initially, Alice may decide to quit. Proposition 2 states formally when this happens.

Proposition 2. *A self-directed person with strictly convex disutility D works on an all-or-nothing task that requires effort E to complete and has benefit B if completed. Let $\tilde{E}$ be the actual effort exerted and $U(e) = \mathbb{1}(e = E) \cdot B - D(e)$ the utility of working e hours. Then there exists a unique $E_H \geq 0$ such that the following statements hold:*

1. *$\forall E$, $\exists B$ s.t. $U(E) < 0$ and $\tilde{E} > 0$*
2. *$\forall E < E_H$ if $\tilde{E} > 0$ then $\tilde{E} = E$.*
3. *$\forall E > E_H$, $\exists B$ s.t. $0 < \tilde{E} < E$.*

The first point of proposition 2 states that for any task, we can find a payment such that a person starts the task even though the task isn't worth doing. The proposition also states that there is a threshold E_H such that if a task requires less work than E_H hours, then Alice finishes the task if she starts it, even if the task isn't worth doing. If on the other hand the task requires more than E_H hours of work, it is always possible to find a benefit such that Alice starts the task, yet she doesn't complete it. In fact, it is possible to show that when $D'(0) = 0$, then $E_H = 0$, so that starting and stopping can happen for all tasks, no matter how small.

Proposition 2 applies more widely to tasks with sufficiently convex benefits, although the mistake is the most obvious in for all-or-nothing tasks. Take a task that has increasing returns to effort, so that the benefits $B(\cdot)$ are convex, and which allows for a maximum effort of E. Let $D'(0) > B'(0)$ and $(1 - \alpha)D''(\cdot) < B''(\cdot)$. Then for both an unbiased and a projection-biased person with projection bias α, it is either optimal to work 0 hours or to work E hours. Yet, even though the person at every moment perceives not doing the task at all, or doing the task fully as the only options, she may end up working a little before giving up. Unlike in all-or-nothing tasks, there are some benefits from the work the person puts in. Since we may often not know that $D(\cdot)$ and thus not know whether the person should either put no work in or do all the work, in these situations it is harder to infer whether the behavior was a mistake or not. One type of situations where benefits are convex is when success is discrete – such as receiving an A in an exam, getting a job or promotion – and the probability of success is S-shaped in the amount of effort exerted.

The behavior described here – that people start a task that they don't finish – is similar to people starting but not finishing multi-stage projects as studied in O'Donoghue and Rabin (2008) driven by naive present bias. The important difference is that in O'Donoghue and Rabin (2008) people procrastinate: they repeatedly fail to do a task today, because they think that they will do the task tomorrow. In the case of projection bias, people quit, fully aware that they will *never* complete the task, since the deadline is the same day. Thus there is no scope for procrastination in this setting.[11] Another difference is that a projection-biased student who quits a task is better off quitting – since she underestimates the disutility of completing the task even at the time of quitting – whereas a present-biased person would benefit from completing the task.[12] Thus, while a naive procrastinator would benefit from committing to complete the task, a projector would be hurt by it.

3.3 Multi-Tasking with Concave Benefits

Let us now revisit the situation with convex disutility and decreasing returns to effort, but with a twist: the person now divides her time between two tasks, each of which has decreasing returns to effort. One of the tasks is more *time-sensitive* than the other task:

Definition 3. *Task A is more* time-sensitive *than task B if the benefits from task B are depreciating faster than those of task A, so that (conditional on doing both tasks) task B should be done first. When there are only two tasks, I call the more time-sensitive task simply the* time-sensitive task *and the less time-sensitive task the* flexible task.

A task with an early deadline is more time-sensitive than a task with a late deadline or no deadline. So is a task where there are very small benefits from early completion – such as impressing one's boss or colleagues by completing a task quickly. To illustrate, suppose that Elaine has two problem sets due the same day, one in economics due at 3pm and one in mathematics due at 8pm. Given these deadlines, she starts working on the economics problem set first. For simplicity, assume that the benefits for each problem set are the same and given

[11] It is still possible that a person with naive (but not sophisticated) present bias starts a task that they don't complete if they have present bias at the hourly time frame, but this is substantially less likely than procrastination, since procrastination entails delaying benefits, whereas dropping a task entails never receiving them. Moreover, if present bias is over really short time intervals (say 10 minutes), then it is unlikely to be a plausible explanation.

[12] Here I apply the welfare criterion based on a person's long-run perspective, as in O'Donoghue and Rabin (1999). Even if one does not apply that criterion, the projection-biased person is always better off quitting in this setup.

by $B(.)$, which has decreasing marginal returns. After working on the first problem set for S hours, she plans to spend $E(S)$ hours on each assignment. She thus stops working on the economics assignment when she thinks that she thinks that she has done half the work. Let's say that this happens after 5 hours, at which point she thinks that she will do another 5 hours on the mathematics assignment. She is of course wrong, and overestimates how long she will keep working. Thus she may stop working after only 3 hours on the mathematics assignment. We know from proposition 1 that this choice is optimal *conditional* on her having spent 5 hours on the economics assignment – so the mistake she makes is to spend too much time on the economics assignment, because she is overly optimistic at that time. As proposition 3 shows, while she works less on the second assignment than she should have, she works more in total than would have been optimal.

Proposition 3. *There are two tasks: a time-sensitive one with stictly concave benefits $B_S(.)$, and a flexible one with strictly concave benefits $B_F(.)$. Let $\tilde{E}_S$ and $\tilde{E}_F$ be the actual effort spent on the time-sensitive and flexible task respectively, and E_F^* and E_S^* be the optimal effort levels. Then $B'(\tilde{E}_F) > B'(E_F^*) = B'(E_S^*) > B'(\tilde{E}_S)$ and $\tilde{E}_F + \tilde{E}_S > E_F^* + E_S^*$.*

The proposition states that, optimally, Elaine should have worked less in total than she did, and she should have spent more time on the later task, and less on the earlier task. She works too much because, by the time she should stop – say after 7 hours – she has only spent 2 hours on the mathematics assignment, due to spending too much time on the economics assignment. Therefore her marginal benefit from working on the mathematics assignment is higher than it would be had she worked optimally, and therefore she continues (correctly, given that she cannot undo her earlier mistake) to work for longer to receive some of these high marginal benefits.

A projection-biased person makes the same mistake when working on a single task consisting of two or more subtasks, as long as each subtask is best done in one continuous session. If the subtasks have a natural sequence, so that one subtask makes the subsequent subtask easier, then Elaine will work too much on the earlier stages than on the later stages. For instance, suppose that Elaine plans to read both the lecture notes and to finish a problem set for the same class today. If she believes that the problem set will be easier after reading the lecture notes, then she reads the lecture notes first and consequently spends too much time on them.

In proposition 3, I assume that the person completes the first task fully before switching to the second task. If instead the person constantly switches to the task with the highest current marginal benefit, then we are back to the result of proposition 1. I already mentioned one situation where this is likely not the case: when there is a natural order in which a person should do the tasks. A more general reason why this is unlikely to happen is switching costs. Even very minor switching costs will keep Elaine from switching, since she doesn't realize that her decision would be beneficially affected by doing so. Moreover, the problem may often be substantially worse than I described it, since most tasks consist of multiple subtasks. A problem set consists of multiple problems, and thus Elaine will spend too much time on the early questions at the expense of later questions, even if she did switch between the mathematics and economics problem set. There is one very situation where Elaine would switch more often: If Elaine is uncertain over how long she will take for the individual tasks or subtasks, and she learns from working on each subtask, then she will quite naturally switch between them in order to identify those tasks that will take her a long time and those that won't.[13] Thus, in this particular instance, Elaine may be better off when she is more uncertain about the benefits of each subtask, because this will lead her to switch more often.

4 Multi-Day Tasks and Multiple Deviations

In single-day tasks, a person with increasing marginal disutility is always overly optimistic about how much she will work. In multi-day tasks, she is overly optimistic at the beginning of each day, but if she works long enough, she becomes overly pessimistic: her marginal disutility is higher than it will be in the future, and she thinks that she therefore underestimates how much she will work on future days. She may therefore repeatedly change her mind about whether a task is worth doing, fluctuating between perceiving it worthwhile when work the marginal disutility is low and perceiving it not worthwhile when her marginal disutility is high. In this section I study the how these fluctuations affect effort choices, as well as the welfare implications thereof. Throughout the section the marginal disutility is strictly increasing during a day, and given by $D(\cdot)$.

4.1 Multi-Day All-or-Nothing Task

Consider Beth, a student who is working on an all-or-nothing task with a deadline in T days. She has an economics exam in 100 days and knows that she will receive a B in her final if she does nothing but attend the required lectures. Getting an A on the final has a value of $1,250$ to her. If she studies 5 hours a day on average, Beth is sure to receive an A.

[13] More formally, what Elaine is uncertain about is how quickly the marginal benefits decrease from each subtask.

The daily disutility is quadratic: $D(E) = \frac{E^2}{2}$ and $D'(E) = E$. First, note that Beth at every moment either plans to complete the task efficiently, or to not do the task at all. After all, at any given moment she plans to do what an unbiased person would do whose actual disutility was given by $\tilde{D}(.|S)$. On the first day, Beth therefore studies so long as she perceives it worthwhile to study 5 hours every day. The disutility of studying 5 hours per day is $100 \cdot D(5) = 1250$, so an unbiased student would be indifferent between studying and not studying. But Beth is projection-biased, with $\alpha = 0.5$. At the start of the first day she underestimates the disutility of the task and starts studying. After 2.5 hours of studying, her marginal disutility is $D'(2.5) \cdot 5 = 2.5 \cdot 5 = 12.5 = D(5)$, and she perceives the disutility of working 5 hours on every future day correctly: $\tilde{D}(5|2.5) = (1-\alpha)D(5) + \alpha D'(2.5) \cdot 5 = D(5)$. She therefore perceives the remaining disutility of studying 5 hours every day almost correctly: she still slightly underestimates it because she underestimates the disutility of the 2.5 hours of work she has to complete on the first day. Thus, she keeps studying a little, and as she does so, she overestimates the disutility of studying 5 hours on future days and therefore soon stops working.[14] At the time she stops studying, she perceives the task no longer as worth doing and believes, mistakenly, that she won't resume it the next day. At the beginning of the next day, the same pattern repeats: she starts studying when effort is not very unpleasant is, planning to get an A; and then she stops once effort becomes sufficiently unpleasant.

When does this happen that Beth starts studying with the intention of studying 5 hours, yet she stops studying before she has done 5 hours? If average daily benefits are strictly lower than $\tilde{D}(5|0) = 6.25$, Beth will not start studying, since she doesn't perceive it worthwhile even at the beginning of the day. Similarly, it can be shown that she will work for a full 5 hours if the benefits are strictly larger than $99 \cdot \tilde{D}(5|5) = 99 \cdot 18.75$ – she perceives the remaining 99 days of work as worthwhile, even at the end of the first day.[15] For daily benefits in the range between $(6.25, \frac{99}{100}18.75)$, Beth will start studying on day 1, yet stop before having done 5 hours.

Every day, Beth thus either doesn't study at all, studies inefficiently given how much work still remains to be done, or studies efficiently. It is not difficult to see that if Beth doesn't study at all on day t, than she won't study on day $t+1$ or any later day either, and therefore not get an A. Similarly, if she studies efficiently on day t, then she will study efficiently on all

[14] More concretely, she certainly will stop working when she perceives the disutility of working 5 hours on all future days as equal to the benefits of getting an A, that is once $\tilde{D}(5|S) = \frac{1250}{99}$, which we can solve for S and find that $S \approx 2.56$. In this particular case, she thus won't work more than 2.56 hours on the first day.

[15] The condition that the disutility of work on future days is perceived worthwhile is necessary, but may appear not sufficient. In this case, it is sufficient because one can show that the perceived disutility of completing the task is strictly increasing over the course of the first day, so that it is enough to know the perceived disutility of completing the task as perceived at the end of the first day, after 5 hours of work – which is given by the $99 \cdot \tilde{D}(5|5)$.

future days and thus get an A. For instance, if after 50 days, Beth had only completed 50 hours of studying, she would have to study 9 hours per day on the remaining days, and she wouldn't start studying any longer. Alternatively, if after 75 days Beth had completed 300 hours of studying, she would have to work 8 hours a day for the remaining 25 days to receive the full benefits worth 1250. She would work 8 hours a day, since she would perceive this as worthwhile even after 8 hours of work: $\tilde{D}(8|8) = 16 + 32 = 48 < 50 = \frac{1250}{25}$. Whether or not these situations also can happen for some initial benefit is answered by proposition 4.

The proposition formally states that for any average daily effort required, each of these two outcomes – wasting effort on a task that won't be completed and working inefficiently for a while on a task that is completed – will happen for some average daily benefit, provided that the number of days is sufficiently large. Thus, in our example, this means that if Beth has to study on average 5 hours per day to receive some benefit $T \cdot b$, then we can pick $b = b_L$ such that Beth will fail to achieve the goal, yet waste time studying, and we can pick another benefit $b = b_H > b_L$ such that she will study 5 hours on average and receive the A, but she will work less initially and more at the end. It shouldn't be surprising that T has to be large, since the result is clearly wrong for $T = 1$ given what we know from the single-day settings. There we had, for instance, the result that people finish all tasks that they start if the task requires effort E less than some E_H (proposition 2).

Definition 4 (Multi-Day Task)**.** *The triplet (e, b, T) denotes the following* multi-day task*: the task has to be completed in T days, requires $e \cdot T$ hours of work to be completed and pays $b \cdot T$ if completed by the end of day T.*

Definition 5 (Partial Work)**.** *Take a multi-day task (e, b, T), let $e_t(e, b, T)$ be the amount of work the person actually exerts on day t, and let $E_t := T \cdot e_0 - \sum_0^{t-1} e_i$ be the amount of work that still remains to be done at the start of day t in order to complete the task. Then a person* works partially *on day t if $e_t > 0$ and $e_t < \frac{E_t}{T-t+1}$, that is she does some work, but stops working earlier than is optimal if she plans to complete the task.*

Next I define the number of days for which a person doesn't work at all, and the number of days for which a person works efficiently, or fully (given the amount of effort remaining).

Definition 6. *Fixing some E, let $g_0(B, T)$ be the number of days for which $e_t = 0$, and let $g_F(B, T)$ be the number of days for which $e_t = \frac{E_t}{T-t+1}$. Let $\tau_i(B, T) = \frac{g_i(B,T)}{T}$ for $i \in \{0, F\}$.*

With these definitions out of the way, I can state the main proposition of this section:

Proposition 4. *The disutility of effort is strictly convex with $D'(0) > 0$ and $D'(E) \to \infty$ as $E \to \infty$. Consider a task (E, B, T) with $E > 0$ fixed. Then there exist $B_H(E_0) > B_C(E_0) > B_L(E_0) > 0$ with $B_H > D(E_0) > B_L$ such that*

- *if $B > B_H$, then the task is completed efficiently.*
- *if $B_H > B > B_C$, then $\lim_{T\to\infty} \tau_F(B,T) = \tau_F(B) \in (0,1)$ and the task is completed inefficiently.*
- *if $B_C > B > B_L$, then $\lim_{T\to\infty} \tau_0(B,T) = \tau_0(B) \in (0,1)$ and the task is not completed.*
- *if $B_L > B$, then no effort is spent on the task.*

where $\tau_0(B)$ is continuous and decreasing in B and $\tau_F(B)$ is continuous and increasing in B.

Moreover, letting $\bar{U}(B,T) = \mathbb{1}(B > B_C)B - \sum_{t=1}^{T} D(e_t(B,T))$ be the average daily utility from task (E,B,T), we have that $\lim_{T\to\infty} \bar{U}(B,T) = \bar{U}(B)$ when $B \in [0, B_C) \cup (B_C, \infty)$, is strictly decreasing on (B_L, B_C) and strictly increasing on (B_C, B_H) with $\lim_{B \to B_C^-} \bar{U}(B) \leq -D(E)$.

What does the proposition mean? Notice first that the *average* daily effort and *average* daily benefit are fixed – rather than total effort and benefit – but the number of days the task requires potentially has to be very large. For instance, a problem set may be due in 5 days, yet require an average of 2 hours of work per day, while the final exam may be in 2 months and also require an average of 2 hours of work per day. The proposition then states that if we fix the average effort for a task, there exist three thresholds $B_H > B_C > B_L > 0$ such that if the actual average benefit of the task is low enough (strictly less than B_L), then Beth never starts the task – intuitively this is because the payment is too low for her to want to start the task on the first day. If the average benefit is large enough (strictly larger than B_H) then Beth will complete the task efficiently, because she perceives the task worth doing at all times.

If the payment B lies in (B_L, B_C) – which is a non-empty interval for every $\alpha > 0$ and every $E_0 > 0$ – and if the task requires sufficiently many days (which depends both on B and on E_0), then Beth spends some days working inefficiently until some day T_0 from which point onward she never does any further studying. The day T_0 is moreover such that $\frac{T_0}{T} \approx \tau_0(E_0, B)$, which means that if T is sufficiently large, and if $\tau_0(E_0, B) = 0.3$ say, then Beth always spends close to the first 30% of days wasting pointless effort before she stops working for good. If the payment B lies in (B_C, B_H), the result is similar, except that Beth spends the first 30% of days working inefficiently before working efficiently on the task and finishing it.

The proposition also states that $\tau_0(E_0, B)$ is continuous and decreasing in B. Since τ_0 is the fraction of days a person spends not working at all, it is clear that it goes from 1 to

0. This means that there is some B such that Beth spends 10% of days not working (the final 10% of days) or 1% if the benefits are larger, or 0.1 if the benefits are even larger. Thus increasing the benefits initially doesn't lead to completion of the task, but simply causes Beth to waste more effort for more days on a task she fails to complete. The second part of the proposition states that this in fact can lead Beth to almost complete the task, yet just fall short, which of course means that she has occurred almost all of the disutility of doing the task – on average $D(E_0)$ per day – yet receives no benefits. Thus the repeated fluctuations may lead her to incur this welfare loss. In the example given previously where Beth needs to study 5 hours per day on average, this states that Beth may incur welfare losses close to $D(5) = 12.5$ per day on average.

Why does the result rely on T being large enough? I prove the result for a continuous-time equivalent of proposition 4, where the actual effort levels, as well as τ_0 and τ_F are continuous, which makes it much easier to prove the results. In a setting with T days, τ_0 clearly need not be continuous in E_0 and B: if for a small increase in B Beth wastes effort for one more day, then τ_0 jumps discontinuously since it is the fraction of two integers. As T becomes large, it is possible to approximate the discrete-time solution with the continuous-time one, which proves the result.

Proposition 5 partially answers the question whether $B_C > D(E_0)$ or $B_C < D(E_0)$. In other words, it gives a partial answer to the question whether Beth will pass the exam if it would be worthwhile for her to do so efficiently, or whether she fails to pass her exam even though it would have been worthwhile doing so. Specifically, the proposition states that if $D''' < 0$, Beth completes all worthwhile tasks, as well as some tasks that are not worthwhile. And if $D''' > 0$, Beth fails to complete some worthwhile tasks. The proposition does not state, though, that if Beth completes a task, she is better off. If $D''' < 0$ and Beth completes task that would have been worthwhile if completed efficiently, it may still be the case that she is worse off than if she had never worked on it, since she may complete the task very inefficiently.

Proposition 5. *Let $D'' > 0$ with $D'(0) > 0$, and consider a task given by (E, b, T). Let $\bar{b}(E,T)$ be the threshold such that if $b > \bar{b}(E,T)$ the task is completed, and if $b < \bar{b}(E,T)$ the task is not completed. Then we have that*

- $D''' < 0 \implies \lim_{T \to \infty} \bar{b}(E,T) < D(E)$
- $D''' > 0 \implies \lim_{T \to \infty} \bar{b}(E,T) > D(E)$

The reason that D''' matters is as follows. Suppose that Beth faces a task with $D(E_0) = B$ – one where an unbiased person is indifferent between completing it and not completing it efficiently – with $E_0 = 5$. Then on the first day, Beth will do some work, but not the full 5 hours if T is large enough (this follows from proposition 4, since $B_H > D(E_0 > B_L)$. Let's say for simplicity that Beth works 2.5 hours and that $T = 100$. Then on the second day she faces effectively a task with a deadline in 99 days, and one that requires $5 + \frac{2.5}{99} \approx 5.025$ hours on average for each of those 99 days. Thus the change in total disutility of completing the task efficiently starting on day 2 compared to completing it efficiently starting on day 1 has two components. First, Beth doesn't need to do the 5 hours on day 1 any more, so the disutility decreases by $D(5)$. Secondly, she has to do an additional 0.025 hours every day from now on, which increases total disutility by $99 \cdot (D(5.025) - D(5))$. When the marginal disutility is concave – when $D''' < 0$ – then the first factor wins out, and the disutility from completing the task is smaller on day 2 than on day 1. Since on day 1 we had that $D(E_0) = B$, this means that the task is now strictly worth doing (if done efficiently). Thus the task becomes strictly better, and thus is completed on the last day, since all worthwhile tasks are completed on the final day. When the marginal disutility is convex – when $D''' > 0$ – then the task is definitely not worth doing on the second day. In this case, the task becomes worse with every passing day, and (if there are sufficiently many days) it becomes so bad that Beth eventually stops working on it for good.

What does it mean that $D''' < 0$ rather than $D''' > 0$? A task that has $D''' < 0$ has concave marginal disutility, which means roughly that the difference in disutility between the 11th hour and the 10th hour of work is smaller than the difference in disutility between the 2nd and the 1st hour. Thus while additional work is more unpleasant, the increase in unpleasantness is smaller. When instead $D''' > 0$, then the difference between the 11th and the 10th hour is larger than that between the 1st and the 2nd.

4.2 Multi-Day Multi-tasking

As we saw previously, a projection-biased person works optimally on a single task with concave benefits, but when she multi-tasks, she works too much on time-sensitive tasks at the expense of more flexible tasks. I now consider multi-tasking over multiple days. Specifically, I consider what happens when Carla, a projection-biased student with projection bias 0.5, has an exam in T days, and has to decide how much time to spend on various short-term tasks each day, such as administrative tasks or attending lectures. As in one-day situations, Carla is overoptimistic about how much she will work each day. This leads her to do too many of the time-sensitive tasks during the day. Unlike in one-day situations, Carla doesn't work as hard at the end of the day to catch up, because she underestimates how many time-sensitive tasks she will do on future days. In short, she thinks that in the future, unlike today, she will focus more on studying for the exam than she did today.

Suppose that Carla has an economics exam in 2 days. She receives benefits of 6 for each of the first 8 hours of studying until then, and 3 for every hour thereafter. Each day, she can also attend 2 hours of lectures for a mathematics class, each hour of which gives her benefits of 4. Carla's disutility is given by $D(E) = \frac{E^2}{2}$.[16] What will Carla do? As long as she has worked less than 2 hours, Carla thinks that the marginal disutility after 6 hours of work (4 from studying, and 2 from attending the lecture) is less than $\tilde{D}'(6|2) = \frac{1}{2}(D'(6) + D'(2)) = 4$. Since the benefits from the lecture are equal to 4, she therefore plans to attend the lecture. Following the lectures, she plans to study for 4 hours, but after 2 hours of studying, she realizes that the marginal benefit of attending the lecture tomorrow is less than the marginal disutility. She therefore no longer plans to attend tomorrow's lecture, and thus plans to do the 8 hours of studying by working 5 hours each day and studies 3 hours the first day. Yet, she starts the next day refreshed and attends the lecture, only to realized after 1.5 hours that it isn't worth staying. After 4.5 hours of additional studying, she gives up, having studied only 7.5 hours in total, and having worked 1 hour more on the second day than on the first.

As proposition 6 shows, if a person first works on short-term tasks before working on long-term tasks, then the person works too much on the short-term tasks at the expense of the long-term task, and works more and more each day. On earlier days, she stops work early because she underestimates how much time she will fritter away on future time-sensitive tasks, and therefore has to work longer to achieve the gains from the long-term task.

Proposition 6. *A person works on one long-term task over T days with benefits $B_L(\boldsymbol{e}) = B_L(\frac{\sum_{t=1}^{T} e_t}{T})$ and each day faces a short-term task with benefits $B_S(e)$. Both B_L and B_S are strictly concave. Let $\tilde{\boldsymbol{e}}_S$ and $\tilde{\boldsymbol{e}}_L$ be the amount of effort spent on the short-term and long-term tasks by a projection-biased person, and $\boldsymbol{e}_S^*$ and $\boldsymbol{e}_L^*$ the optimal amounts.*

When the person works first on the long-term task, then the behavior is equivalent to the one-day setting in proposition 3 repeated T days in a row, where the long-term task takes the role of the time-sensitive task, with daily effort given by the average daily effort B_L.

When people work on the short-term task first, then $\tilde{e}_{L,t} \geq \tilde{e}_L^*$ and $e_S^* \geq \tilde{e}_{S,t}$, and $\tilde{e}_{S,t}$ is decreasing over time, while $\tilde{e}_{L,t}$ and $\tilde{e}_{S,t} + \tilde{e}_{L,t}$ are increasing in t.

In the example, Carla attends the lectures first because I assumed the lectures are early in the day. This need of course not always be the case, although there is one natural setting

[16] As before, effort on economics studying or attending lectures has the same effect on her disutility, which is a simplification. As long as the disutility for studying economics depends on some kind of 'tiredness' that increases with the amount of time spent on doing mathematics, the results are qualitatively the same.

where Carla would first work on the short-term task before working on the long-term task. Suppose that Carla is uncertain both about the marginal benefits of the short-term task and the long-term task, and she learns about those benefits while working on the tasks. If a task turns out to be more beneficial than she expected, she will want to work for longer. If the task she learns about is a short-term task, then she should work on the short-term task first, since she will be able to smooth the shock over all the remaining days, by working less on the long-term task today and make up for it on future days. For instance, suppose that Carla first works on the problem set and learns that she will work 1 hour more than expected. Then she can work one hour more on the problem set, yet only work one tenth of an hour more in total the first day by spreading the shock across the remaining days. If on the other hand she first studies for the exam, then once she finds out that she will work one additional hour, she cannot undo her studying, and so can no longer smooth work across days. Thus, in the case of uncertainty it will be quite natural for Carla to work first on the short-term task — and thus to end up spending too much time on short-term tasks.

5 Careless Timing: Misallocation of Effort

In the previous section I showed that, even though they always plan on completing the task inefficiently or not at all, projection-biased people may complete tasks inefficiently. This is because their plans may fluctuate between completing a task and dropping it. In this section I show that projection-biased people make plans that allocate effort inefficiently, even if they total effort is fixed and they never decide to drop the task. The reason is that projection-biased people underestimate the difference in marginal disutility at different times and therefore are too willing to work more on days when they receive larger marginal benefits for their work or on days when their productivity is higher.

5.1 Misallocation in General

A person on day 0 plans how much to work between days 1 and T. She has worked for S hours on day 0, so that the instantaneous disutility from working is $D'(S)$. She plans to work $\tilde{\boldsymbol{e}}(S) = (\tilde{e}_1(S), ..., \tilde{e}_T(S))$ where $\tilde{e}_t(S)$ is the day-t effort. These plans maximize her current perceived utility:

$$\tilde{U}(\tilde{\boldsymbol{e}}|S) := B(\tilde{\boldsymbol{e}}) - \sum_{t=1}^{T} \tilde{D}(\tilde{e}_t|S) = B(\tilde{\boldsymbol{e}}) - (1-\alpha)\sum_{t=1}^{T} D(\tilde{e}_t) - \alpha D'(S) \cdot \tilde{E} \tag{8}$$

where $B(\tilde{\boldsymbol{e}})$ is the benefit from working and $\tilde{E} := \sum_{t=1}^{T} \tilde{e}_t$ is total effort across all days. I will use the following lemma:

Lemma 1. *Let $U_a(\boldsymbol{e}) = X(\boldsymbol{e}) + a \cdot Y(\boldsymbol{e})$, with X and Y continuous (real-valued) functions of the vector $\boldsymbol{e}$, and $a \in \mathbb{R}$ a fixed parameter. Let $\boldsymbol{e}(a) \in \arg\max_{e \in \mathcal{E}} U_a(\boldsymbol{e})$ for some compact set $\mathcal{E}$. If $a_H > a_L$, then $X(\boldsymbol{e}(a_H)) \leq X(\boldsymbol{e}(a_L))$ and $Y(\boldsymbol{e}(a_H)) \geq Y(\boldsymbol{e}(a_L))$.*

Lemma 1 says that if a person maximizes a sum of two utilities then the person who puts more weight on the second dimension chooses a bundle with higher utility in that dimension.

Now suppose a person works on a given task requiring a fixed amount of effort $E := \sum_{t=1}^{T} e_t$. We can rewrite total utility from equation (8) as follows

$$\tilde{U}(\tilde{\boldsymbol{e}}|S) = B(\tilde{\boldsymbol{e}}) - \sum_{t=1}^{T} D(\tilde{e}_t) + \alpha \left(\mathbb{E}_0 \left(\sum_{t=1}^{T} D(\tilde{e}_t) \right) - D'(S) \cdot E \right) \tag{9}$$

We can apply lemma 1 to see that a more biased person chooses effort such that $\mathbb{E}_0 \left(\sum_{t=1}^{T} D(\tilde{e}_t) \right) - D'(S) \cdot E$ increases with α. Thus total disutility increases or total effort E decreases or both. When total required effort $E := \sum_{t=1}^{T} e_t$ is fixed, it must be the case that total disutility increases. Of course, a projection-biased person only increases the disutility if this also increases the benefits. She thus overreacts to benefits and opportunity costs, underestimating how much this will cost her, as shown in the next proposition.

Proposition 7. *Suppose that the person has to complete total effort E by day T. Then we have that $\tilde{\mathbf{e}}^*(S) = \tilde{\mathbf{e}}^*$, so that planned effort does not depend on S. Moreover, $\tilde{\mathbf{e}}$ is the same as that of an unbiased person with actual disutility $(1-\alpha)D$ or, equivalently, with net benefits $\frac{B}{1-\alpha}$.*

Let us consider what kinds of mistake this can lead to. Doris, a projection-biased student, has to decide how to split 10 hours of studying across two days. Her disutility of effort is convex, so that it becomes more unpleasant to keep studying the longer she has studied. A good friend is visiting town tomorrow, so that Doris has higher opportunity costs of time tomorrow than today. She therefore should work more today than tomorrow, say 7 hours today and 3 hours tomorrow. But Doris underestimates how much more unpleasant it will be to work an additional hour tomorrow and thus works even more today than she should, say 8 hours.[17] In the next section, I show that when the marginal disutility is decreasing, then this underestimation leads projection-biased people to take too many breaks.

[17] This underestimation in the domain of consumption choices would lead a person to consume goods too fast, since she would not realize how much more she would enjoy eating something once she is hungry rather than sated. This is exactly what Galak, Kruger, and Loewenstein (2013) find: people consume chocolates faster when they choose themselves when to eat, yet report enjoying them less.

5.2 Uncertainty, Productivity, and Time Discounting

Let us now move to a setting where people discount time. Doris has to complete an assignment by tomorrow night that requires $e_1 + e_2 = E$ hours. She discounts time: her disutility is $D(e_1) + \delta D(e_2)$, with $\delta < 1$. When she has worked for S hours, she plans to stop after completing $e_1(S)$ hours today, given by

$$\begin{aligned} & \tilde{D}'(e_1(S)|S) = \delta \tilde{D}'(e_2(S)|S) \\ \iff & (1-\alpha)D'(e_1(S)) + \alpha D'(S) = \delta(1-\alpha)D'(e_2(S)) + \delta\alpha D'(S) \\ \iff & D'(e_1(S)) - \delta D'(e_2(S)) = -\frac{\alpha}{1-\alpha} D'(S)(1-\delta) \end{aligned} \tag{10}$$

which shows that $e_1(S) > e_1^*$ when $\delta < 1$ and $D'(S) > 0$. She stops working when her current perceived plan is equal to (or less than) what she has done, that is when $e_1(S) = S$. Substituting this into (10), we get

$$\begin{aligned} & D'(\tilde{e}_1) = \delta(1-\alpha)D'(\tilde{e}_2) + \delta\alpha D'(\tilde{e}_1) \\ \iff & D'(\tilde{e}_1) = \frac{1-\alpha}{1-\delta\alpha} \delta D'(\tilde{e}_2) \end{aligned}$$

and so she acts as if her discount rate was $\frac{1-\alpha}{1-\delta\alpha}\delta$, which is strictly less than δ.

A similar overreaction happens when Doris doesn't discount time but may learn how to complete the assignment more efficiently tomorrow. For instance, the next lecture might provide a shortcut about how to solve the assignment. Say that Doris knows that she can solve the problem in 4 hours, but she expects that with probability p she may learn a shortcut tomorrow, in which case she will only have to spend 15 minutes to write the solution up. If she does no work at all today and she learns no shortcut tomorrow, then she has to do the full 4 hours tomorrow. Her optimization problem is

$$\min_{e_1} \tilde{D}(e_1) + (1-p)\tilde{D}(4-e_1) + pD(0.25)$$

with first order conditions given by

$$\tilde{D}'(e_1) = (1-p)\tilde{D}'(4-e_1)$$

The projection-biased person stops again when $s = e_1(s)$

$$D'(e_1) = (1-p)(1-\alpha)D'(4-e_1) + (1-p)\alpha D'(e_1)$$
$$\iff D'(e_1) = \frac{1-\alpha}{1-\alpha q} q D'(4-e_1)$$

where $q = 1-p$. As in the case of discounting, since $\frac{1-\alpha}{1-\alpha q} < 1$ she works too little today.

Both of these results are special cases of proposition 8.

Proposition 8. *Suppose a person allocates effort over T days, with uncertainty over what the total amount of required effort is. Uncertainty resolution is independent of the person's actions. On the final day, the state is known and the constraint $E_s = \sum_{t=1}^{T} p_t \cdot e_t$ must hold, where p_t is her (known, exogenously given) productivity on day t. Then on the first day, she perceives the first order conditions after s hours of work as*

$$D'(\tilde{e}_1(s)) - q_t \mathbb{E}(D'(\tilde{e}_t(s))) = -\frac{\alpha}{1-\alpha} D'(s)(1-q_t)$$

where $q_t = \frac{p_t}{p_1}$ and $\tilde{e}_1$, thus solves

$$D'(\tilde{e}_1^*) = \frac{1-\alpha}{1-\alpha q_t} \mathbb{E}(D'(\tilde{e}_{t|1}^*))$$

where $\tilde{e}_{t|1}^*$ is the amount of work on day 6 the person plans at the end of day 1.

When $T = 2$, if $q_2 < 1$, we have that $\tilde{e}_1(s)$ decreases with s, and if $q_2 < 1$, we have that $\tilde{e}_1(s)$ increases with s.

Using these first order conditions, we can show the following:

Proposition 9. *Consider the same setup as in proposition 8, with q_t weakly increasing and the optimal solution having $e_1^* > 0$. If q_t is strictly increasing, or the optimal solution is such that there will be no work on at least one day, then denoting by $\tilde{E}_t^*$ the total amount of work done by the beginning of day t, we have that $\tilde{E}_t^* < E_t^*$ $\forall t > 1$. Moreover, the person overestimates on day t how much work she will have done by day $t' > t$.*

Similarly, if there is a maximum amount of work a person can do in a day, say $\bar{E}$, then if q_t is strictly decreasing or if optimal effort is equal to $\bar{E}$ on some days but never zero, then $\tilde{E}_t^* > E_t^*$. Moreover, the person underestimates on day t how much work she will have done by day $t' > t$.

Here is an example of proposition 9 in action. Betsy has to complete an assignment that would take her $E = 18$ hours of work if each day she was as productive as she is today. Fortunately for her, she has lectures tomorrow and office hours the day after that: during lectures and office hours she will learn shortcuts for completing the problems on the assignment. Specifically, every hour of work done tomorrow is worth 1.5 hours of work today, while every hour of work done in two days is worth 1.5 hours of work tomorrow.

Obviously, she should work less today than tomorrow, since she will become more efficient at solving questions. Suppose that her disutility is quadratic. The optimal effort levels should satisfy the first order conditions $D'(e_1^*) = \frac{1}{p}D'(e_2^*) = \frac{1}{p^2}D'(e_3^*)$, which leads to $e_1^* \approx 2.2$, $e_2^* \approx 3.33$, and $e_3^* \approx 4.95$. On day 1, Betsy instead solves her perceived first order conditions, which we can derive as in the 2-day case to be $D'(\tilde{e}_1) = \frac{1-\alpha}{1-\alpha\frac{1}{p}}\frac{1}{p}D'(\tilde{e}_{2|1}) = \frac{1-\alpha}{1-\alpha\frac{1}{p^2}}\frac{1}{p^2}D'(\tilde{e}_{3|1})$, where $\tilde{e}_{i|1}$ indicates that it is the effort Betsy perceives to be optimal at the end of day 1. These are given by $\tilde{e}_1 \approx 1.52$, $\tilde{e}_{2|1} \approx 3.03$, and $\tilde{e}_{3|1} \approx 5.29$.

Yet, on day 2 she will not do what she thought she would do. She solves her new perceived first order condition, which is now exactly as in the 2-day case, taking into account that she worked roughly 1.52 hours on day 1: $D'(\tilde{e}_2) = \frac{1-\alpha}{1-\alpha\frac{1}{p}}\frac{1}{p}D'(\tilde{e}_3)$. Solving this, we find that $\tilde{e}_2 \approx 2.75$ and that $\tilde{e}_3 \approx 5.50$. Betsy was already planning to work less than she should, planning to do 3.03 instead of 3.33, yet she ends up doing even less, namely 2.75. Thus, Betsy postpones too much work, and thinks that she will have done more by the end of day 2 than will be the case. The reason is that Betsy wants to delay more effort, the more unpleasant effort is at that time. Betsy correctly understands that doing 1 minute less of work requires her to do 40 seconds more work tomorrow. Thus she saves 20 seconds, which she perceives as more unpleasant the more unpleasant effort is right now. Therefore she is willing to delay more work until tomorrow to take advantage of her higher productivity. Since tomorrow she will work more, she will be more tired at the end of the day when she decides to stop, and therefore she will want to delay more at the end of day 2 than at the end of day 1 and stops working earlier than anticipated.

6 Concave Disutility

Until now I have assumed that the daily disutility D is convex, so that work becomes more unpleasant the longer a person works. While this is often correct, it is also true that some tasks become easier as we warm up or become focused, before eventually becoming harder as fatigue and boredom take over. Warm up is an integral part of both sports and music performance, and many tasks that require focus, such as writing or programming, get easier after an initial time of settling in. For this reason, in this section I explore daily disutility that, at least initially, is concave.

When a projection-biased person with concave disutility faces a single all-or-nothing task, she overestimates the task's disutility at the start of the day and therefore fails to do some worthwhile tasks. Let us consider the same situation as in section 3, where a high-school student named Alice has a deadline to finish a college application by midnight. Her daily disutility is $D(E) = 6 \cdot E - \frac{E^2}{2}$ for $E \leq 6$. The application takes her exactly 6 hours to complete and is worth $B = 20$. Since $\tilde{D}(6|0) = 27 > 18 = D(6)$, Alice decides not to complete the application: she is willing to start the application only if $B > 27$.

This result also holds when D is first concave and then convex. Concretely, suppose that $D(\cdot)$ is concave for $E < \bar{E}$ and convex on $E > \bar{E}$ for some $\bar{E} > 0$. Then there is a threshold E_L such that Alice overestimates the task's disutility if it takes fewer than E_L hours to complete, and she underestimates it if it takes more. We thus obtain the following proposition which, when the task is large enough, mirrors proposition 2 and, when the task is small enough, mirrors the preceding example.

Proposition 10. *Let D be concave for $E < \bar{E}$ and convex for $E > \bar{E}$, with $\bar{E} > 0$. The marginal disutility eventually becomes larger than it is initially: $\lim_{E\to\infty} D'(E) = \bar{D}$ with $\bar{D} > D'(0)$. A self-directed person faces an all-or-nothing task requiring effort E_0 and paying B if completed. Let $\tilde{E}$ be the actual effort exerted and $U(\tilde{E}) = \mathbb{1}(\tilde{E} = E_0) \cdot B - D(\tilde{E})$ the total utility. Then there exist unique E_H and E_L with $E_H > E_L > \bar{E}$ and $D'(0) = \frac{D(E_L)}{E_L}$ such that the following hold:*

- $\forall E_0 > E_L$, $\exists B$ *s.t* $\tilde{E} > 0$ *and* $B - D(E_0) < 0$.
- $\forall E_0 < E_L$, $\exists B$ *s.t* $\tilde{E} = 0$ *and* $B - D(E_0) > 0$.
- $\forall E_0 < E_H$ *if* $\tilde{E} > 0$ *then* $\tilde{E} = E_0$.
- $\forall E_0 > E_H$, $\exists B$ *s.t.* $0 < \tilde{E} < E_0$.

Now suppose that we know that Alice will complete the application for sure, but she can take a break, whether to have coffee with a friend, respond to emails, or do homework. Since her disutility is concave, she should work without interruptions, unless the benefits B from taking a break are positive. Suppose that after 3 hours of work, a friend asks Alice if she wants to grab coffee. As we have seen in 7, Alice underestimates the change in disutility by a factor of $1 - \alpha = 0.5$. Thus, since the actual change in disutility from taking the break is 3 for each of the 3 hours she will have to do after the break, she perceives the increase in disutility to be $(1-\alpha)(3 \cdot 3) = 4.5$ and joins her friend if $B > 4.5$, even though she should only do so if $B > 9$.

Now consider the case where Alice can work up to 6 hours on her application, with each hour of work having a benefit of 5. Then Alice may take a break expecting to continue working on the application afterwards, yet fail to resume the task. She starts the application, since $5 \cdot 6 = 30 > 27 = \tilde{D}(6|0)$. After 3 hours of work, her friend again asks her if she wants to have a break. If she takes a break, she will fail to resume the task, since her perception of resuming again will be $\tilde{D}(3|0) = \frac{1}{2}(D(3)+D'(0)\cdot 3) = \frac{1}{2}(13.5+18) = 15.25 > 15 = 3\cdot 5$. But after 3 hours of work, Alice mistakenly thinks that she will resume the task, since she currently perceives the disutility of completing 3 hours after the break as $\tilde{D}(3|3) = \frac{1}{2}(13.5 + D'(3) \cdot 3) = 11.25 < 15$. As we just saw, she thinks that the increase in disutility from taking the break is only 4.5 and thus takes a break if the benefits exceed 4.5. Since she doesn't resume the task, she in fact is worse of by $3 \cdot 5 - (D(6) - D(3)) = 11.5$, so that if $B \in (4.5, 11.5)$ she is strictly worse off from taking the break.

Moreover, Alice may fail to resume a task even if the task becomes easier only over very short time intervals – if $\bar{E}$ is small. To take an extreme example, suppose $D'(0) = 6$ and $D'(E) = 2 \; \forall E > 0$, so that the task is instantaneously easier after Alice gets started. Then as long as Alice resumes the task, she doesn't incur any costs. But because there are no costs incurred from resuming, she always thinks that she will resume the task if she currently would continue, and therefore she will take every opportunity for beneficial breaks, even though she may not resume the task, since she overestimates the disutility when she has to start.

One way that Alice can overcome the failure to start a worthwhile task is to front-load benefits. Let us stick with the example where $D'(0) = 6$ and $D'(E) = 2 \; \forall E > 0$. Suppose the benefits per hour are b and that Alice can work for at most 6 hours on the task. Then she perceives the task worthwhile doing if $b > 4 = \tilde{D}'(E|0)$, and not worthwhile if $b < 4$, even though it is worth doing as long as $b > 2$. If instead, Alice perceived the rewards of the first hour to be $b + 1$ and the rewards of the other hours to be equal to $b - \frac{1}{5}$, then she would start working for $b > 3$ – and once she starts working, she will continue to work as long as $b - \frac{1}{5} > 2$. In short, Alice would complete all 6 hours if $b > 3$, even though the total benefits from doing so did not increase. Whenever Alice doesn't start a worthwhile task, there is a way to front-load benefits (without raising total benefits) such that Alice completes all the work.

7 Discussion and Conclusion

Throughout the paper, I made three assumptions on the instantaneous disutility. First I assumed that a person either works or doesn't work, ruling out intensity of effort. Second I assumed that the instantaneous disutility depends only on total time a person has worked so far, ruling out breaks and rest during a day. And third, I assumed that people know their disutility, but misperceive it. I now discuss each of these assumptions in turn.

It is straightforward to extend the results in sections 3 and 4 to allow for intensity of effort. These results rely on the person being overly optimistic when work is currently easy and overly pessimistic when it is currently hard, which remains true when we allow for intensity of effort. In the notation of section 2, the instantaneous disutility is $d(e, s)$ where e is no longer restricted to 0 or 1, and a projection-biased person perceives her future instantaneous disutility as $\tilde{d}(e, s|s_0) = (1 - \alpha)d(e, s) + \alpha d(e, s_0)$ when she is in state s_0. If $d(e, s)$ is increasing in s, then a person starts some all-or-nothing tasks that are overly ambitious, yet may quit once she becomes tired. She will also work too much on time-sensitive tasks when multi-tasking, underestimating how soon she will stop working on other tasks afterwards.

On the other hand, the results in section 5 do change in meaningful ways. Whereas in the simple framework, a person can only change the timing of effort by working more on one day – by changing the extensive margin of time – now she can also work more on a given day by working at a higher intensity – by changing the intensive margin. When a person responds to changes in incentives and opportunity costs primarily along the extensive margin – by working longer, but not harder – the results in section 5 hold: the person overreacts to incentives and opportunity costs, working too much on days where incentives are higher or costs lower. If however, she responds primarily along the intensive margin, this need no longer be true. Consider, to take the extreme case, a person who has to work for exactly 8 hours every day, but can choose the intensity of effort at every moment. She has to complete an all-or-nothing task, requiring a fixed amount of work, so that she only decides when to work, but not how much. Then, at the end of the first day, being tired she underestimates how easy work will be the next day and therefore works more on the first day, even when there are no incentives to do so and no differences in opportunity cost of time. Thus the intensive margin has new implications, and may push behavior in a different direction from the extensive margin studied in section 5.

Let us now relax the assumption of no rest during a day and consider a person who can take a fixed number of breaks in a day. This is identical to the situation analyzed in section 4. The main lesson from section 4 was that repeated fluctuations in plans can lead to inconsistent behavior with large welfare consequences for all-or-nothing tasks, and that people are systematically overly optimistic about how well they will complete long-term tasks with decreasing returns to effort. If a person can take breaks during a day, these results extend to daily tasks. Since a person doesn't become fully rested after a break and can choose how many breaks to take, the fluctuations in marginal disutility during a day will be less severe than those studied in section 4, which may attenuate – but not reverse – these results.

Finally, people are often uncertain about their disutility of effort, in which case a projection-biased person may mislearn what her actual disutility is. She may mistakenly attribute her dislike of a task she always does when tired to the task itself, rather than to the fact that she is always tired when doing it. This may lead to the type of attribution bias as described in Haggag and Pope (2016). This type of incorrect belief updating may lead her to become too confident that some tasks are better than others – or fail to realize that this is the case.

All of this suggests that the basic logic drawn out in this paper – the repeated fluctuation between overly optimistic and pessimistic, the inconsistent plans, the inefficient effort allocation – are robust to extensions that allow for effort intensity and flexible breaks.

References

Acland, Dan, and Matthew R. Levy. 2015. “Naiveté, Projection Bias, and Habit Formation in Gym Attendance.” *Management Science* 61 (1): 146–160.

Ariely, Dan, and George Loewenstein. 2006. “The Heat of the Moment: The Effect of Sexual Arousal on Sexual Decision Making.” *Journal of Behavioral Decision Making* 19 (2): 87–98.

Augenblick, Ned, and Matthew Rabin. 2016. “An Experiment on Time Preference and Misprediction in Unpleasant Tasks.”

Badger, Gary J., Warren K. Bickel, Louis A. Giordano, Eric A. Jacobs, George Loewenstein, and Lisa Marsch. 2007. “Altered States: The Impact of Immediate Craving on the Valuation of Current and Future Opioids.” *Journal of Health Economics* 26 (5): 865–876.

Buchheim, Lukas, and Thomas Kolaska. forthcoming. “Weather and the Psychology of Purchasing Outdoor-Movie Tickets.” *Management Science*.

Buehler, Roger, Dale Griffin, and Johanna Peetz. 2010. “Chapter One-the Planning Fallacy: Cognitive, Motivational, and Social Origins.” *Advances in Experimental Social Psychology* 43. Elsevier: 1–62.

Buehler, Roger, Dale Griffin, and Michael Ross. 1994. “Exploring the‘ Planning Fallacy’: Why People Underestimate Their Task Completion Times.” *Journal of Personality and Social Psychology* 67 (3). American Psychological Association: 366.

Busse, Meghan R, Devin G Pope, Jaren C Pope, and Jorge Silva-Risso. 2015. “The Psychological Effect of Weather on Car Purchases.” *The Quarterly Journal of Economics* 130 (1). Oxford University Press: 371–414.

Conlin, Michael, Ted O’Donoghue, and Timothy J. Vogelsang. 2007. “Projection Bias in Catalog Orders.” *American Economic Review* 97 (4): 1217–1249.

Fisher, Geoffrey, and Antonio Rangel. 2013. “Symmetry in Cold-to-Hot and Hot-to-Cold Valuation Gaps.” *Psychological Science*. Sage Publications, 0956797613502362.

Galak, Jeff, Justin Kruger, and George Loewenstein. 2013. “Slow Down! Insensitivity to Rate of Consumption Leads to Avoidable Satiation.” *Journal of Consumer Research* 39 (5): 993–1009.

Gilbert, Daniel T., Elizabeth C. Pinel, Timothy D. Wilson, Stephen J. Blumberg, and Thalia P. Wheatley. 1998. “Immune Neglect: A Source of Durability Bias in Affective

Forecasting." *Journal of Personality and Social Psychology* 75 (3): 617.

Haggag, Kareem, and Devin G Pope. 2016. "Attribution Bias in Consumer Choice." Working Paper.

Herrnstein, Richard J, and Dražen Prelec. 1991. "Melioration: A Theory of Distributed Choice." *The Journal of Economic Perspectives*. JSTOR, 137–156.

Levy, Matthew. 2009. "An Empirical Analysis of Biases in Cigarette Addiction." Working Paper.

Loewenstein, George, and Daniel Adler. 1995. "A Bias in the Prediction of Tastes." *The Economic Journal*. JSTOR, 929–937.

Loewenstein, George, Daniel Nagin, and Raymond Paternoster. 1997. "The Effect of Sexual Arousal on Expectations of Sexual Forcefulness." *Journal of Research in Crime and Delinquency* 34 (4). Sage Publications: 443–473.

Loewenstein, George, Ted O'Donoghue, and Matthew Rabin. 2003. "Projection Bias in Predicting Future Utility." *The Quarterly Journal of Economics* 118 (4): 1209–1248.

Nordgren, Loran F, Joop van der Pligt, and Frenk van Harreveld. 2008. "The Instability of Health Cognitions: Visceral States Influence Self-Efficacy and Related Health Beliefs." *Health Psychology* 27 (6). American Psychological Association: 722.

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

O'Donoghue, Ted, and Matthew Rabin. 2008. "Procrastination on Long-Term Projects." *Journal of Economic Behavior & Organization* 66 (2): 161–175.

Read, Daniel, and Barbara Van Leeuwen. 1998. "Predicting Hunger: The Effects of Appetite and Delay on Choice." *Organizational Behavior and Human Decision Processes* 76 (2): 189–205.

Van Boven, Leaf, and George Loewenstein. 2003. "Social Projection of Transient Drive States." *Personality and Social Psychology Bulletin* 29 (9). Sage Publications: 1159–1168.

A Proofs

A.1 Proofs for Section 3

Proof of proposition 1.

Proof. A projection-biased person solves the same maximization problem as an unbiased person, but perceives her disutility to be $\tilde{D}(.|S)$ at time S. The first order condition thus depends on time S:

$$\tilde{D}(\tilde{E}^*(S)|S) = B'(\tilde{E}^*(S)) \iff (1-\alpha)D'(\tilde{E}^*(S)) + \alpha D'(S) = B'(\tilde{E}^*(S))$$

Let E^* be the optimal level of effort, so that $D'(E^*) = B'(E^*)$. Then $S < E^* \implies D'(S) < D'(E^*) \implies \tilde{D}(E^*|S) < D'(E^*) = B'(E^*)$, so that $\tilde{E}^*(S) > E^*$, since B is concave and D strictly convex. Similarly, if $S = E^*$, then $\tilde{D}'(E^*|S) = D'(E^*) = B'(E^*)$. Thus as long as the person has worked less than E^*, she plans to work more than E^* and continues working, and once she has worked E^* she stops. □

Here is the proof of proposition 2.

Proof. Let $\tilde{R}(E|S) := \tilde{D}(E|S) - \tilde{D}(S|S)$, the perceived remaining disutility of completing the task after S hours of work have already been completed. Notice that the person works as long as $\tilde{R}(E|S) < B$, never works when $\tilde{R}(E|S) > B$.

The first part of the proposition claims that for all $E > 0$, it is possible to find a $B > 0$ such that the person starts working on the task. Notice that $\tilde{D}(E|0) = (1-\alpha)D(E) + \alpha D'(0) \cdot E < D(E)$, since $D(E) = \int_0^E D'(S)dS > \int_0^E D'(0)dS$ since $D(\cdot)$ is strictly convex and $E > 0$. Pick $B \in (\tilde{D}(E|0), D(E))$. Then, since $\tilde{R}(E|0) = \tilde{D}(E|0)$ and since $\tilde{R}(E|S)$ is continuous in S, we have that $\tilde{R}(E|\varepsilon) < B$ for some sufficiently small $\varepsilon > 0$, so that the person will work at least for a time ε. This proves the first part of the proposition.

Now define $\tilde{R}^*(E) := \max_{S\in[0,E]} \tilde{R}(E|S)$, that is, the worst perceived remaining disutility of completing the task. Of course, for an unbiased person, the worst remaining disutility is always at the start when the most work remains to be done, but this won't necessarily hold for projection-biased people. Notice that if $\tilde{R}^*(E) < B$, then the task is always perceived worth doing and therefore is completed. If $\tilde{R}^*(E) > B$, then the task is definitely not completed, since at some point the person perceives it not worth doing. Finally, if $\tilde{R}(E|0) < B$ and $\tilde{R}^*(E) > B$, then the person starts the task, but does not complete it.

Let $\mathcal{E} := \{E \geq 0 : \tilde{R}^*(E) > B > \tilde{R}(E|0)\}$. I will show that $\mathcal{E} = (E_H, \infty)$ for some finite $E_H > 0$, which proves that if $E > E_H$, then we can pick B in the non-empty interval $(\tilde{R}(E|0), \tilde{R}^*(E))$ and the person starts the task but fails to complete it. Moreover, I will show that if $E < E_H$, then $\tilde{R}(E|0) > \tilde{R}(E|S)\ \forall S \in (0, E]$, which means that if the person starts the task, she also completes it.

First, let us show that $\mathcal{E}$ is not the empty set. Pick some $S > 0$ such that $D'(S) > 0$. Notice that $\tilde{R}(E|S) - \tilde{R}(E|0) = \tilde{D}(E|S) - \tilde{D}(S|S) - \tilde{D}(E|0) = (1-\alpha)(D(E) - D(S) -$

$D(E)) + \alpha E(D'(S) - D'(0)) + \alpha SD'(S) = -D(S)(1-\alpha) + \alpha SD'(S) + \alpha E(D'(S) - D'(0))$. Since $D'(S) - D'(0) > 0$, this expression becomes positive for sufficiently large E, say for $E > \bar{E}$, so that $\tilde{R}(E|S) - \tilde{R}(E|0) > 0$ for all $E > \bar{E}$. Thus $\mathcal{E}$ is not empty.

Further, notice that $\tilde{R}^*(E) - \tilde{R}(E|0) > 0$ and $E' > E$ we have

$$
\begin{aligned}
&(1-\alpha)D(E) + \alpha ED'(S) - (1-\alpha)D(S) - \alpha SD'(S) > (1-\alpha)D(E) + \alpha ED'(0) \\
\iff &\alpha E(D'(S) - D'(0)) > (1-\alpha)D(S) + \alpha SD'(S) \\
\iff &\alpha E'(D'(S) - D'(0)) > (1-\alpha)D(S) + \alpha SD'(S) \\
&\iff \tilde{R}^*(E') > \tilde{R}(E'|0)
\end{aligned}
$$

Thus if $E \in \mathcal{E}$, then $E' \in \mathcal{E}$. Let $E_H = \liminf \mathcal{E}$. Then if $E > E_H$, since by definition of E_H, there is some $E' \in (E_H, E)$ such that $E' \in \mathcal{E}$, so that $E \in \mathcal{E}$. Moreover, $E_H \notin \mathcal{E}$, since either $E_H = 0$ (in which case it is obvious) or $E_H > 0$. If $E_H > 0$ and $E_H \in \mathcal{E}$, then $\tilde{R}(E_H|S) > \tilde{R}(E_H|0)$ for some $S > 0$, and thus $\tilde{R}(E_H - \varepsilon|S) > \tilde{R}(E_H - \varepsilon|0)$ for sufficiently small ε, which contradicts the definition of E_H as $\liminf$.

Finally, note that when $E < E_H$, we must have that $0 > \tilde{R}(E|S) - \tilde{R}(E|0)\ \forall S > 0$. If not, then $\tilde{R}(E|S) - \tilde{R}(E|0) = 0$ and we know that the LHS strictly increases in E, which would imply that $E_H \in \mathcal{E}$. And thus we are done.

□

Proof of proposition 3:

Proof. The person – by assumption – first works on the first task, and then on the second task. As long as she works on first task she solves the following first order conditions:

$$\tilde{D}'(\tilde{E}_1^*(S) + \tilde{E}_2^*(S)|S) = B_1'(\tilde{E}_1^*(S)) = B_2'(\tilde{E}_2^*)$$

She switches to the second task when the amount she has worked on the first task so far, S, is equal to how much she thinks she should optimally work on the first task, $\tilde{E}_1^*(S)$. Let $\tilde{E}_1^*$ be the actual time by which she switches to the second task , which must satisfy the first order condition when S is equal to it. Thus:

$$\tilde{D}'(\tilde{E}_1^* + \tilde{E}_{2|1}^*|\tilde{E}_1^*) = B_1'(\tilde{E}_1^*) = B_2'(\tilde{E}_{2|1}^*)$$

where $\tilde{E}_{2|1}^*$ is the amount she plans to work on task 2 when she switches. Note that $B_1'(\tilde{E}_1^*) = B_2'(\tilde{E}_{2|1}^*)$ implies that $\tilde{E}_1^* > E_1^* \iff \tilde{E}_{2|1}^* > E_2^* \iff \tilde{E}_1^* + \tilde{E}_{2|1}^* > E_1^* + E_2^*$, since the B_i^* are strictly convex.

We can now show that $\tilde{E}_1^* > E_1^*$. Suppose not. Then we have

$$\begin{aligned}
\tilde{D}'(\tilde{E}_1^* + \tilde{E}_{2|1}^* | \tilde{E}_1^*) &< \tilde{D}'(\tilde{E}_1^* + \tilde{E}_{2|1}^* | \tilde{E}_1^* + \tilde{E}_{2|1}^*) \\
&= D'(\tilde{E}_1^* + \tilde{E}_{2|1}^*) \\
&< D'(E_1^* + E_2^*) \\
&= B'(E_1^*) \\
&< B'(\tilde{E}_1^*)
\end{aligned}$$

which shows that it does not satisfy the first order condition. Thus $\tilde{E}_1^* > E_1^*$.

Once she switches, she keeps working on the second task, she will only want to reduce E_1, which she cannot do. Thus she takes $\tilde{E}_1^*$ as a given. Thus she now simply solves the first order condition

$$\tilde{D}'(\tilde{E}_1^* + \tilde{E}_2^*(S)|S) = B_2'(\tilde{E}_2^*(S))$$

and as before, she stops once S is equal to the total effort she feels she should exert, that is $S = \tilde{E}_1^* + \tilde{E}_2^*$:

$$\begin{aligned}
&\tilde{D}'(\tilde{E}_1^* + \tilde{E}_2^* | \tilde{E}_1^* + \tilde{E}_2^*) = B_2'(\tilde{E}_2^*) \\
\iff\ & D'(\tilde{E}_1^* + \tilde{E}_2^*) = B'(\tilde{E}_2^*) \\
\iff\ & D'(\tilde{E}_1^* + \tilde{E}_2^*) - B'(\tilde{E}_2^*) = 0 \\
\iff\ & D'(E_1^* + \tilde{E}_2^* + x) - B'(\tilde{E}_2^*) = 0 \text{ for some } x > 0
\end{aligned}$$

Since $D'(E_1^* + E_2^* + x) > D'(E_1^* + E_2^*) = B'(E_2^*)$, we must have that $\tilde{E}_2^* < E_2^*$. Otherwise, $D'(E_1^* + \tilde{E}_2^* + x) > B'(E_2^*) > B'(\tilde{E}^*)$ and the first order condition can't hold.

Therefore, $D'(\tilde{E}_1^* + \tilde{E}_2^*) = B'(\tilde{E}_2^*) > B'(E_2^*) = D'(E_1^* + E_2^*)$, so that $\tilde{E}_1^* + \tilde{E}_2^* > E_1^* + E_2^*$ and we are done. □

A.2 Proofs for Section 4

A.2.1 Proofs of Results on Multi-Day All-or-Nothing Task

In the first part of this section, I prove the propositions 4 and 5 for a continuous-time version of the task, which I show in the second part to be the limit case as $T \to \infty$, which then proves the original propositions.

Proofs of Results on Multi-Day All-or-Nothing Task: Continuous Time Setup

First, I need some notation to talk succinctly about tasks and I need to define the continuous-time problem.

Definition 7. *A discrete-time task requiring total effort $E_0 \cdot T$, paying rewards $B_0 \cdot T$ if completed by the end of day T is written as task (E_0, B_0, T). A continuous-time task requiring total effort E_0 and paying total rewards B_0 if completed by time 1 is written as (E_0, B_0).*

Definition 8. *Consider a person facing a task (E, B, T). Then the* continuous time problem *corresponding to this discrete time problem is as follows. At every time $x \in [0, 1)$ a person chooses instantaneous (flow) effort e_x, based on receiving total benefits B for completing total effort E by time $x = 1$. Let E_x for $x \in [0, 1)$ be the effort remaining at time x – that is $E_x = E - \int_0^x e_{x'} dx'$. The initial condition is $E_0 = E$ and a task is completed if $E_1 = 0$. Instantaneous effort e_x satisfies the following:*

$$e_x = \begin{cases} 0, \text{ if } G(x, 0, E_x) > B \\ \frac{E_x}{1-x}, \text{ if } G(x, \frac{E_x}{1-x}, E_x) < B \\ e_x^* \text{ otherwise, with } G(x, e_x^*, E_x) = B \end{cases}$$

where $G(x, s, E) = (1 - \alpha)(1 - x) \cdot D(\frac{E}{1-x}) + \alpha D'(s)E$.

Notice that e_x solves a similar maximization problem, as if at instant x she did work more and more and perceive effort at later times as more costly. Intuitively, the difference with the discrete time setup is that we do not need to take into account that the more a person has worked, the less work there remains to do, since the instantaneous work doesn't matter, which ensures that the perceived remaining disutility always increases during the 'period' at time x. Readers who do not like this interpretation can simply treat the continuous time problem as an analytical device.

The proofs will refer to the times τ_0 and τ_F, which are (roughly) the total time a person spends not working at all or the total time a person works efficiently *given* how much she worked up to a given time x. Formally and concretely, we have the following definition:

Definition 9. *Let $\tau_0(E_0, B_0) := \liminf\{1 - x : x \in [0, 1)$ and $G(x', 0, E_{x'}(E_0, B_0)) < B_0 \; \forall x' < x\}$.*

Let $\tau_F(E_0, B_0) := \liminf\{1 - x : x \in [0, 1)$ and $G(x', \frac{E_{x'}(E_0,B_0)}{1-x'}, E_{x'}(E_0, B_0)) > B_0 \; \forall x' < x\}$.

With this, let us first prove that $E_x(E_0, B_0)$ for $x < 1$ is Lipschitz continuous in a neighborhood of (E_0, B_0, x) and that it is increasing in E_0.

I will use the following theorem (from https://www.math.washington.edu/~burke/crs/555/555_notes/continuity.pdf) to prove continuity.

Theorem 1. *Consider the initial value problem*

$$x' = f(t, x, \mu), \; x(t_0) = y$$

If f is continuous in t, x, μ and Lipschitz in x with Lipschitz constant independent of t and μ, then $x(t, \mu, y)$ is continuous in (t, μ, y) jointly.

Then we get continuity as follows:

Lemma 2. *Suppose that $D'(0) > 0$. The solution $E_x(E, B)$ to the continuous-time problem restricted to $x \in [0, 1-\varepsilon]$ with $\varepsilon > 0$ exists and is Lipschitz continuous in x, E, and B, on $[0, 1-\varepsilon] \times [\underline{E}, \bar{E}] \times [0, \infty]$, for some $\bar{E} > \underline{E} > 0$.*

Proof. We have that $\dot{E}_x = -e(x, E_x, B)$ with

$$e(x, E_x, B) = \begin{cases} 0, \text{ if } G(x, 0, E_x) > B \\ \frac{E_x}{1-x}, \text{ if } G(x, \frac{E_x}{1-x}, E_x) < B \\ f(x, E_x, B) \text{ otherwise} \end{cases}$$

where $G(x, s, E) = (1-\alpha)(1-x) \cdot D(\frac{E}{1-x}) + \alpha D'(s)E$, and $f(x, E, B) := (D')^{-1}\left(\frac{B-(1-\alpha)(1-x)D(E)}{\alpha E}\right)$. Given theorem 1, we only need to show that $e_x(x, E, B)$ is continuous in t, E, and B, and Lipschitz continuous in E. Notice that G and f are continuous functions.

First, notice that when $G(x, 0, E) = B$, by definition of G and f we have that $f(x, E, B) = (D')^{-1}(D'(0)) = 0$, and similarly when $G(x, \frac{E}{1-x}, E) = B$, we have that $f(x, E, B) = \frac{E}{1-x}$. Thus $e(x, E, B)$ restricted to $A := \{(x, E, B) : G(x, 0, E) \geq B\}$ is the constant 0 function, $e(x, E, B)$ restricted to $B := \{(x, E, B) : G(x, \frac{E}{1-x}, E) \leq B\}$ is equal to $\frac{E}{1-x}$, and $e(x, E, B)$ restricted to $C := \{(x, E, B) : G(x, 0, E) \leq B \text{ and } G(x, \frac{E}{1-x}, E) \geq B\}$ is equal to $f(x, E, B)$.

If we can show that $e(x, E, B)$ restricted to A, B, and C is Lipschitz in all parameters (which is stronger than what we need), then $e(x, E, B)$ is Lipschitz continuous in all parameters over the union of A, B, and C. The reason is that all three regions are closed, and thus contain their limit points. Here is why. Suppose we have two points $\boldsymbol{x} = (x, E, B)$ and $\boldsymbol{x'} = (x', E', B')$ and we want to show that $|e(x, E, B) - e(x', E', B')| < K(|x-x'| + |E-E'| + |B-B'|)$ for some K. First, if both points are in the same region, then this immediately holds, by the assumption that the function is Lipschitz in that region. Now suppose that the two points are in regions A and C. These two regions share a common border. Thus there exists some point $\boldsymbol{x''} = (x'', E'', B'')$ on the line connecting the two points that belongs to both regions (this is the part that requires both A and C to be closed), so that $|e(x, E, B) - e(x', E', B')| = |e(x, E, B) - e(x'', E'', B'') + e(x'', E'', B'') - e(x', E', B')| \leq |e(x, E, B) - e(x'', E'', B'')| + |e(x'', E'', B'') - e(x', E', B')| < K(|x-x''| + |E-E''| + |B-B''| + |x''-x'| + |E''-E'| + |B''-B'|) = K(|x-x'| + |E-E'| + |B-B'|)$, where $|x-x''| + |x''-x'| = |x-x'|$ because the point $\vec{x''}$ lies between the two points (is a convex combination of) $\boldsymbol{x}$ and $\boldsymbol{x'}$.

Restricting ourselves to $E_0 \in [\underline{E}, \bar{E}]$, it is clear that $e(x, E, B)$ is Lipschitz on A, where it is constant. It is equally clear that $e(x, E, B)$ is Lipschitz continuous on B since (by assumption) we are only considering $x \leq 1-\varepsilon$, that is $1-x \geq \varepsilon$.

Finally, $e(x, E, B)$ is Lipschitz continuous on C, since (by assumption) $D'(0) > 0$. It is clear that $E_x \leq E_0$, so we can pick $\bar{E} > E_0$. We also need to show that, starting with $E_0 \in [\underline{E}, \bar{E}]$, we will not fall below $\underline{E}$ before time x. Given that the maximum instantaneous effort is given by $\frac{E_x}{1-x}$ it is not hard to see that at most a fraction x of the total effort will be completed by time x (the efficient amount, conditional on trying to complete the task).[18] Thus if $E_0 \geq \frac{1}{\varepsilon}\underline{E}$, then E_x will be larger than $\underline{E}$ for all $x \leq 1-\varepsilon$. Thus we have shown that $e(x, E, B)$ is Lipschitz continuous when $x \leq 1-\varepsilon$, $E \in [\underline{E}, \bar{E}]$ and $B \geq 0$, for any $\varepsilon > 0$, $\underline{E} > 0$, $\bar{E} > 0$. $\square$

Lemma 3. *If $G(x, 0, E_x) > B$, then $G(x', 0, E_{x'}) > B$ for all $x' \geq x$. Similarly, if $G(x, \frac{E_x}{1-x}, E_x) < B$, then $G(x', \frac{E_{x'}}{1-x'}, E_{x'}) < B$ for all $x' \geq x$.*

Proof. Suppose not. Then there exists $1 > x' > x$ such that $G(x', 0, E_{x'}) \leq B$. Note that because E_x is continuous on $[0, x'+\varepsilon]$ for sufficiently small ε, and because G is continuous in all its arguments, we know that $G(x+\varepsilon_1, 0, E_{x+\varepsilon_1}) > B$ for sufficiently small ε_1. Now let $x^* := \liminf\{x' > x : G(x', 0, E_{x'}) \leq B\}$, then $x^* > x$. Moreover, $G(y, 0, E_y) > B$ for all $x \leq y < x^*$ and therefore $e_y = 0$. Thus $E_{x^*} = E_x - \int_x^{x^*} e_y dy = E_x$. Hence we have that $G(x^*, 0, E_{x^*}) = G(x^*, 0, E_x) > G(x, 0, E_x) > B$, since G is strictly increasing in x. But then by continuity of E_x and G, we have that $G(x^*+\varepsilon_2, 0, E_{x^*+\varepsilon_2}) > B$ for sufficiently small $varepsilon_2$, which contradicts the definition of x^*.

A similar argument works for the second part of the lemma. $\square$

Now let us prove that E_x is increasing in E_0:

Lemma 4. *For a fixed $x < 1$ and $B_0 > 0$, $E_x(E_0, B_0)$ is increasing in E_0.*

Proof. Let $\Delta_x = E_x(E_0', B_0) - E_x(E_0, B_0)$ for some $E_0' > E_0 > 0$. We need to show that $\Delta_x \geq 0$.

Notice that $\Delta_0 = E_0' - E_0 > 0$ and that $\frac{d\Delta}{dx} = e_x' - e_x$. Since E_x is continuous, we have that $\Delta_x > 0$ for all $x < \varepsilon$ at least. Suppose that the claim is false, so that there $x^* := \liminf\{x : \Delta_{x^*} = 0\}$ exists. Then $x^* \geq \varepsilon > 0$. Note that for all $x < x^*$, $\Delta_{x^*} > 0$. Thus $E_x' > E_x$. Therefore if $G(x, 0, E_x) > B$, then $G(x, 0, E_x')$, and hence $e_x = e_x' = 0$. If $G(x, e_x, E_x) = B$, then $G(x, e_x, E_x') > B$, since G is increasing in E, and therefore $e_x' < e_x$ because G is increasing in its second argument. (Note that it could have been possible that the third condition holds, i.e. $G(x, \frac{E_x'}{1-x}, E_x') < B$, but this would imply $G(x, \frac{E_x}{1-x}, E_x) < B$ contradicting $G(x, e_x, E_x) = B$.) Finally, if $G(x, \frac{E_x}{1-x}, E_x) < B$, then $e_x = \frac{E_x}{1-x}$ and $e_x' \leq \frac{E_x'}{1-x}$. Thus it is easy to see that $\frac{d\Delta_x}{dx} \geq \frac{E_x - E_x'}{1-x} = \frac{\Delta_x}{1-x}$, and therefore for $x < x^*$ $\Delta_{x^*} = D_x + \int_x^{x^*} \frac{d\Delta_x}{dx} dx \geq D_x + \int_x^{x^*} \frac{\Delta_y}{1-y} dy$. We know by the definition of x^* that $\Delta_x > 0$ for $x < x^*$. Pick $x \in [x^* - \delta, x^*)$ that achieves a maximum of Δ_x (possible since Δ_x is continuous). Then we have that $\Delta_{x^*} \geq D_x - \int_x^{x^*} \frac{\Delta_y}{1-y} dy \geq D_x - D_x \frac{\delta}{1-x^*} > \frac{1}{2} D_x > 0$ when $\delta < \frac{1}{2}(1-x^*)$. Thus $\Delta_{x^*} > 0$ and therefore (by continuity) $\Delta_{x^*+\varepsilon} > 0$ for some small $\varepsilon > 0$, which contradicts the definition of x^*. Thus the claim is proved. $\square$

Lemma 5. $E_x(E_0, B_0) \geq E_{x'} \frac{1-x'}{1-x'}$ *for* $1 > x > x' \geq 0$.

[18] This statement is proved in Lemma 5

Proof. Notice that for any ε, we can choose sufficiently large B_0 so that $G(x', \frac{E_{x'}}{1-x'}, E_{x'}) < B_0$ for all $x' < 1-\varepsilon$. Therefore by lemma 3 we know that $G(x, \frac{E_x}{1-x}, E_x) < B_0$ for all $x' < x$, and thus $e_x = \frac{E_x}{1-x}$. We can solve the differential equation

$$\dot{E}_x = -\frac{E_x}{1-x} \iff \frac{\dot{E}_x}{E_x} = -\frac{1}{1-x} \iff \log(E_x) - \log(E_{x'}) = -\int_{x'}^{x} \frac{1}{1-y}dy$$
$$\iff \frac{E_x}{E_{x'}} = \frac{1-x}{1-x'}$$

which proves that equality is possible. The same argument holds for all $B > B_0$. Thus, if we can show that $E_x(E_0, B_0)$ is decreasing in B_0 we are done, since then $E_x(E_0, B_0) \geq E_x(E_0, B^*) = E_{x'}(E_0, B_0)\frac{1-x}{1-x'}$.

The proof that $E_x(E_0, B_0)$ is decreasing in B_0 is similar to the proof of lemma 4, and thus I omit it. □

Now let us prove that $\tau_0(E_0, B_0)$ and $\tau_F(E_0, B_0)$ are continuous in E_0.

Lemma 6. *Suppose $D'(E) \to \infty$ as $E \to \infty$ and $D'(0) > 0$. If $\tau_0(E_0, B_0) \in (0, 1)$, then $\tau_0(E, B)$ is continuous and increasing in E in a neighborhood of E_0. If $\tau_F(E_0, B_0) \in (0, 1)$, then $\tau_F(E, B)$ is continuous and decreasing in E in a neighborhood of E_0.*

Proof. The proofs are essentially identical for τ_0 and τ_F, so I only prove the first. Remember that

$$\tau_0(E_0, B_0) = \limsup\{\tau : 1-\tau \in [0, 1) \text{ and } G(x', 0, E_{x'}(E_0, B_0)) < B_0 \ \forall x' > 1-\tau\} = \limsup \Gamma_0$$

Notice that $0 \in \Gamma_0$, thus τ_0 always exists. Suppose $\tau_0 \in (0, 1)$. Then take $x > 1-\tau_0$. Suppose $E_0^{'} > E_0$ and let $\tau_0 := \tau_0(E_0, B_0)$ and $\tau_0^{'} := \tau_0(E_0^{'}, B_0)$ and similarly for Γ_0 and $\Gamma_0^{'}$. Note that if $G(x, 0, E_x(E_0, B_0)) > B$ then, by lemma 4, $E_x^{'} \geq E_x$, and thus (since G is increasing in its third argument) $G(x, 0, E_x(E_0, B_0)) > B$. Therefore if $\tau \notin \Gamma_0$, then there exists some $x' < 1-\tau$ with $G(x, 0, E_x) > B$ and therefore $G(x, 0, E_x^{'}) > B$ so that $\tau \notin \Gamma_0^{'}$. Hence $\Gamma_0^{'} \subset \Gamma_0$ and thus $\tau_0^{'} \geq \tau_0$.

Further note that if $\delta > 0$, then we must have that there is some $x' \in [1-\tau_0^{'}, 1-\tau_0^{'}+\delta]$ such that $G(x', 0, E_{x'}^{'}) > B$, since if this is not the case then $G(x', 0, E_{x'}^{'}) \leq B$ for all $x' \leq 1-\tau_0^{'}+\delta$ (if not, then pick some counterexample x'' and then by lemma 3 all $x' > x''$ have $G(x', 0, E_{x'}^{'}) > B$ contradicting the initial statement). But this means that $\tau_0^{'}-\delta \in \Gamma_0^{'}$ contradicting the definition of $\tau_0^{'}$. Thus there is some $x' \in [1-\tau_0^{'}, 1-\tau_0^{'}+\delta]$ with $G(x', 0, E_{x'}^{'}) > B$. By continuity of $E_{x'}$ in E_0 and of G in E, we can find $E_0^{'}$ close to E_0 such that $G(x', 0, E_{x'}) > B$, so that $1-\tau_0 < x' < 1-\tau_0^{'}+\delta$. Thus we have that $\tau_0^{'}-\delta < \tau_0 \leq \tau_0^{'}$, proving the claim. □

Now let us prove the equivalent to proposition 4 but for continuous time tasks. Let us split the proof in two parts.

Proposition 11. *The disutility of effort is strictly convex. Consider a task* (E_0, B_0) *with* $E_0 > 0$ *fixed. Then there exist* $B_H(E_0) > B_C(E_0) > B_L(E_0) > 0$ *such that*

- *if* $B > B_H$*, then the task is completed efficiently, i.e.* $\tau_F = 1$*.*
- *if* $B_H > B > B_C$*, then* $\tau_F(E_0, B_0) \in (0,1)$ *and the task is completed.*
- *if* $B_C > B > B_L$*, then* $\tau_0(E_0, B_0) \in (0,1)$ *and the task is not completed.*
- *if* $B_L > B$*, then no effort is spent on the task, i.e.* $\tau_0 = 1$*.*

Proof. Let $B_L = (1-\alpha)D(E_0) + \alpha D'(0)E_0$. Then $G(0,0,E_0) = B_L$ and therefore if $B < B_L$ we have $G(0,0,E_0) > B_L$ and hence by lemma 3 we know that $G(x,0,E_x) > B_L$ for all $x \geq 0$. Hence $e_x = 0$ and $\tau_0(E_0, B_0) = 0$. Similarly, if $B_H = (1-\alpha)D(E_0) + \alpha D'(E_0)E_0$, then $G(0, E_0, E_0) = B_H$. Hence if $B > B_H$ we have $G(0, E_0, E_0) < B$ and again by lemma 3 this holds for all $x \geq 0$ and thus $\tau_F = 1$ and $e_x = E_0$ (this last part in effect requires solving the same differential equation as we did in lemma 5, which I omit).

Moreover, note that if $B < B_H$ then we have that $G(0, E_0, E_0) > B$ and thus (by continuity of E_x and G) we have that $G(x, \frac{E_x}{1-x}, E_x) > B$ for all sufficiently small x. Therefore, $\tau_F < 1$. Similarly, if $B > B_L$ we have that $\tau_0 < 1$.

It is clear that if $\tau_0 > 0$ then $\tau_F = 0$ and if $\tau_F > 0$ then $\tau_0 = 0$. Let $B_{C,0} = \liminf\{B : \tau_0(E_0, B) = 0\}$. Then because τ_0 is decreasing in B_0, we know that if $B < B_{C,0}$ then $\tau_0(E_0, B) > 0$, since if $\tau_0(E_0, B) = 0$, then $\tau_0(E_0, B') = 0$ for all $B' \geq B$, contradicting the definition of $B_{C,0}$. Similarly we can define $B_{C,F} = \limsup\{B : \tau_F(E_0, B) = 0\}$ and show that if $B > B_{C,F}$ then $\tau_F > 0$.

To finish the proof, we need to show that $B_{C,F} = B_{C,0}$. Notice that if $B_0 \in [B_{C,0}, B_{C,F}]$ we have that $\tau_0 = 0$ and $\tau_F = 0$. Therefore $G(x, e_x, E_x(E_0, B_0)) = B$ for all $x < 1$. Suppose that $B_{C,0} < B_{C,F}$. Since $G(x, e_{x,0}, E_{x,0}) = B_{C,0} < B_{C,F} = G(x, e_{x,F}, E_{x,F})$ for all x, we must have that $e_{x,F} > e_{x,0}$ or $E_{x,F} > E_{x,0}$ for every x. By continuity of E_x in x and G in E, we can pick $\varepsilon > 0$ such that $G(x, e_{x,0}, E_{x,F})$ is arbitrarily close to $G(x, e_{x,0}, E_{x,0}) = B_{C,0}$ so that $e_{x,F} > e_{x,0}$ for all $x < \varepsilon$. Thus $E_{x,F} < E_{x,0}$ and we can show that $e_{x,F} > e_{x,0}$ for all x. Suppose not, then we must have that $E_{x,F} > E_{x,0}$ for some x and therefore there exists a smallest $x^* > \varepsilon$ such that $E_{x^*,F} = E_{x^*,0}$. But $e_{x,F} > e_{x,0}$ for all $x < x^*$, therefore $E_{x^*,F} < E_{x^*,0}$, which is a contradiction.

Thus we have shown that $E_{x,F} < E_{x,0}$ and that $e_{x,F} > e_{x,0}$ for all $x > 0$. Let $\delta = E_{\frac{1}{2},0} - E_{\frac{1}{2},F} > 0$, then $E_{x,0} - E_{x,F} \geq \delta$ for $x > \frac{1}{2}$ and therefore $E_{x,0} \geq \delta > 0$ for all x. Therefore $D(\frac{E_{x,0}}{1-x})(1-x) \geq D(\frac{\frac{1}{2}}{1-x})\frac{1-x}{\frac{1}{2}} \to \infty$ by lemma 7. But this means that $G(x, 0, E_{x,0}) \to \infty$ and therefore that $G(x, 0, E_{x,0}) > B$ as $x \to 1$, so that $\tau_0 > 0$. Therefore, we cannot have that $B_{C,0} < B_{C,F}$ and we are done. ◻

Here is the lemma I referred to at the end of the previous proof.

Lemma 7. *Let* D *be convex with* $D'(e) \to \infty$ *as* $e \to \infty$*. Then* $\forall K > 0$, $\exists E$ *s.t.* $D(e) > K \cdot e$ $\forall e > E$*. That is,* $D(e)/e \to \infty$ *as* $e \to \infty$*.*

Proof. Since $D'(e) \to \infty$, pick E s.t. $D'(\frac{E}{2}) > 2 \cdot K$. Then for $e > E$

$$D(e) = \int_0^e D'(s)ds \geq \int_{E/2}^e D'(s)ds \geq \int_{E/2}^E 2 \cdot Kds \geq \frac{e}{2} 2 \cdot K = e \cdot K$$

□

Now let us show that the utility is continuous and decreasing on (B_L, B_C) and continuous and increasing on (B_C, B_H).

Lemma 8. *The utility* $u_0(E_0, B_0) := -\int_0^1 D(e_x)dx$ *is continuous and decreasing on* $(B_L(E_0), B_C(E_0))$ *and the utility* $u_F(E_0, B_0) := B - \int_0^1 D(e_x)dx$ *is continuous and increasing on* $(B_C(E_0), B_H(E_0))$.

Proof. Notice that when $B \in (B_L, B_C)$ then we know that $\tau_0 \in (0,1)$ and the task is not completed, hence the definition of the utility as u_0 is correct. Moreover $u_0 = \int_0^{1-\tau_0} D(e_x)dx$. We can show that τ_0 and E_x are continuous and decreasing in B_0. Picking $B_0 < B_0^{'}$, we therefore have that $\tau_0^{'} < \tau_0$ and that for $x \leq 1-\tau_0$ we have $G(x, e_x, E_x) = B_0 < B_0^{'} = G(x, e_x^{'}, E_x^{'})$. Since $E_x^{'} \leq E_x$ we therefore have that $e_x^{'} > e_x$ and therefore $u_0^{'} > \int_0^{1-\tau_0} D(e_x^{'})dx > \int_0^{1-\tau_0} D(e_x)dx = u_0$. Moreover, if $B_0^{'}$ is close to B_0 then E_x is close to $E_x^{'}$ by Lipschitz continuity and therefore $e_x^{'}$ and e_x are close together, since e_x is Lipschitz continuous in all the parameters as well (I haven't shown this in detail, but this is where I use the condition $D'(0) > 0$). Therefore the $u_0^{'}$ and u_0 are close.

Now suppose B_0, $B_0^{'} \in (B_C, B_H)$ then $\tau_F \in (0,1)$. Let $B_0 < B_0^{'}$. We can show in a similar way as before that $\tau_F^{'} > \tau_0^{'}$ and that $e_x^{'} > e_x$ for $x \leq 1 - \tau_0^{'}$. Then notice that $\int_0^1 e_x = E_0 = \int_0^1 e_x^{'}$. Let $F(e) = \int_0^1 \mathbb{1}(e_x \leq e)de$ and $G(e) = \int_0^1 \mathbb{1}(e_x^{'} \leq e)de$. Let $\bar{e} = e_{1-\tau_F^{'}}$. Then if $e < \bar{e}$, $F(e) < G(e)$ (this can be proved rigorously using continuity of e_x, but it is intuitive noting that $e_x^{'} > e_x$ for all $x \leq 1 - \tau_0^{'}$). And since $F(\bar{e}) = 1$ and $G(\bar{e}) < 1$ we have that $F(e) \geq G(e)$ for $e \geq \bar{e}$. Therefore G is a mean-preserving spread of F and thus the disutility for e_x is higher than for $e_x^{'}$. Continuity follows again by noting that, until time $1 - \tau_0$, e_x is Lipschitz continuous in all parameters, and thereafter it is constant. Therefore the utility is Lipschitz continuous. □

I will need the following two lemmas to prove the second part of the proposition.

Lemma 9. *Let D be convex and such that $D'(e) \to \infty$ as $e \to \infty$. Fix B and $\varepsilon > 0$. Let e_ε be s.t.*

$$D(e_\varepsilon) \cdot \varepsilon = B \tag{11}$$

Then $e_\varepsilon \cdot \varepsilon \to 0$ as $\varepsilon \to 0$.

Proof. First note that as ε goes to 0, e_ε goes to ∞, since if it was bounded, then $D(e_\varepsilon) \cdot \varepsilon$ would go to 0. By lemma 7, we know that $\frac{D(e_\varepsilon)}{e_\varepsilon} \to \infty$. Dividing both sides of equation 11 by $\varepsilon \cdot e_\varepsilon$ yields

$$D(e_\varepsilon)/e_\varepsilon = \frac{B}{e_\varepsilon \cdot \varepsilon} \iff e_\varepsilon \cdot \varepsilon = \frac{B}{D(e_\varepsilon)/e_\varepsilon} \to 0$$

which proves the claim. □

Lemma 10. *Take the continuous time problem with effort $E > 0$. Then for every $\varepsilon > 0$ there exists B^* s.t.*

$$D\left(\frac{E_{1-\varepsilon}}{\varepsilon}\right) \cdot \varepsilon = B^* \tag{12}$$

Proof. We know from lemma 2 that $E_t(B)$ is a continuous and decreasing function in B. Moreover $E_t(0) = E$, and $E_t(B) \to (1-t)E$ as $B \to \infty$, since the task eventually will be completed efficiently. Let $\bar{B}$ be such that it is completed efficiently and worth it.

Moreover, $D\left(\frac{E}{1-t}\right)(1-t) > 0$ and $D(E)(1-t) < \bar{B}$. Since both sides are continuous in B, there is a $B^* \in (0, \bar{B})$ s.t. equation (12) holds, and we are done. □

Here is the proof of the second part of proposition 4 in the continuous-time setting.

Proposition 12. *Suppose $D'(0) > 0$ and $D'(E) \to \infty$ as $E \to \infty$. Then $\lim_{B_0 \to B_C^-} u_0(E_0, B_0) \leq -D(E_0)$.*

Proof. Since τ_0 is continuous and decreasing on $(B_L(E_0), B_H(E_0))$, and since τ_0 can be 0 and 1, we know that for every $\tau \in (0,1)$ there is some $B_0 \in (B_L(E_0), B_H(E_0))$ such that $\tau_0(E_0, B_0) = \tau$. Notice that at time τ_0 we have that $G(1-\tau_0, 0, \frac{E_{1-\tau_0}}{\tau_0}) = B_0$. Therefore $(1-\alpha)D(\frac{E_{1-\tau_0}}{\tau_0})\tau_0 + \alpha D'(0)E_{1-\tau_0} = B_0$. As $\tau_0 \to 0$, we must therefore have that $E_{1-\tau_0} \to 0$, since otherwise $G \to \infty$. But this means that almost all the work gets done before time $1-\tau_0$ for which the least disutility is $D(E_0 - \varepsilon) > D(E_0) - \delta$ for sufficiently small ε (i.e. τ_0 sufficiently close to 1). Therefore the disutility is at least $D(E_0) - \delta$ for arbitrary δ. Hence the result holds. □

Here is the proof of proposition 5.

Proof. First, I show that in the continuous time problem the statement holds. That is, if $D''' < 0$ and $D(E) \leq B$, the task is completed, whereas if $D''' > 0$ and $D(E) \geq B$, the task is not completed.

Claim: When $D''' < 0$ and $D(E) < B$, then $D(\frac{E_x}{1-x})(1-x)$ strictly decreases with x. When $D''' > 0$ and $D(E) > B$, then $D(\frac{E_x}{1-x})(1-x)$ strictly increases with x.

Proof of the claim:

$$\begin{aligned}\frac{d}{dx} D\left(\frac{E_x}{1-x}\right)(1-x) &= D'\left(\frac{E_x}{1-x}\right)\left(\frac{\dot{E}_x}{1-x} + \frac{E_x}{(1-x)^2}\right)(1-x) - D\left(\frac{E_x}{1-x}\right) \\ &= D'\left(\frac{E_x}{1-x}\right)\left(\dot{E}_x + \frac{E_x}{1-x}\right) - D\left(\frac{E_x}{1-x}\right)\end{aligned}$$

First, notice that $\dot{E}_x$ is how much the person works instantaneously. Since the person stops working only once she overestimates the disutility of the task (before that, she won't stop since the task is worth doing), we have that in picture 1 we have $0 < \alpha B - \alpha A$, since $\alpha B - \alpha A$ is the difference between the perceived and the actual disutility $D\left(\frac{E_x}{1-x}\right)$. In picture 2 I highlight both $D(\frac{E_x}{1-x})$ and $D'(\frac{E_x}{1-x})(\frac{E_x}{1-x} + \dot{E}_x)$. This shows that $C - D$ in picture 3 is equal to $D'(\frac{E_x}{1-x})\left(\dot{E}_x + \frac{E_x}{1-x}\right) - D(\frac{E_x}{1-x})$. Finally, because D' is concave, we have that $A > D$ and $C > B$, so that $C > D$ and hence we are done.

The second part follows the same way, except that $\dot{E}_x$ is now such that the person still *underestimates* total disutility (since the task is not worth doing), and that $A < D$ and $B < C$ since D' is strictly convex.

This proves the claim.

From lemma 10 we know that for every x there is some B_x such that $D\left(\frac{E_x(E_0,B_x)}{1-x}\right)(1-x) = B_x$. By the claim we have just proved, we know that the LHS is decreasing in x if $D''' < 0$. Therefore if $x' > x$, we have that $D\left(\frac{E_{x'}(E_0,B_x)}{1-x'}\right)(1-x') < B_x$. Since $E_x(E_0, B_x)$ is decreasining in B_x and B_x is (trivially) increasing in B_x, we must have that $B_{x'} < B_x$ in order for $D\left(\frac{E_{x'}(E_0,B_{x'})}{1-x'}\right)(1-x') = B_{x'}$ to hold. Now I will show that $\lim_{x\to 1} B_x = B_C$ and therefore $B_C < B_x$, so that the task is completed for all B_x, which means the task is completed whenever the task is just weakly worth doing at time x, in particular at time 0, which proves the claim. A similar argument works for $D''' > 0$, as long as we can prove that $B_x \to B_C$.

Proof that $B_x \to B_C$:

Notice that $\tau_0(E_0, B_x) \leq 1 - x$, since $G(x, 0, \frac{E_x}{1-x}) \leq D\left(\frac{E_x(E_0,B_x)}{1-x}\right)(1-x) = B_x$. Similarly, $\tau_F(E_0, B_x) \leq 1 - x$. Thus, pick some $B \in (B_L, B_C)$. Then for sufficiently large x, we have that $\tau_0(E_0, B) > 1 - x$, and therefore (since τ_0 is decreasing in B) we have that $B_x > B$. Thus $\limsup B_x > B$ for all $B \in (B_L, B_C)$ so that $\limsup B_x \geq B_C$. A similar argument shows that $\liminf B_x \leq B_C$, so that $\lim B_x = B_C$. And we are done. $\square$

The propositions then follow if we can show that in the limit, the discrete time solution approaches the continuous time solution.

Multi-Day All-or-Nothing Tasks: Continuous to Discrete Time

With the following lemma, we are done, since it shows that as T goes to infinity, everything converges to the quantities in the continuous-time problem, for which we have shown that the statements in proposition 4 and in proposition 5 hold.

Lemma 11. *If $B_0 \in (B_L(E_0), B_C(E_0))$, then $\lim_{T\to\infty} \tau_0^D(E_0, B_0, T) = \tau_0(E_0, B_0)$ and $\lim_{T\to\infty} u^D(E_0, B_0, T) = u(E_0, B_0)$. If $B_0 \in (B_C(E_0), B_H(E_0))$, then $\lim_{T\to\infty} \tau_F(E_0, B_0, T) = \tau_F(E_0, B_0)$ and $\lim_{T\to\infty} u(E_0, B_0, T) = u(E_0, B_0)$.*

Proof. When $B_0 \in (B_L(E_0), B_C(E_0))$, we want to show that $\forall \delta > 0$, $\exists T^* > 0$ such that $\forall T > T^*$ we have both of the following

$$|\tau_0^D - \tau_0| < \delta$$
$$|u_0^D - u_0| < \delta$$

where $\tau_0^D := \tau_0^D(E_0, B_0, T)$ and $\tau_0 := \tau_0(E_0, B_0)$, and $u_0^D := u^D(E_0, B_0, T)$ and $u_0 := u(E_0, B_0)$. I will show this by finding large N and T^* such that $\tau_0^- := \tau_0(E_0 - \frac{1}{N}, B_0)$ and $\tau_0^+ := \tau_0(E_0 + \frac{1}{N}, B_0)$ τ_0 and $\tau_0^D \in [\tau_0^-, \tau_0^+]$ with $|\tau_0^- - \tau_0^+| < \delta$ and u_0, $u_0^D \in [u_0^-, u_0^+]$ and $|u_0^- - u_0^+| < \delta$, which will prove the claim.

First, by lemma 6, we know that τ_0 is continuous and decreasing and u_0 is continuous and increasing on $[0, E_C(B_0))$, thus (since $E_0 < E_C(B_0)$ given that the task is not completed) we can find a neighborhood $\mathcal{E}$ of E such that τ_0 and u_0 are continuous and increasing. Pick N_1 large enough so that $E_0 - \frac{1}{N}$ and $E_0 + \frac{1}{N} \in \mathcal{E}$. τ_0 and u_0 are continuous for $E \in \mathcal{E}$, thus we can choose N larger than N_1 large enough such that $|\tau_0^+ - \tau_0^-| < \delta$ and $|u_0^+ - u_0^-| < \delta$. Moreover, we can pick N_1 large enough as well so that $|\tau_0^- - \tau_0| < \frac{1}{2}\tau_0$. Since τ_0 is decreasing in in E, we have that $\tau_0 \in [\tau_0^+, \tau_0^-]$ and since u_0 is increasing in E we have that $u_0 \in [u_0^-, u_0^+]$.

To prove the remainder, pick T^* large enough so that $\frac{E_0+1}{\frac{1}{2}\tau_0^- T^*} < \frac{1}{N}$ and consider any $T > T^*$. Moreover, pick T^* large enough such that $\frac{1}{T} < \frac{1}{2}\tau_0$. Let $x_t = \frac{t-1}{T}$ for $t \in \{1, ..., T\}$. First, let us show that if $E_{x_t}^- \leq E_{x_t}^D - \frac{1}{N}$ and $E_{x_t}^D + \frac{1}{N} \leq E_{x_t}^+$, and $x_t \leq 1 - \frac{1}{2}\tau_0$, then

$$G^D(x_t, s, E_{x_t}^D) > G(x, s, E_x^-) \forall x \in [x_t, x_{t+1}]$$
$$G^D(x_t, s, E_{x_t}^D) < G(x, s, E_x^+) \forall x \in [x_t, x_{t+1}]$$

We have the following:

$$\begin{aligned} G(x, s, E_x^-) &\leq G(x_{t+1}, s, E_x^-) \text{ since } G \text{ increases in } x && (13)\\ &\leq G(x_{t+1}, s, E_{x_t}^-) \text{ since } G \text{ increases in } E \text{ which decreases in } x && (14)\\ &= (1-\alpha)D\left(\frac{E_{x_t}^-}{1-x_{t+1}}\right)(1-x_{t+1}) + \alpha D'(s)E_{x_t}^- && (15)\end{aligned}$$

Now note the following:

$$
\begin{aligned}
\frac{E^-_{x_t}}{1-x_{t+1}} = \frac{E^-_{x_t}\frac{1-x_t}{1-x_{t+1}}}{1-x_t} &= \frac{E^-_{x_t}(1+\frac{x_{t+1}-x_t}{1-x_{t+1}})}{1-x_t} \\
&= \frac{E^-_{x_t}+\frac{E^-_{x_t}}{T(1-x_{t+1})}}{1-x_t} \\
&\le \frac{E^-_{x_t}+\frac{E^-_{x_t}}{T\tau_0\frac{1}{2}}}{1-x_t} \text{ since } 1-x_{t+1} \ge \frac{1}{2}\tau_0 \\
&\le \frac{E^-_{x_t}+\frac{E_0+1}{T\tau_0\frac{1}{2}}}{1-x_t} \text{ since } E^-_{x_t} \le E^-_0 \le E_0 + \frac{1}{N} \le E_0+1 \\
&< \frac{E^-_{x_t}+\frac{1}{N}}{1-x_t} \text{ since } T > T^* \text{ and by how } T^* \text{ is chosen} \\
&\le \frac{E^D_{x_t}}{1-x_t} \text{ by the assumption we made}
\end{aligned}
$$

Therefore we have that $\forall s \le \frac{E^D_{x_t}}{1-x_t}$

$$
\begin{aligned}
(1-x_{t+1})D\left(\frac{E^-_{x_t}}{1-x_{t+1}}\right) < (1-x_{t+1})D\left(\frac{E^D_{x_t}}{1-x_t}\right) &= (1-x_t)D\left(\frac{E^D_{x_t}}{1-x_t}\right) - \frac{1}{T}D\left(\frac{E^D_{x_t}}{1-x_t}\right) \\
&\le (1-x_t)D\left(\frac{E^D_{x_t}}{1-x_t}\right) - \frac{1}{T}D(s)
\end{aligned}
$$

and that

$$
E^-_{x_t} \le E^D_{x_t} - \frac{1}{N} < E^D_{x_t} - \frac{E_0+1}{T\frac{1}{2}\tau_0} < E^D_{x_t} - \frac{E^D_{x_t}}{T(1-x_t)} \le E^D_{x_t} - \frac{s}{T}
$$

Thus, plugging this into equation (15), we get

$$
\begin{aligned}
G(x,s,E^-_x) &< (1-\alpha)\left(D(\frac{E^D_{x_t}}{1-x_t})(1-x_t) - \frac{1}{T}D(s)\right) + \alpha D'(s)\left(E^D_{x_t} - \frac{s}{T}\right) \\
&= G^D(x,s,E^D_{x_t})
\end{aligned}
$$

which proves the first of the two statements. Next we have:

$$
\begin{aligned}
G(x,s,E^+_x) &\ge G(x_t,s,E^+_x) \text{ because } G \text{ is increasing in } x && (16)\\
&\ge G(x_t,s,E^+_{x_{t+1}}) \text{ because } G \text{ is increasing in } E && (17)
\end{aligned}
$$

We also have that $E^+_{x_{t+1}} \geq E^+_{x_t} - \frac{E^+_{x_t}}{T(1-x_t)}$ by lemma 5 with

$$E^+_{x_t} - \frac{E^+_{x_t}}{T(1-x_t)} \geq E^D_{x_t} + \frac{1}{N} - \frac{E^+_{x_t}}{T(1-x_t)} \geq E^D_{x_t} + \frac{1}{N} - \frac{E^+_0}{T\frac{1}{2}\tau_0} \geq E^D_{x_t} + \frac{1}{N} - \frac{E_0+1}{T\frac{1}{2}\tau_0} > E^D_{x_t}$$

where I used the fact $E^+_{x_t} \geq E^D_{x_t} + \frac{1}{N}$, that $E^+_{x_t} \leq E^+_0$, that $1 - x_t > \frac{1}{2}\tau_0$, that $E^+_0 = E_0 + \frac{1}{N} \leq E_0 + 1$ and that $\frac{E_0+1}{T\frac{1}{2}\tau_0} < \frac{1}{N}$. Plugging this into equation (17), we get

$$G(x, s, E^+_x) > G(x_t, s, E^D_{x_t}) \geq G^D(x_t, s, E^D_{x_t})$$

which proves the second of the two statements.

Let P^-_t be the statement that is true $E^-_{x_{t'}} \leq E^D_{x_{t'}} - \frac{1}{N}$ for all $t' \leq t$ and that $e^D_x \leq e^-_x$ for all $x < x_t$. By assumption, P_1 holds. Suppose P_t holds. Then if $x_{t+1} \leq 1 - \frac{1}{2}\tau_0$, we have that $e^D_{x_t} \leq e^-_x$ and therefore that P_{t+1} holds. First, note that $G(x, e^-_x, E^-_x) \geq B$ for all $x \in [x_t, x_{t+1}]$, since if $G(x, \frac{E^-_x}{1-x}, E^-_x) < B$, then $\tau^-_0 = 0$, which we ruled out by choosing N sufficiently large. Since $\tau^-_0 > \frac{1}{2}\tau_0$ we can use the bounds between G and G^D just proved, thus we have $B \leq G(x, e^-_x, E^-_x) < G^D(x_t, e^-_{x_{t+1}}, E^-_{x_t})$, which implies that $e^D_{x_t} \leq e^-_x$ and therefore $E^D_{x_{t+1}} \geq E^-_{x_{t+1}} + \frac{1}{N}$. Therefore P_t holds for all t such that $x_t \leq 1 - \frac{1}{2}\tau_0$.

Now note that because of how we picked N and T, we have that $0 < \tau_0 - \tau^-_0 < \frac{1}{2}\tau_0$ and that $\frac{1}{T} < \frac{1}{2}\tau_0$. Therefore we have that $\frac{1}{2}\tau_0 < \tau^-_0 < \tau_0$, thus $1 - \tau_0 < 1 - \tau^-_0 < 1 - \frac{1}{2}\tau_0$ and finally $(1-\frac{1}{2}\tau_0) - (1-\tau^-_0) = (1-\frac{1}{2}\tau_0) - (1-\tau_0) + (1-\tau_0) - (1-\tau^-_0) = \frac{1}{2}\tau_0 + \tau^-_0 - \tau_0 > \frac{1}{2}\tau_0 > \frac{1}{T}$. Let t^* be the largest t such that $x_t \leq 1 - \tau^-_0$. Then $x_{t+1} = x_t + \frac{1}{T} \leq 1 - \tau^-_0 + \frac{1}{T} < 1 - \frac{1}{2}\tau_0$, and therefore P_{t^*} holds and P_{t^*+1} holds. Therefore $e^D_{x_{t^*}} \leq e^-_x$ for all $x \in [x_{t^*}, x_{t^*+1}]$, which means that $e_{x^*_t} = 0$ since $e^-_{x_{t^*+1}} = 0$. This proves that $\tau^D_0 \geq \tau^-_0$.

A similar argument definining P^+_t as the statement that holds that $E^+_{x_{t'}} \geq E^D_{x_{t'}} + \frac{1}{N}$ for all $t' \leq t$ and that $e^D_x \geq e^+_x$ for all $x < x_t$ can be made to show that P^-_t holds for all t with $x_t \leq 1 - \frac{1}{2}\tau_0$ and therefore $\tau^D_0 \leq \tau^+_0$. Moreover, since the effort is always in between, and the task is never completed, it is clear that the disutility u^D_0 lies in between u^- and u^+.

This proves the first half of the proposition, when $B_0 \in (B_L(E_0), B_C(E_0))$.

Now let us look at the case when $B_0 \in (B_C(E_0), B_H(E_0))$. The proof is identical if we replace τ_0 by τ_F, until we get to the final step regarding property P^-_t and property P^+_t.

Let P^+_t be the statement that $E^+_{x_{t'}} \geq E^D_{x_{t'}} + \frac{1}{N}$ for all $t' \leq t$ and that $e^D_x \geq e^-_x$ for all $x < x_t$. By assumption P_1 holds. Suppose P_t holds. Then if $x_t \leq 1 - \tau^+_F$, we have that $x_{t+1} < 1 - \frac{1}{2}\tau_F$ by our choice of N and T (this uses a similar argument as in the τ_0 case). Therefore we know that $G(x, s, E^+_x) \geq G^D(x_t, s, E^D_{x_t})$ $\forall x \in [x_t, x_{t+1}]$. We have that $B = G(x, e^+_x, E_x)$ (since we know that the person works partially until time $1 - \tau^+_F$) and thus $B > G^D(x_t, e^+_x, E^D_{x_t})$. Therefore we have that $e^D_{x_t} \geq e^+_x$ or $e^D_{x_t} = \frac{E^D_{x_t}}{1-x_t}$, i.e. work is done efficiently from then onwards. Thus, either we have that P_{t+1} holds or that $\tau^D_F \geq \tau^+_F$. Suppose that t^* is the largest t for which P_t holds. Then if $x_{t^*} \leq 1 - \tau^+_F$ we know that either P_{t^*+1} holds or that $\tau^D_F \geq \tau^+_F$. Thus

$\tau_F^D \geq \tau_F^+$ since by the definition of t^*, P_{t^*} cannot hold. If we have that $x_{t^*} > 1 - \tau_F^+$, then we know that P_{t^*} holds, and thus $e^D_{x_{t^*}} \geq e^-_x$ for all $x < x_{t^*}$, i.e. for all $x \leq 1 - \tau_F^+$ as well. But since $e_{\tau_F^+} = e_x = \frac{E^+_{\tau_F^+}}{\tau_F^+}$, this means that $e^D_{\tau_F^+} \geq \frac{E^+_{\tau_F^+}}{\tau_F^+} > \frac{E^D_{\tau_F^+}}{\tau_F^+}$, which implies that the person works more than is efficient, which is a contradition. Thus $x_{t^*} \leq 1 - \tau_F^+$ and $\tau_F^D \geq \tau_F^+$, with P_t holding for all $x_t \leq 1 - \tau_F^D$.

A similar argument establishes that $\tau_F^D \leq \tau_F^-$ and that $e^-_x \geq e^D_x$ for all $x \leq 1 - \tau_F^-$. Then we can apply lemma 8 which shows that in this case $u_F^+ \leq u_F^D \leq u_F^-$. This completes the proof. □

A.2.2 Proofs of Results on Multi-Day Multi-Tasking

Proof of proposition 6.

Proof. The first part of the proof relies on proposition 3. By assumption, all days are the same, so the person plans to work the same amounts on the short-term and the long-term task. Therefore on the first day, she switches exactly at the time where she would if $T = 1$. The same is true on the second day, because as long as she has worked less than she worked on day 1, she is more optimistic, and therefore plans to work more, and thus switch later than on day 1. Once she has worked as much as on day 1, she solves exactly the same maximization problem as she did at that time on day 1 – the fact that she worked less on the short-term task afterwards does not change her benefits today or on any other future day. Therefore she switches at the same time. The same is true for all future days.

The situation is more interesting when the person works each day first on the short-term task – the intuition is that on the second day, she switches earlier, because she did not work as much at the end of day 1 on the long-term task, and therefore she does, in fact, face different benefits on day 2 than she thought.

First, let us prove that on the first day, she works just the same on the short-term task as she would in the single-day case, but that she works less on the long-term task.

Note that the on the first day, they switch at the same time as when $T = 1$. The reason is as follows: call this time S_1^*. Then at S_1^* they solve the same maximization problem that an unbiased person would solve who actually had the disutility $\tilde{D}(.|S^*)$. But this person would switch at the same time when $T = 1$ or when $T > 1$, since it is optimal to work the same amount each day, and what is optimal for 1 day is optimal for T days. Therefore, on day 1 she switches to the long-term task at the same time as she would in the 1-day case.

Let us now show that she stops working on the long-term task sooner when $T > 1$ than when $T = 1$.

Step 1: The higher S is on day 1 after switching (the longer she has worked), the less she plans to work in total on all future days.

Let $\hat{E}_1(S)$ and $\hat{E}_2(S)$ be the amount the person plans to work on day 2 (and thus on all following days) when she has worked S hours so far on day 1. Notice that the amount she plans to work on the first day on the long-term task is given by taking the total amount of work she plans to do (which is the same as on all future days) minus how much she already has worked on the short-term task, that is $\hat{E}_1(S) + \hat{E}_2(S) - \tilde{E}_1^*$. Thus she thinks that in total she will work $(T-1)\hat{E}_2(S) + (\hat{E}_1(S) + \hat{E}_2(S) - \tilde{E}_1^*) = T\hat{E}_2(S) + (\hat{E}_1(S) - \tilde{E}_1^*)$ on the long-term task.

The first order conditions are given by

$$\tilde{D}'(\hat{E}_1(S) + \hat{E}_2(S)|S) = B_1'(\hat{E}_1(S)) = B_2'\left(\frac{(T-1)\hat{E}_2(S) + \hat{E}_2(S) + \hat{E}_1(S) - \tilde{E}_1^*}{T}\right) \tag{18}$$

Notice that if $\hat{E}_1(S) + \hat{E}_2(S)$ weakly increases as S strictly increases, then $\tilde{D}'(\hat{E}_1(S) + \hat{E}_2(S)|S)$ strictly increases. The first equality implies that $\hat{E}_1(S)$ must strictly decrease (since B_1 is strictly concave), and the second equalilty implies that $\hat{E}_2(S)$ must strictly decrease (since the remainder of the argument increases). But that contradicts the statement that $\hat{E}_1(S) + \hat{E}_2(S)$ weakly increases. Therefore, it must strictly decrease as S strictly increases.

Step 2: $\hat{E}_1(S)$ strictly decreases as S strictly increases.

Suppose not. We know from step 1 that $\hat{E}_1(S) + \hat{E}_2(S)$ strictly decreases. If $\hat{E}_1(S)$ does not strictly decrease, then $\hat{E}_2(S)$ must strictly decreases. Moreover, by equation (18), if $\hat{E}_1(S)$ stays the same or increases, then so must $(T-1)\hat{E}_2(S) + (\hat{E}_1(S) + \hat{E}_2(S))$. But this is impossible, since it is the sum of two terms, both of which strictly decrease. Therefore, $\hat{E}_1(S)$ must strictly decrease.

Step 3: Find the first order conditions for $T = 1$ and $T > 1$:

Let $\tilde{E}_{2,i}^*$ be the amount the person actually works on the second task on day 1 for $i \in \{L, S\}$ for "long-term" and "short-term". Let $\hat{E}_i^* := \hat{E}_i(\tilde{E}_2^*$, the amount the person plans (at the end of the first day) to work on task $i \in \{1, 2\}$ on all future days. The first order conditions are then as follows:

$$\begin{aligned}
&T = 1: \;\; D'(\tilde{E}_1^* + \tilde{E}_{2,S}^*) = B_2'(\tilde{E}_{2,S}^*) \\
&T > 1: \;\; D'(\tilde{E}_1^* + \tilde{E}_{2,L}^*) = B_2'\left(\frac{(T-1)\hat{E}_2^* + \tilde{E}_2^*}{T}\right) = D'(\hat{E}_1^* + \hat{E}_2^*)
\end{aligned}$$

Step 4: $\tilde{E}_2^*(T > 1)$ is lower than $\tilde{E}_2^*(T = 1)$.

From the first order condition, we have that $\tilde{E}_1^* + \tilde{E}_{2,L}^* = \hat{E}_1^* + \hat{E}_2^*$. But since $\tilde{E}_1^* = \hat{E}_1(S)$ at $S = \tilde{E}_1^*$, and $\hat{E}_1^* = \hat{E}_1(S)$ at $S = \tilde{E}_1^* + \tilde{E}_{2,L}^*$, we know from step 2 that $\tilde{E}_1^* > \hat{E}_1^*$ and therefore $\tilde{E}_{2,L}^* < \hat{E}_2^*$. This in turn means that $\frac{(T-1)\hat{E}_2^* + \tilde{E}_{2,L}^*}{T} = \tilde{E}_{2,L}^* + a$ for some $a > 0$.

We can now rewrite the first order condition for $T > 1$ as $D'(\tilde{E}_1^* + \tilde{E}_{2,L}^*) = B_2'(\tilde{E}_{2,L} + a)$, whereas $\tilde{E}_{2,S}$ solves the first order condition $D'(\tilde{E}_1^* + \tilde{E}_{2,S}^*) = B_2'(\tilde{E}_{2,S})$. Thus $\tilde{E}_{2,L}^* < \tilde{E}_{2,S}^*$: if it is the same or larger, then the LHS of the first order condition is the same or larger, and the RHS strictly smaller, since $a > 0$, and thus it cannot hold. Thus, on day 1, the person works less on the long-term task with $T > 1$ than with $T = 1$.

Part II: The amount spent working on the long-term task is lower on average with $T > 1$ than with $T = 1$.

Let $E_{i,L}(t)$ the work done on day t on task $i \in \{L, S\}$ task when $T > 1$. Let $E_{i,F}(t)$ be the amount the person would choose that day if she was told that she would have to work *exactly* the same on all future days too (F stands for 'fixed'). Then she works the same as she does in the single-day choice where her daily benefits are her average benefits – taking the effort she has worked so far as a given.

Then I claim that $E_{2,S}(t+1) \cdot (T-t) + E_{2,L}(t) < E_{2,S}(t)(T-t+1)$, so that $E_{2,S}(t+1)(T-t) + E_{2,L}(t) + E_{2,L}(t-1) < E_{2,S}(t)(T-t+1) + E_{2,L}(t-1) < E_{2,S}(t-1)(T-t+2)$, and so on until $E_{2,S}(t+1)(T-t) + \sum_{i=1}^{t} E_{2,L}(i) < T \cdot E_{2,S}(1)$. Thus for $t = T-1$, we have $E_{2,S}(T) + \sum_{i=1}^{T-1} E_{2,L}(i) = \sum_{i=1}^{T} E_{2,L}(i) < T \cdot E_{S,1}$, since $E_{2,S}(T) = E_{2,L}(T)$ as there is only one day left. This proves the claim, since the LHS is equal to the work done when $T > 1$ and the RHS is equal to T times the work done when $T = 1$ on day 1. We now have to prove the claim.

Claim: $E_{2,S}(t+1) \cdot (T-t) + E_{2,L}(t) < E_{2,S}(t)(T-t+1)$ for all $t < T$.

In fact, let us prove a stronger claim. Let $E_i(x)$ be the work done on task $i \in \{L, S\}$ on day 2 if on day 1 the person has worked for x hours on the long-term task, and if she again will be required to work exactly the same across all future days as she does on day 2. Then I will show that for all $x < E_{2,S}(1)$ we have

$$E_2(x)(T-1) + x < E_{2,S}(1) \cdot T$$

This proves the claim when we relabel day t as day 1 and focus only on the continuation problem, and set $x = E_{2,L}(t)$, since we know that $E_{2,L}(t) < E_{2,S}(t)$, as we showed in part I that the this holds for the first day – and day t is the first day of the remaining task. Intuitively, this says that if the person works on day 1 knowing that she has to work the same on all future days, she works more than if she had only completed x on day 1 (for whatever reasons), and knew that she had to work the same on all days after day 2.

This statement is true for an unbiased person. Suppose it wasn't true. We have that $D'(E_{2,S}(1) + E_{1,S}(1)) = B_1'(E_{1,S}(1)) = B_2'(E_{2,S}(1))$ and $D'(E_2(x) + E_1(x)) = B_1'(E_1(x)) = B_2'(\frac{(T-1)E_2(x)+x}{T})$. If $E_2(x) \leq E_{2,S}(1)$ then we are done. So suppose $E_2(x) > E_{2,S}(1) > x$ (which is true, but that is irrelevant). Then we can write $B_2'(\frac{(T-1)E_2(x)+x}{T}) = B_2'(E_2(x) - a)$ for some $a > 0$. Note that $E_2(x) + E_1(x)$ is strictly increasing in a: suppose not, so that it stays or decreases. Then $E_1(x)$ stays or increases. Therefore $E_2(x) - a$ stays or increases. Therefore $E_2(x)$ must strictly increase, which is a contradiction, since $E_1(x) + E_2(x)$ – assumed to weakly decrease – is then the sum of two terms, both of which increase, and one strictly increasing. Therefore, we have that $E_1(x) + E_2(x)$ strictly increases, and hence that $B'(E_2(x) - a)$ strictly increases, and thus that $E_2(x) - a$ is strictly smaller the larger a is. Since $E_{2,S}(1)$ solves the same equations when we set $a = 0$, we have that $\frac{(T-1)E_2(x)+x}{T} < E_{2,S}(1)$ which proves the claim.

The statement is also true for a projection-biased person. Let us write down the first order conditions satisfied by the different amounts of effort exerted. For $E_{1,S}$, $E_{2,S}$ and $E_{2|1,S}$ we have the following:

$$\tilde{D}'(E_{1,S} + E_{2|1,S}|E_{1,S}) = B_1^{'}(E_{1,S}) = B_2^{'}(E_{2,S}) \tag{19}$$

$$D'(E_{1,S} + E_{2,S}) = B_2^{'}(E_{2,S}) \tag{20}$$

For $E_1(x)$, $E_2(x)$ and $E_{2|1}(x)$ we have:

$$\tilde{D}'(E_1(x) + E_{2|1}(x)|E_1(x)) = B_1^{'}(E_1(x)) = B_2^{'}\left(\frac{(T-1)E_{2|1}(x) + x}{T}\right)$$

$$D'(E_1(x) + E_2(x)) = B_2'\left(\frac{(T-1)E_2(x) + x}{T}\right)$$

Case 1: $E_1(x) \geq E_{1,S}$. Then I claim that $E_1(x) + E_2(x) > E_{1,S} + E_{2,S}$. If so, then we have we have $D'(E_1(x) + E_2(x)) > D'(E_{1,S} + E_{2,S}) \implies B_2'(\frac{(T-1)E_2(x)+x}{T}) > B_2'(E_{2,S}) \implies (T-1)E_2(x)+x < TE_{2,S}$ as we want. To prove the claim, suppose that $E_1(x)+E_2(x) \leq E_{1,S}+E_{2,S}$ when $E_1(x) \geq E_{1,S}$. Then $E_2(x) \leq E_{2,S}$ and thus $(T-1)E_{2,S} + x < TE_{2,S}$ since $x < E_{2,S}$.

Case 2: $E_1(x) < E_{1,S}$. Let us look at the plans the person would ideally make if she had to choose at $S = E_{1,S}$, that is at the moment she decides to switch when her future behavior mirrors behavior on day 1. Then she solves the first order condition in equation (19) when choosing on day 1. If she could change her behavior earlier on day 2 (after working for x hours on day 1), on day 2 she would choose

$$\tilde{D}'(E_1(x|S) + E_{2|1}(x|S)|S = E_{1,S})) = B_1^{'}(E_1(x|S)) = B_2^{'}\left(\frac{(T-1)E_{2|1}(x|S) + x}{T}\right)$$

Since $E_{1,S} > E_1(x)$, she would choose to work less in total, and therefore to switch earlier than she actually did switch on day 2. Thus $E_1(x) > E_1(x|S)$. Now we can note that these plans, $E_1(x|S)$ and $E_{2|1}(x|S)$ in one case, and $E_{1,S}$ and $E_{2|1,S}$ in the other are both as if done by an unbiased person in two different situations. We have shown that such a person works less in total on the long-term task, that is $(T-1)E_{2|1}(x) + x < TE_{2|1,S}$, so we have the following

$$
\begin{aligned}
&(T-1)E_{2|1}(x|S) + x < TE_{2|1,S} && (21)\\
\implies & B_2^{'}((T-1)E_{2|1}(x|S)+x) > B_2^{'}(E_{2|1,S}) && (22)\\
\implies & \tilde{D}'(E_1(x|S)+E_{2|1}(x|S)|S=E_{1,S}) > \tilde{D}'(E_{1,S}+E_{2|1,S}|S=E_{1,S}) && (23)\\
\implies & (1-\alpha)(D'(E_1(x|S)+E_{2|1}(x|S)) - D'(E_{1,S}+E_{2|1,S}))+ && (24)\\
& \alpha D'(E_{1,S})(E_1(x|S)+E_{2|1}(x|S) - E_{1,S} - E_{2|1,s}) > 0 && (25)\\
\implies & E_1(x|S) + E_{2|1}(S) > E_{1,S} + E_{2|1,S} && (26)\\
\implies & -E_1(x|S) - E_{2|1}(S) < -E_{1,S} - E_{2|1,S} && (27)
\end{aligned}
$$

where the last line holds since both terms are of the same sign. We want to show that $E_1(x)+E_2(x) > E_{1,S}+E_{2,S}$. Suppose that this is not true, so that $E_1(x)+E_2(x) \leq E_{1,S}+E_{2,S}$. Adding inequality (27) to this we obtain:

$$
\begin{aligned}
E_1(x|S) + E_{2|1}(x|S) - E_1(x) - E_2(x) &> E_{2|1,S} - E_{2,S}\\
\implies E_2(x) - E_{2|1}(x|S) &< \underbrace{E_{2,S} - E_{2|1,S} + E_1(x|S) - E_1(x)}_{<0}
\end{aligned}
$$

where the final inequality holds because we know that $E_1(x|S) < E_1(x)$ and $E_{2,S} < E_{2|1,S}$ – since they are choices over exactly the same choice variable made when the person is more tired, and thus plans to work less. Finally, we have that

$$
(T-1)E_2(x) + x = (T-1)(E_2(x) - E_{2|1,S}(x|S)) + (T-1)E_{2|1,S} + x < (T-1)E_{2|1,S} + x < TE_{2,S}
$$

and we are done with part II.

Part III: Total effort exerted on both tasks is strictly lower on day 1 than on day 2, and thus is strictly increasing over time.

The second half follows directly from the first, since we can relabel day t as day 1 of the remaining task and we are done.

Let us show that total effort increases. At the end of day 1, the person plans to switch sooner on day 2 than she will switch (similar proof to all such statements so far). On day 2 once she has worked as much as she worked on day 1, she therefore has worked less on the long-term task than she had planned the day before, so that the marginal benefit from continuing are strictly larger, so that she strictly continues working.

Part IV: Effort on task 1 is strictly lower on day 2 than on day 1, and is therefore strictly decreasing over time.

The second part is easy. If $E_{1,S}(1) > E_{1,S}(2)$, then $E_{1,S}(t) > E_{1,S}(t+1)$ since day t is simply the first day of the task starting that day, which satisfies all the same conditions.

Let us prove that on day 2, the person switches earlier. If she had worked as much on the long-term task as she planned when she switched, then on the second day she would switch at exactly the same time (all first order conditions hold as they did on day 1). But she worked less on the long-term task. Therefore the marginal benefits from working on the long-term task are higher, and an argument as in step 4 of part I (with $a > 0$) will go through, showing that she works indeed less, in order to have more time working on the long-term task.

Part V: Effort on task 2 is strictly larger on day 2 than on day 1, and is therefore strictly increasing over time.

As for part IV, the second half of the claim follows by relabeling day t as the first day of the remaining days. The first half follows immediately from parts III and IV, since the person works strictly less on the first task, but strictly more on both together, so she must work strictly more on the second task.

This completes the proof. $\square$

A.3 Proofs for Section 5

Lemma 12. *Let $U_a(\boldsymbol{e}) = X(\boldsymbol{e}) + a \cdot Y(\boldsymbol{e})$, with X and Y continuous (real-valued) functions of the vector $\boldsymbol{e}$, and $a \in \mathbb{R}$. Then, for all a, $\arg\max_{e \in \mathcal{E}} U_a(\boldsymbol{e})$ is not empty when $\mathcal{E}$ is a compact set. Let $\boldsymbol{e}(a) \in \arg\max_{e \in \mathcal{E}} U_a(\boldsymbol{e})$. If $a_H > a_L \geq 0$, then $X(\boldsymbol{e}(a_H)) \leq X(\boldsymbol{e}(a_L))$ and $Y(\boldsymbol{e}(a_H)) \geq Y(\boldsymbol{e}(a_L))$.*

Proof. By compactness of $\mathcal{E}$ and continuity of X and Y, the maximimum is achieved in $\mathcal{E}$ we can find $\boldsymbol{e}(a)$ as stated. Denote $X(\boldsymbol{e}(a_i))$ by X_i, $Y(\boldsymbol{e}(a_i))$ by Y_i for $i \in \{H, L\}$. Since $\boldsymbol{e}(a_i)$ maximizes U_{a_i}, we have that

$$X_H + a_H \cdot Y_H \geq X_L + a_H \cdot Y_L \tag{28}$$
$$X_L + a_L \cdot Y_L \geq X_H + a_L \cdot Y_H \tag{29}$$

Adding equations (28) and (29), we find that

$$\begin{aligned} a_H \cdot Y_H + a_L \cdot Y_L \geq a_H \cdot Y_L + a_L \cdot Y_H &\iff (a_H - a_L)Y_H \geq (a_H - a_L)Y_L \\ &\iff Y_H \geq Y_L \end{aligned}$$

since $a_H - a_L > 0$. If $a_H > a_L \geq 0$, then by adding a_L times equation (28) and a_H times equation (29), we find that

$$\begin{aligned} a_L \cdot X_H + a_H \cdot X_L \geq a_L \cdot X_L + a_H \cdot X_H &\iff (a_H - a_L) \cdot X_L \geq (a_H - a_L) \cdot X_H \\ &\iff X_L \geq X_H \end{aligned}$$

which completes the proof. $\square$

Proof of 7:

Proof. The person at every moment maximizes her perceived utility, which is given by

$$B(\tilde{\boldsymbol{e}}) - \sum_{t=1}^{T} \tilde{D}(\tilde{e}_t) = B(\tilde{\boldsymbol{e}}) - (1-\alpha)\sum_{t=1}^{T} D(\tilde{e}_t) - \alpha D'(s)\sum_{t=1}^{T} \tilde{e}_t$$

Since $\sum_{t=1}^{T} \tilde{e}_t = E$ by assumption – the person needs to complete a given amount of work – at every moment she maximizes $B(\tilde{\boldsymbol{e}}) - (1-\alpha)\sum_{t=1}^{T} D(\tilde{e}_t)$ and the claim follows. $\square$

Proof of 8:

Proof. The actual first order condition is

$$D'(e_1) - q_t \mathbb{E}(D'(e_t)) = 0$$

Replacing D' by $\tilde{D}'$ and expanding this yields the first part (similar to the derivation for the examples). The actual effort $\tilde{e}_1^*$ is determined when the person stops working – when $s = \tilde{e}_1^*$. Substituting this for s in

$$D'(\tilde{e}_1(s)) - q_t \mathbb{E}(D'(\tilde{e}_t(s))) = -\frac{\alpha}{1-\alpha} D'(s)(1-q_t)$$

yields

$$D'(\tilde{e}_1) = \frac{1-\alpha}{1-\alpha q_t} \mathbb{E}(D'(\tilde{e}_{t|1}^*))$$

Finally, when $q_2 < 1$, then $(1-q_t) > 0$ and so $-\frac{\alpha}{1-\alpha} D'(s)(1-q_2)$ is decreasing in s, which requires $\tilde{e}_1(s)$ to decrease with s. $\square$

Proof of 9.

Proof. After period t, the person receives one of $k_t < \infty$ signals based on which she updates her beliefs over E. Every sequence of signals thus determines a single final outcome, which I denote by $s \in S$. Let $n(s,t) \in \{1, ..., N\}$ for some N index all the possible effort levels the person can choose, where $n(s,t)$ is the effort level at time t when the final outcome will be state s. Notice that $n(s,t) = n(s',t)$ is possible for $s \neq s'$ when at time t, the person puts strictly positive probability on both states s and s' happening. Let e_t denote the random variable representing effort chosen at time t. Then the optimization problem is given by:

$$\min_{e_n, n \in \{1,\ldots,N\}} \mathbb{E}\left(\sum_{t=1}^{T} \delta_t D(e_t)\right)$$
$$\text{s.t. } \sum e_{n(s,t)} q_t = E_s, \ \forall s \in S$$

Let us first prove that $\tilde{e}_1^* \leq e_1^*$, that the projection-biased person works less on day 1 than the unbiased person. Suppose not. Then $\tilde{e}_1^* > e_1^*$. First, this means that the projection-biased person is not at a corner solution. Suppose that $e_1 > 0$, then the first order condition is

$$\frac{\delta_1 D'(e_1)}{q_1} \leq \frac{\delta_t \mathbb{E}_1(D'(e_t))}{q_t}$$

Now the perceived first order condition when the person has worked for $H \geq 0$ hours is:

$$\begin{aligned}
&\frac{\delta_1 \tilde{D}'(e_1|H)}{q_1} \leq \frac{\delta_t \mathbb{E}_1(\tilde{D}'(e_t|H))}{q_t} \\
\iff &\frac{\delta_1}{q_1} D'(\tilde{e}_1^*) - \frac{\delta_t}{q_t}\mathbb{E}(D'(\tilde{e}_{t|1}^*)) \leq \left(\frac{\delta_1}{q_1} - \frac{\delta_t}{q_t}\right) D'(H) \left(-\frac{\alpha}{1-\alpha}\right) \leq 0
\end{aligned}$$

which implies that $\mathbb{E}_1(D'(\tilde{e}_{t|1}^*))\frac{\delta_t}{q_t} \geq D'(\tilde{e}_1^*)\frac{\delta_1}{q_1} > D'(e_1^*)\frac{\delta_1}{q_1} \geq \mathbb{E}(D'(e_t^*))\frac{\delta_t}{q_t}$. In particular, this means that there is some $n_2 = n(s,2)$ s.t. $D'(\tilde{e}_{n_2}) > D'(e_{n_2})$. If not, then we can sum over $\{n : n(s,2) \forall s \in S\}$, over all possible effort levels in period 2 and obtain that $\mathbb{E}_1(D'(\tilde{e}_{2|1}^*) \leq \mathbb{E}_1(D'(e_2^*))$, which is a contradition. We can now repeat the same argument for day 2 and show that, starting with the subproblem on day 2 with the state n_2, that there exist possible continuations n_3, ..., n_t s.t. $n_i = n(s,i)$ for a given $s \in S$, such that $\tilde{e}_{n_i}^* > e_{n_i}^*$. But this is impossible, since it means that the person works strictly more than she must, which cannot minimize the disutility. Therefore $\tilde{e}_1^* \leq e_1^*$ (and in fact the inequality is strict if $e_1^* > 0$).

It is therefore clear that the person plans to have worked less by day t in total than the unbiased person. But if at time t, she is behind, and at time $t+1$ has pulled ahead, she must have worked strictly more, and therefore (by a similar argument as above) work strictly more in some other periods as well, in which case she works strictly more again. This is impossible.

Let us now show that the projection-biased person works less on day 2 than she thought she would when there is no uncertainty in E, so that $|S| = 1$. Since $\frac{\delta_t}{q_t}$ is strictly decreasing, she should work strictly more on day 2. Denote $\frac{\delta_t}{q_t}$ simply by p_t, so that p_t is decreasing (at some point strictly) in t. I assume that the person exerts some effort on day 1 (and thus on all days). Then the FOCs are

$$p_1 \tilde{D}'(\tilde{e}_1|s) = p_t \tilde{D}'(e_t|s) \iff \qquad q_1 D'(\tilde{e}_1(s)) - q_t D'(\tilde{e}_t(s)) = -\frac{\alpha}{1-\alpha}(q_1 - q_t) D'(s)$$

Thus $\tilde{e}_1(s)$ is strictly increasing in s. (If it was increasing, then $\tilde{e}_t(s)$ strictly increases, which means that total effort strictly increases, a contradiction since it doesn't minimize perceived disutility.) Thus, on day 2, the person plans to work at least as much as she planned to do at the end of day 1, until she has worked more than she worked on day 1, at which point she plans to work strictly less.

We can again show that the person also plans to have less work done by day t than she did at the end of day 1. Here is why. Let $\bar{t}$ be the first day on which the person plans to work more at the end of day 2 than at the end of day 1. (This must exist in order to complete all the work.) Then she works strictly more on all subsequent days, since the FOC for day $t' > \bar{t}$ is

$$q_{t'} D'(\tilde{e}^*_{t'|i}) - q_{\bar{t}} D'(\tilde{e}^*_{\bar{t}|i}) = -\frac{\alpha}{1-\alpha}(q_{t'} - q_{\bar{t}}) D'(\tilde{e}^*_i)$$

Thus

$$\begin{aligned}
D'(\tilde{e}^*_{t'|2}) &= \frac{q_{\bar{t}}}{q_{t'}} D'(\tilde{e}^*_{\bar{t}|2}) - \frac{\alpha}{1-\alpha}(q_{t'} - q_{\bar{t}}) D'(\tilde{e}^*_2) \\
&\geq \frac{q_{\bar{t}}}{q_{t'}} D'(\tilde{e}^*_{\bar{t}|1}) - \frac{\alpha}{1-\alpha}(q_{t'} - q_{\bar{t}}) D'(\tilde{e}^*_2) \\
&> \frac{q_{\bar{t}}}{q_{t'}} D'(\tilde{e}^*_{\bar{t}|1}) - \frac{\alpha}{1-\alpha}(q_{t'} - q_{\bar{t}}) D'(\tilde{e}^*_1) \\
&= D'(\tilde{e}^*_{t'|1})
\end{aligned}$$

so that the person plans to work more on all days after $\bar{t}$, and therefore must have done less work until that point, and at the start of every day until the final day. □

A.4 Proofs for Section 6

Proof for proposition 10.

Proof. The claims regarding E_H follow the same as in the proof of proposition 2, using the fact that there is an S such that $D'(S) > D'(0)$.

Thus I focus on the statements around E_L.

Claim 1: If $\bar{e} > 0$, then there is a unique $E_0 > 0$ such that $D'(E_0) = D'(0)$ and $D'(E) < D'(0)$ for $E \in (0, E_0)$ and $D'(E) > D'(0)$ for $E > E_0$.

Proof of claim 1: $D''(E) < 0 \ \forall E \in [0, \bar{e})$, and $D''(\bar{e}) = 0$. Thus $D'(E) - D'(0) = \int_0^E D''(e)de < 0$ for all $E \in [0, \bar{e}]$. By assumption, $D'(E) \to \bar{D} > D'(0)$. Therefore, by the intermediate value theorem, there exists some $E_0 \in (\bar{e}, \infty)$ such that $D'(E_0) = D'(0)$. Similarly, $D''(E) > 0$ for $E > \bar{e}$, therefore $D'(E)$ is strictly increasing on $E > \bar{e}$, so there cannot be two such E_0. Since $D'(\bar{e}) < D'(0)$ and since $D'(E)$ is increasing for $E > \bar{e}$, the claim follows.

Claim 2: If $\bar{e} > 0$, there is a unique strictly positive number (call it E_L) such that $D(E)/E = D'(0)$. (Note that if $\bar{e} = 0$, then $D(E)/E > D'(0)$ for all $E > 0$, and thus there is no such E.)

By claim 1, we know that there is a unique $E_0 > 0$ such that $D'(E_0) = D'(0)$, and $D'(E) > D'(0)$ when $E > E_0$ and $D'(E) < D'(0)$ when $0 < E < E_0$. Therefore $D(E)/E < D'(0)$

for $E < E_0$. Moreover, since $D'(E_0+\varepsilon) > D'(E_0) = D'(0)$, we have that $D(E)/E_0$ must eventually become arbitrarily close to $D'(E_0+\varepsilon)$ and thus exceed $D'(0)$. By the intermediate value theorem there must therefore be a point at which $D(E)/E$ equals $D'(0)$. Denote the first time this happens by E_L. Note that $E_L > E_0$. Then it is easy to see that there is only one such E_L. Suppose there was a second, $E_L^{'} > E_L > E_0$. Then

$$\begin{aligned} D'(0) = \frac{D(E_L^{'})}{E_L^{'}} &= \frac{1}{E_L^{'}} \int_0^{E_L^{'}} D'(e)de \\ &= \frac{1}{E_L^{'}} \left(\int_0^{E_L} D'(e)de + \int_{E_L}^{E_L^{'}} D'(e)de \right) \\ &= \frac{1}{E_L^{'}} \left(D'(0) \cdot E_L + \int_{E_L}^{E_L^{'}} D'(e)de \right) \\ &> \frac{1}{E_L^{'}} \left(D'(0) \cdot E_L + D'(0)(E_L^{'} - E_L)de \right) \\ &= D'(0) \end{aligned}$$

which is a contradiction. The inequality comes from the fact that $D'(e) > D'(0)$ for $e > E_0$, which holds when $e > E_L$ since $E_L > E_0$. Note that the proof also shows that $D(E)/E < D'(0)$ when $E < E_L$ and $D(E)/E > D'(0)$ when $E > E_H$.

Suppose that $e < E_L$. We want to show that people only start tasks that are worth it, that if they start a task, they finish it, and that they fail to start some worthwhile tasks.

People start a task if the initial perceived disutility of the task is lower than the benefits, that is if

$$B \geq \tilde{D}_0(E) = (1-\alpha)D(E) + \alpha D'(0) \cdot E = D(E) + \alpha(D'(0) - D(E)/E)$$

Since $D'(0) > D(E)/E$ when $E < E_L$, this means that the task needs to be strictly worth it for projection-biased people to start it, and if $B \in (D(E), D(E) + \alpha(D'(0) - D(E)/E))$, then people don't start the task.

We only need to show (for the first part) that people finish all worthwhile tasks that they start. Since this result should hold for all E, not just for $E < E_L$, that's what I shall prove.

People stop a task if the remaining perceived disutility at any point, $\tilde{D}_s(E) - \tilde{D}_s(s)$, is larger than the benefits.

$$\begin{aligned} B &\leq \tilde{D}_s(E) - \tilde{D}_s(s) \\ &= (1-\alpha)(D(E) - D(s)) + \alpha D'(s)(E-s) \end{aligned}$$

For tasks that are worthwhile, $D(E) \leq B$, thus $D(E) - D(s) \leq B$. If $D'(s) \leq D'(0)$ then

$(1-\alpha)(D(E)-D(s))+\alpha D'(s)(E-s)<(1-\alpha)D(E)+\alpha D'(0)E\leq B$, since the task was started at time 0. This means that people do not stop before reaching $\bar{e}$, as $D'(e)<D'(0)$ for $e\in(0,\bar{e}]$.

For $e>\bar{e}$, we have that D' is strictly increasing, which implies that $D'(s)(E-s)<D(E)-D(s)<D(E)$, and therefore $(1-\alpha)(D(E)-D(s))+\alpha D'(s)(E-s)<(1-\alpha)D(E)+\alpha D(E)=D(E)<B$, so people do not stop. $\square$