Goal Bracketing and Self-Control

Alice Hsiaw, , International Business School, Brandeis University

Goal Bracketing and Self-Control[*]

Alice Hsiaw[†]
Brandeis University

August 2015

Abstract

This paper studies the role of goal bracketing to attenuate time inconsistency. When setting non-binding goals in multi-stage project, an agent must also decide how and when to evaluate himself against such goals. In particular, he can bracket broadly by setting an aggregate goal for the entire project, or narrowly by setting incremental goals for individual stages. In the presence of loss aversion and uncertainty over outcomes, this decision involves a trade-off between motivation and comparative disutility due to ex-ante uncertainty. Narrow goal bracketing can be used as an instrument to counteract the self-control problem, while broad goal bracketing can itself generate apparently erroneous behavior such as the sunk cost fallacy. The sequential nature of decision-making introduces a differential reaction to outcome uncertainty based on its timing, which determines the optimal bracketing choice.

1 Introduction

When contemplating a long-term project, individuals often realize that they may lack the willpower to achieve a desired outcome, despite the foreknowledge that persistence is beneficial. One commonly suggested method of improving perseverance is to set goals to motivate oneself. In addition to choosing goal levels for achievement, setting appropriate goals also requires a decision regarding how they should be structured to complete a multi-stage project. One option is to take each stage "one step at a time," setting and evaluating goals at each stage. For example, an entrepreneur building a business can set goals for each potential stage of expansion. A student pursuing an education can set achievement goals for each stage before deciding whether to continue onto the next. Alternatively, one could evaluate "the big picture" by setting a holistic goal that is evaluated only upon the project's completion.

Neoclassical theory assumes that when making many choices, an agent maximizes utility by considering a choice in conjunction with the consequences of other choices. However, there exists extensive evidence that individuals generally do not make decisions on such a global basis. Rather than considering choices broadly, they tend to consider each choice in isolation - that is, they bracket narrowly - in the domains of both consumption (Heath and Soll 1996) and risk (Tversky and Kahneman 1981, Gneezy and Potters 1997, Rabin and Weizsäcker 2009). Similarly, the concept of mental accounting posits that individuals allocate expenses into various categories with separate budget constraints, so that they essentially engage in narrow bracketing within each category (Thaler 1999, Heath and Soll 1996). While this body of work has focused on the decision errors that narrow bracketing can produce and the benefits of broad bracketing, formal study of when and why individuals might want to engage in narrow bracketing deliberately has been limited to qualitative discussion of the practice as a means to counteract a self-control problem, if they anticipate a lack of willpower to implement ex-ante plans.

This paper develops a theoretical model to study the endogenous choice of level and bracketing of outcome-based goals, which are outcome targets that the agent expects to achieve, and makes testable predictions regarding phenomena such as the sunk cost fallacy, based on the timing and magnitude of environmental uncertainty. Since frequent goal evaluation can be expected to increase perseverance, a natural question is why an individual might choose to set broader, aggregate goals

that are evaluated less frequently and more holistically. Intuitively, the agent may dislike frequent evaluation if there is ex-ante uncertainty regarding how he will fare relative to his goal, since there is a greater likelihood that he incurs disutility from falling short at each milestone. Setting broader goals allows the agent to pool the risk from uncertain endeavors, minimizing the expected comparative disutility from evaluation.[1] For example, consider an entrepreneur who is building a business. He may enforce stronger motivation if he sets and evaluates goals for each potential stage of expansion. But given ex-ante uncertainty about the profitability of each stage, he is more likely to disappoint himself at each point, so he might choose to evaluate himself less frequently to avoid this. For example, he could instead evaluate the profitability of the stages collectively once he has completed them.

Formally, I consider the optimal bracketing of self-set goals, which I model as reference points. An agent derives utility from both standard consumption, arising from outcomes, and comparison utility, arising from the comparison of these outcomes to a self-set goal. At each point in time, his goal is based on his ex-ante expectations about the outcome of future behavior, which his next "self" will inherit as a reference point in his utility function. In this way, he provides a degree of internal motivation that is desirable if he exhibits time inconsistency due to present-biased time preferences. In addition, the agent determines ex-ante whether to set incremental goals, evaluating "one step at a time," or to set aggregate goals, evaluating less frequently and more holistically. The agent is sophisticated, implying that his goals must ultimately coincide with the rational expectation of what he will actually do. I assume that the agent is loss averse, which leads him to prefer one-shot rather than gradual resolution of ex-ante uncertainty over outcomes.

The economic setting I study is a sequential stopping problem in continuous time with an infinite horizon, in which there exists an option value of waiting due to uncertainty. In the context of entrepreneurship, a business owner must first establish a storefront before opening a second location or expanding a product line. Just as in a single optimal stopping problem, present-biasedness leads the agent to stop too early in each stage because he undervalues this option relative to his time-consistent counterpart.

---

[1]This intuition has been previously been explored in psychology, where more immediate goals are "proximal" and temporally distant goals are "distal." Kirschenbaum (1985, p. 503) states "Using very specific and proximal plans may lead to many failures to reach short-term subgoals. Failures to achieve such subgoals will occur frequently because people often fail to predict accurately fluctuations in situational demands ..."

How an agent with a self-control problem chooses to bracket his goals is determined by a tradeoff between motivational power and expected disutility from ex-ante uncertainty over outcomes. In addition, the intertemporal separation between stopping stages creates an asymmetry in the agent's reaction to ex-ante uncertainty across stages, because an aggregate goal both delays interim outcome evaluation and provides the ability to react to interim news.

Formally, I solve for the unique stationary Markov equilibrium of the intrapersonal game when the agent sets either incremental or aggregate goals, then determine the conditions under which each form of goal bracketing maximizes the value of the sequential stopping option ex-ante. I find that when an agent sets an aggregate goal, he responds to payoff-irrelevant news, exhibiting behavior such as the sunk cost fallacy and "coasting." Either form of goal bracketing attenuates the self-control problem, but setting an aggregate goal is less effective at improving patience than setting an incremental goal because aggregating outcome uncertainty across stages both decreases and delays the anticipated disutility incurred upon stopping. Thus the incentive to wait longer to compensate for such disutility is weaker. However, aggregating goals enables risk pooling and responses to interim news, so doing so becomes more appealing when initial ex-ante outcome uncertainty about a project is high. In contrast, greater ex-ante outcome uncertainty regarding later project stages primarily weakens the motivation to be patient by delaying aggregate goal evaluation, so setting incremental goals becomes more attractive when ex-ante uncertainty about later project stages is high. Thus, I find that narrow goal bracketing can be used as an instrument to counteract the self-control problem, and that broad goal bracketing can itself generate apparently erroneous behavior such as the sunk cost fallacy. Surprisingly, in contrast to a static model, the sequential nature of decision-making interacts with motivational power and risk pooling to introduce a differential reaction to outcome uncertainty based on its timing. This framework also demonstrates the existence of environments in which loss aversion can be beneficial to the agent from an ex-ante perspective.

The paper proceeds as follows. Section 2 links this paper to related lines of research. Section 3 describes the model. Section 4 describes the general method of equilibrium construction. Section 5 characterizes the stationary Markov equilibrium when the agent sets incremental goals, while Section 6 considers the case of aggregate goals. Section 7 derives the conditions under which each form of goal bracketing maximizes the ex-ante value of the sequential stopping option. Section 8

elaborates on the role of loss aversion. Section 9 summarizes the results and discusses avenues for future research. Proofs are gathered in the Appendix.

2   Related Literature

This paper studies choice bracketing in the context of self-control. It also relates to reference dependence through the concept of self-imposed, non-binding goals.

There is extensive evidence that individuals often do not make decisions on the global basis assumed in neoclassical theory, and tend to consider choices in isolation (i.e., to bracket narrowly) in the domains of consumption (Heath and Soll 1996) and risk (Tversky and Kahneman 1981, Gneezy and Potters 1997, Rabin and Weizsäcker 2009). Narrow bracketing, in conjunction with loss aversion, has been used to explain the equity premium puzzle (Benartzi and Thaler 1995), low stock market participation (Barberis, Huang and Thaler 2006), and individual investors' portfolio choices (Odean 1998). The concept of mental accounting posits that individuals create budgets by earmarking expenses into categories, essentially engaging in narrow bracketing within each category (Thaler 1999, Heath and Soll 1996).

While it is evident that individuals engage in narrow bracketing across a variety of contexts, the question of when and why they would *choose* to do so has been relatively unexplored formally. Consistent with the interpretation of bracketing as a result of cognitive limitations, most empirical work has inferred the degree of bracketing necessary to explain observed phenomena (Benartzi and Thaler 1995, Odean 1998), while existing theoretical work has assumed it as given (Barberis, Huang and Santos 2001, Barberis et al. 2006). In contrast, Read, Loewenstein and Rabin (1999) suggest the motivated use of narrow versus broad bracketing intertemporally as a means of self-control. Similarly, Thaler (1999) discusses how people may deliberately assign "tempting" goods to mental accounts with small budgets. The psychological literature includes studies of intertemporal bracketing in the form of short-term versus long-term goals, referred to as proximal versus distal goals, respectively, for repeated or lengthy tasks.[2] Koch and Nafziger (2014) study an agent's bracketing

---

[2]These studies usually involve the comparison of subjects' performance when proximal versus distal goals are *assigned* by the experimenter, and tend to find a larger response to goals when they are proximal (Bandura and Simon 1977, Locke and Latham 2002). Recognizing that the terms "proximal" and "distal" are inherently relative, Kirschenbaum (1985) considers the circumstances under which proximal versus distal goals are preferable and pinpoints the central factor as a trade-off between motivation and uncertainty.

decision across simultaneous activities, and consider sequential activities with anticipatory gain-loss utility in the absence of time discounting, leading to very different incentives. In contrast, I consider intertemporal bracketing with endogenous timing of goal evaluation, in which the interaction between the timing of outcome uncertainty and the value of time plays a central role in decision-making.

Psychologists have long posited that goals serve as a reference standard in a cognitive comparison process of self-evaluation and satisfaction (Latham and Locke 1991, Locke and Latham 2002, Bandura 1989), while Heath, Larrick and Wu (1999) explicitly argue that a goal acts as the reference point in the prospect theory value function formulated by Kahneman and Tversky (1979). Goal-setting as a mechanism for self-control is discussed by Loewenstein (2007) with respect to mountaineering, and has been the subject of recent interest by economists. Hsiaw (2013) finds that even in the absence of loss aversion or diminishing sensitivity, outcome-based goal-setting can attenuate the hyperbolic agent's tendency to undervalue the option and stop too early, if there is sufficient commitment to expectation-based goals. However, too much reference dependence can itself be a source of intrapersonal conflict, causing an agent to wait longer than the first best. Suvorov and van de Ven (2008) and Koch and Nafziger (2011) study a three-period problem where a sophisticated agent sets a goal regarding both current effort and delayed task benefit, while Koch, Nafziger, Suvorov and van de Ven (2014) extend the goal-setting framework to contingent self-rewards.

Kőszegi and Rabin (2006) develop a model of reference dependent preferences in which an individual derives utility from both final outcomes and comparison to a reference point endogenously determined by rational expectations, and extend it to a dynamic setting (Kőszegi and Rabin 2009). There is much experimental (Abeler, Falk, Goette and Huffman 2011, Ericson and Fuster 2011, Sprenger in press) and empirical evidence that people use reference points (Camerer, Babcock, Loewenstein and Thaler 1997, Goette and Huffman 2005, Farber 2005, Farber 2008, Crawford and Meng 2008, Pope and Schweitzer 2011, Allen, Dechow, Pope and Wu 2014).

3 The Model

I first describe the economic environment, followed by the agent's preferences, which may include hyperbolic discounting and reference dependence. I study a sequential stopping problem, where the self-control problem arises purely from the tension between stopping today and waiting for a better outcome. A number of stopping problems can realistically be cast as sequential stages. For example, an entrepreneur must build a core business before considering expansion to another location or a brand extension. A venture capitalist often must invest in an early funding round in order to participate in the next. A student must obtain a college degree before continuing to post-graduate studies. A worker searches for a series of jobs to pursue a career.

3.1 Sequential Stopping

I consider a sequential continuous-time stopping problem, in which an infinitely-lived agent is engaged in a two-stage project.[3] For example, consider an entrepreneur who is building a business. First, he can establish a storefront at some irreversible, fixed cost I_1. Completing it buys him the option to open a second location at another irreversible, fixed cost I_2. Thus, he must complete the first stage in order to continue to the second.

Formally, the agent decides whether to stop or to wait, based on an observation of the current value of the stage payoff. In the latter case, the payoff of stage i, $x_{it} \in [0, \infty)$, evolves as a geometric Brownian motion:

$$dx_{it} = \mu_i x_{it} dt + \sigma_i x_{it} dz_i, \tag{1}$$

where z_i is a standard Wiener process, μ_i the average growth rate of x_{it}, and σ_i its standard deviation per unit time. The cost of completing stage i at any time is $I_i > 0$, and is incurred only at the stopping time $\bar{t}_i$.[4] Completion of stage i at time $\bar{t}_i$ yields the lump-sum terminal payoff $x_{i\bar{t}_i}$. The second-stage payoff process x_{2t} only starts upon completion of the first stage, at time $\bar{t}_1$, and evolves thereafter independently of x_{1t}, which terminates at time $\bar{t}_1$. The processes x_{1t} and x_{2t} are

---

[3]Of course, the two-stage problem can naturally be extended into several more stages, but all intuitions remain the same.

[4]Here, there is no intertemporal separation of the costs and benefit. The agent's self-control problem arises purely from the tension between the option value of waiting for an uncertain period of time and stopping today at a known project value. In contrast, Brocas and Carrillo (2005) and Miao (2008) study irreversible consumption in discrete time models where costs are delayed until after consumption. Separating the costs and benefits of stopping in such a manner certainly exacerbates the self-control problem, but is not necessary to produce intrapersonal conflict.

only linked at one point in time, $\bar{t}_1$.

Without loss of generality, there is no interim flow payoff nor any direct cost incurred prior to stopping for each stage.[5] In a number of settings, including the one described above, it seems natural that the initial value of a payoff process depends on the payoff from the preceding stage. For example, the initial value of the second potential location may be lower if the entrepreneur's first storefront, from which he receives a payoff from opening it, was not as successful.

Thus, each stage of the project is a standard optimal stopping problem, where the agent can only complete the second stage by completing the first. Note, however, that there is nothing to preclude the agent from completing both stages simultaneously if it is optimal to do so.

I assume that $\mu_1 = \mu_2 \equiv \mu$ and $\sigma_1 = \sigma_2 \equiv \sigma$, but the processes x_{1t} and x_{2t} still evolve independently. None of the qualitative results rely on or require this simplification, which is made for simplicity.

3.2   Outcome Uncertainty

In many contexts, an agent must make stopping decisions in each stage $i = 1, 2$ based on imperfect observation of the project payoff value, and learns the true realization of x_{it} after stopping. For example, an entrepreneur relies on observed market conditions to form an estimate of how successful his business may be and decides when to invest accordingly, but his realized payoff could certainly differ.

Let $\tilde{x}_{it}$ be the observed payoff value in stage i. I assume a discrete, two-point distribution over the noise regarding the payoff processes for tractability. In particular, $\tilde{x}_{it} = (1 + \zeta_i)x_{it}$ for $i = 1, 2$, where ζ_i is a discrete random variable with the following distribution, which is known to the agent:

$$\zeta_i = \begin{cases} \epsilon_i & \text{with probability} \quad \frac{1}{2} \\ -\frac{\epsilon_i}{1+2\epsilon_i} & \text{with probability} \quad \frac{1}{2}, \end{cases} \tag{2}$$

where $\epsilon_i \in [0, \infty)$ and $i = 1, 2$. Uncertainty over the true payoff values x_{1t} and x_{2t} increases in ϵ, with perfect observation when $\epsilon = 0$. Given an observed $\tilde{x}_{it}$, the true value of x_{it} is either $(\frac{1}{1+\epsilon_i})\tilde{x}_{it}$

---

[5]An alternative setting might include a constant flow payoff $y \in (\underline{y}, \infty)$, where $\underline{y} < 0$ is the minimal flow payoff such that the agent stops immediately for any $\tilde{x}_{it} \geq 0$. Given the other assumptions, the inclusion of a constant flow payoff has no qualitative effect on the results, so I assume $y = 0$ for simplicity of exposition. Likewise, incorporating a stochastic flow payoff that follows a known process with known current value leads to the same qualitative results.

or $(\frac{1+2\epsilon_i}{1+\epsilon_i})\tilde{x}_{it}$ with equal probability. Thus, $E(x_{it}|\tilde{x}_{it}) = \tilde{x}_{it}$, so the agent expects to receive $\tilde{x}_{i\bar{t}_i} - I_i$ if he stops at time $\bar{t}_i$, for $i = 1,2$.[6] That is, the agent's expectation of the true payoff based on his observation is correct on average in either stage. I construct the noise as a mean-preserving spread over the expected outcome in order to isolate the effect of observational uncertainty alone. Ex-ante, nature chooses the realizations of ζ_1 and ζ_2, which are i.i.d. and fixed throughout, but unknown to the agent, who knows ϵ_1 and ϵ_2.[7] Multiplicative, rather than additive, noise prevents an observed payoff value from falling below the boundary value of zero unless the true payoff value equals zero, since this case would clearly allow the agent to infer a noise realization ζ_i. It also implies that the effect of uncertainty on comparative utility and the observed payoff process is bounded. Unless otherwise noted, qualitative results apply to both multiplicative and additive noise.

The stage 2 process (x_{2t}) does not begin until stage 1 is completed. The observed processes $\tilde{x}_{1t}$ and $\tilde{x}_{2t}$ are linked at only one point in time, $\bar{t}_1$. In particular, let $\tilde{x}_{2\bar{t}_1} = k\tilde{x}_{1\bar{t}_1}$ where $k > 0$ is known to the agent. Thereafter, the evolution of the stage 2 process is unaffected by the events of stage 1.

3.3 Time Preferences

The agent may have present-biased preferences, creating a self-control problem. I model this present-biasedness with a continuous time version of quasi-hyperbolic preferences (Harris and Laibson 2013). At any time s, an agent's preferences are divided into a "present," which lasts from time s to time $s + \tau_s$, and a "future," which arrives at time $s + \tau_s$ and persists forever. The length of the present, τ_s, is stochastic and distributed exponentially with parameter $\lambda \in [0, \infty)$.[8] When the future for this self s arrives at time $s + \tau_s$, he is replaced by a new self who takes control of decision-making. Likewise, the preferences of this self $s + \tau_s$ are divided into a "present" of length $\tau_{s+\tau_s}$ and a "future" that arrives at time $(s + \tau_s) + \tau_{s+\tau_s}$ and persists forever. Hence, when each self's "future" arrives, it "dies" and is replaced by a new self.

---

[6]More generally, any distribution of ζ_i such that $E(\frac{1}{1+\zeta_i}) = 1$ yields $E(x_{it}|\tilde{x}_{it}) = \tilde{x}_{it}$.

[7]The assumption that uncertainty is fixed ex-ante is a technical necessity, to prevent the observed payoff processes $\tilde{x}_{1t}$ and $\tilde{x}_{2t}$ from jumping discontinuously from one instant to the next.

[8]The assumption of a stochastic arrival time of the future allows for a parsimonious stationary solution to the stopping problem, but is not necessary to obtain qualitative results.

Each self s has a stochastic discount function $D_s(t)$:

$$D_s(t) = \begin{cases} e^{-\rho(t-s)} & \text{if} \quad t \in [s, s + \tau_s) \\ \beta e^{-\rho(t-s)} & \text{if} \quad t \in [s + \tau_s, \infty), \end{cases} \tag{3}$$

where $\beta \in [0, 1]$ and $\rho > 0$. To ensure that the agent never finds it optimal to wait forever in the optimal stopping problem, let $\rho > \mu$. The function $D_s(t)$ decays exponentially at the rate ρ throughout, but drops discontinuously at time $s + \tau_s$ to a fraction β of its prior level. In contrast to the discrete time version, there are two parameters that determine the degree to which an agent's behavior deviates from that of a time-consistent individual. First, the parameter β retains the same role it plays in the discrete-time version, measuring how much the future is valued relative to the present. Second, the parameter λ determines the arrival rate of the future, and thus how often preferences change. When $\beta = 1$ or $\lambda = 0$, the preferences described by Equation (3) are equivalent to those of an exponential discounter with discount rate ρ.

I assume that the agent is sophisticated, so he is fully aware of his dynamic inconsistency and would like to bring his future selves' behavior in line with his own preferences.

3.4 Reference Dependent Preferences

The agent's preferences are reference-dependent: his utility is composed of both standard consumption utility, which is based on absolute levels, and of comparison utility, which is concerned with gains and losses relative to a reference point, which here corresponds to a goal. In the optimal-stopping context with zero flow payoffs, the agent's expected consumption utility upon completing stage i at time $\bar{t}_i$ is simply his expected net terminal payoff: $E(x_{i\bar{t}_i} - I_i | \tilde{x}_{i\bar{t}_i}) = \tilde{x}_{i\bar{t}_i} - I_i$.

The agent's comparison utility is closely related to his consumption utility. It is derived by comparing his actual net terminal payoff at time $\bar{t}_i$ against his goal at that time, $r_{\bar{t}_i}$, and is governed by a piece-wise linear function $\psi(\cdot)$, given by

$$\psi(y) = \begin{cases} \alpha \eta y & \text{if} \quad y < 0 \\ \eta y & \text{if} \quad y \geq 0, \end{cases}$$

where $\alpha \geq 1$ and $\eta \geq 0$. The parameter η measures the agent's degree of reference dependence,

and can be interpreted as the degree to which he *cares about*, or *pays attention to*, the difference between his outcome and his goal. The parameter α captures his degree of loss aversion, where $\alpha = 1$ when loss aversion is absent. If the agent's goal for the completion of stage i at time $\bar{t}_i$ is $r_{\bar{t}_i}$, then the argument y is given by $x_{i\bar{t}_i} - I_i - r_{\bar{t}_i}$ and his expected comparison utility is given by $E_{\bar{t}_i}[\psi(x_{i\bar{t}_i} - I_i - r_{\bar{t}_i})|\tilde{x}_{i\bar{t}_i}]$. Thus, given an observed $\tilde{x}_{i\bar{t}}$, his expected comparison utility upon stopping at time $\bar{t}_i$ is derived by applying the comparison function $\psi(\cdot)$ to the difference between each possible realization of $x_{i\bar{t}_i}$ against this goal and weighting these comparisons linearly by their respective probabilities. However, possible losses relative to the goal are additionally weighted by α. In the absence of loss aversion, mean-zero ex-ante uncertainty over x_{1t} and x_{2t} clearly has no effect on behavior, since the consumption-utility component of the agent's preferences is linear. Moreover, it is the presence of loss aversion that leads him to prefer aggregated rather than gradual resolution of ex-ante uncertainty over outcomes.[9]

I assume that the agent only incurs comparison utility at the time at which he stops and receives the net terminal payoff. That is, although he is always aware that he will incur comparison utility at the moment of stopping, he does *not* directly experience it while waiting. This assumption accords with the notion from mental accounting that individuals do not necessarily "feel" gains and losses until they have been realized (Thaler 1999). For example, the disposition effect, where stockholders are reluctant to sell losing stocks, and hence realize losses relative to their original buying prices, is consistent with this idea (Odean 1998, Barberis and Xiong 2012). Imas (2015) finds that individuals' risk attitudes differ when prior losses are realized or paper.

For simplicity, overall utility is taken to be additively separable in its two components. Thus, given a goal r_{it_i} to be evaluated upon completion of stage i, the agent's expected total utility upon completing stage i is

$$E_{\bar{t}_i}\left(x_{i\bar{t}_i} - I_i + \psi(x_{i\bar{t}_i} - I_i - r_{i\bar{t}_i})|\tilde{x}_{i\bar{t}_i}\right). \tag{4}$$

In the absence of such a goal for stage i, the agent does not make any comparison upon completion of stage i and the last term of Equation (4) is omitted.

---

[9]Similarly, Kőszegi and Rabin (2009) find that loss aversion leads agents to prefer information to be received in clumps rather than spread apart. Palacios-Huerta (1999) demonstrates that an agent with Gul's (1991) disappointment aversion is also averse to the sequential resolution of uncertainty. Dillenberger (2010) finds that having recursive, non-expected utility preferences is not sufficient to exhibit a preference for one-shot resolution of uncertainty. An individual must also exhibit Negative Certainty Independence, a feature analogous to the "certainty effect" (Kahneman and Tversky 1979) without requiring non-linear probability weighting.

3.5 Goal Bracketing

In addition to setting the level of his goals, the agent can choose when and how he evaluates them. I assume that there exists a "self 0," an ex-ante self, who learns that the sequential investment opportunity will present itself in future and forms an expectation of how he will behave once the option becomes available for exercise. Self 0 determines how to bracket goals and the corresponding goal levels given the bracketing choice. Because self 0 brackets and sets goals from an ex-ante perspective for his descendants, this decision is made by maximizing ex-ante welfare. That is, self 0 prefers that all future selves behave as though they were exponential discounters. At any time s, the goal r_s is taken as given by self s and cannot be changed during his entire "lifetime," having been set by self 0. The assumption that the agent cannot change an inherited goal implies that such a goal can provide a degree of internal motivation to his (present-biased) future selves.

In the two-stage problem, his options for bracketing and setting goals are intuitive. First, he can specify incremental goals for each stage of the project, framing the problem narrowly. That is, he can set goals for the net terminal payoffs of stages i, denoted r_i^{inc} for $i = 1, 2$, and evaluate himself against r_i^{inc} upon completion of each stage i. In the context of the entrepreneur, he can set goals for each stage of his business project, evaluating himself against individual goals for each. Alternatively, he can specify an aggregate goal for the entire project, framing the problem broadly. That is, he can set a goal regarding the total net payoff from the entire project, denoted r_2^{agg}, and evaluate the sum of net payoffs from both stages against r_2^{agg} upon completion of the *entire* project. When the agent sets an aggregate goal, he only derives comparison utility at the end of the second stage. However, when making the stopping decision in the first stage, he is aware that he will make a comparison at the end of the entire project, so his behavior in each stage will contribute to his overall evaluation.

Because the agent is sophisticated and correctly anticipates his actions, I assume that each self, including self 0, has rational expectations about goal achievement. That is, he cannot consistently fool himself about what he can or cannot achieve - he sets goals that are realistic.[10] Thus, self 0 prefers that future selves behave like exponential discounters, but brackets and sets goals with

---

[10]Based on the results of lab and field experiments, Latham and Locke (1991) conclude that goal choice integrates what one wants and what one believes is possible, suggesting that goals must be, and are, realistic to the agent. In Hsiaw (2013), I study behavior when the agent's goals are not constrained to satisfy rational expectations.

the knowledge that they will respond to such goals with present bias. Because he has ex-ante uncertainty over the realized terminal payoff when setting a goal, I assume that his goal is the expectation of his net terminal payoff, given the observed payoff upon stopping. If the agent has a goal regarding his stage i payoff and stops at $\bar{t}_i$, then rational expectations require that $r_{\bar{t}_i} = E(x_{i\bar{t}_i}|\tilde{x}_{i\bar{t}_i}) - I_i$. The particular formulation of reference point as a degenerate distribution that is the expectation of his payoff is not essential to the intuitions that drive the main results. Although it determines the magnitude of distaste over ex-ante uncertainty, it does not affect the relevant comparative statics. The key requirements are that the agent has distaste over ex-ante uncertainty, which arises from loss aversion, and that he holds rational (endogenous) expectations, comparing possible realizations against his reference point.[11]

4   Equilibrium Construction

To determine the conditions under which setting incremental versus aggregate goals is optimal, I analyze the agent's behavior under each form of bracketing, then consider the intertemporal bracketing choice from an ex-ante perspective. Because the agent is quasi-hyperbolic and sophisticated, the problem takes on the nature of a dynamic game between successive selves. I focus on the stationary Markov equilibrium, in which each self employs the same threshold strategy in each stage.

To construct such an equilibrium, I solve the intrapersonal game backwards in the manner delineated in Grenadier and Wang (2007) and Hsiaw (2013), which study a single optimal stopping problem.[12] I apply backwards induction to determine the agent's behavior in the second stage upon completion of the first stage, then consider his behavior in the first stage. Each self anticipates that his descendants will act according to a threshold that maximizes their own current benefit of waiting, so they will face a problem that is identical to his own. Constructing the stationary solution involves, within each stage of the sequential stopping problem, searching for a fixed point such that

---

[11]There are a number of other proposed formulations of the reference point. Although there is some evidence supporting the theory of expectations as a reference point (Abeler et al. 2011, Crawford and Meng 2008), the precise formulation that individuals actually use is an unresolved empirical question. Kőszegi and Rabin (2006) assume that an agent holds a stochastic reference point when there is ex-ante uncertainty, where rational (endogenous) expectations imply that it must be the probability measure over realized outcomes. Gul's (1991) model implies that the reference point is the certainty equivalent of a chosen lottery.

[12]Grenadier and Wang (2007) solve for the stationary Markov equilibrium when the agent has quasi-hyperbolic time preferences, which is equivalent to the $\eta = 0$ case in Hsiaw (2013).

current and future selves stop at a common threshold. Thus, the agent faces ex-ante uncertainty about the stopping time, but not his stopping threshold, in a stationary Markov equilibrium. In the absence of observational noise over the payoff process, the agent would have no uncertainty over the payoff he received upon stopping. Henceforth, I refer to ex-ante *outcome* uncertainty due to $\epsilon_i > 0$ simply as ex-ante uncertainty for brevity.

Since the agent incurs no flow utility while waiting, the Bellman equations describing his decision problem only differ in his total utility upon stopping, given his bracketing and goal choices and the current stage. Let the function $\Phi_i^b(x_i, r_i^b)$ describe the current self's utility upon stopping in stage $i = 1, 2$, given his bracketing choice b of either incremental goals (inc) or an aggregate goal (agg), his corresponding goal r_i^b, and the current payoff value x_i. Likewise, let $\phi_i^b(x_i, r_i^b)$ denote his consideration of future selves' utility from stopping. Using these generic stopping values, I will solve for his optimal threshold, then obtain specific expressions for each bracketing choice b and stage i by substituting for $\Phi_i^b(x_i, r_i^b)$ and $\phi_i^b(x_i, r_i^b)$ appropriately in the following sections.

Because each self controls the stopping decision in the present, and cares about - but cannot directly control - those of the future, two value functions are required to describe the intrapersonal problem in a given stage. The <u>continuation</u> value function, denoted $v_i^b(\cdot)$ where $i = \{1, 2\}$ and $b = \{inc, agg\}$, describes each self s's consideration (or internalization) of his future selves, following the random arrival of the future at time τ_s. Denoting the goal inherited by future selves as $\hat{r}_i$, the <u>continuation</u> value function in stage i is

$$\rho v_i^b(x_i, \hat{r}_i^b) = \mu x_i \left(\frac{\partial v_i^b}{\partial x_i}\right) + \frac{1}{2}\sigma^2 x_i^2 \left(\frac{\partial^2 v_i^b}{\partial x_i^2}\right). \tag{5}$$

Beyond time τ_s, he discounts any future utility flows exponentially at rate ρ. For this reason, it also describes his preference for future selves to behave as exponential discounters. That is, he prefers that future selves choose the maximum of the current total utility from stopping stage i, described by $\phi_i^b(x_i, \hat{r}_i^b)$, and the expected present discounted value of waiting for a higher realization of $\tilde{x}_i$, where this discounting is exponential. If the agent were time consistent ($\beta = 1$ or $\lambda = 0$), then all selves' preferences would coincide and he would choose the optimal strategy by maximizing v_i^b.

However, if the agent is present-biased ($\beta < 1$ and $\lambda > 0$), he maximizes the <u>current</u> value function, denoted $w_i^b(\cdot)$ where $i = 1, 2$, which overweights the present relative to the future. Denoting

the goal inherited by the current self as r_i^b, the Bellman equation for the <u>current</u> value function is

$$w_i^b(x_i, r_i^b) = \max\{E[\Phi_i(x_i, r_i^b)|\tilde{x}_i], (1 - e^{-\lambda dt})e^{-\rho dt}\beta E[v_i^b(x_i + dx_i, \hat{r}_i^b)|\tilde{x}_i]$$
$$+ (e^{-\lambda dt})e^{-\rho dt}E[w_i^b(x_i + dx_i, r_i^b)|\tilde{x}_i]\}. \tag{6}$$

Given the observed $\tilde{x}_i$ and an inherited goal r_i^b, and anticipating that his future selves will inherit $\hat{r}_i^b$ (with the knowledge that he sets $\hat{r}_i^b$ for his immediate descendant), the current self chooses the maximum of the current total utility from stopping, described by $\Phi_i^b(x_i, r_i^b)$, and the expected present discounted value of waiting for a higher realization of $\tilde{x}_i$, where this discounting discontinuously drops by the factor β upon future's arrival. A future self arrives in the next instant dt with probability $1 - e^{-\lambda dt}$, while the current self remains in control with probability $e^{-\lambda dt}$.

The agent's expectation of his terminal payoff, and hence his goal, is dependent on whether he completes the second stage simultaneously with or strictly after the first. In this paper, I assume that he completes stages sequentially, then find the conditions required for such a strategy to be optimal. Because the case of simultaneous completion includes the same intuitions but lacks the richer dynamics that arise from intertemporal separation, I omit that analysis and focus on the more interesting sequential scenario.

5 Incremental Goals

First, consider the case in which the agent sets incremental goals for the net terminal payoffs of stages i, denoted r_i^{inc} for $i = 1, 2$. He evaluates himself against a goal r_i^{inc} only upon completion of stage i. Given that he sets incremental goals for himself, let $\overline{x}_i^{inc}$ be the stopping threshold that the agent employs to complete stage $i = 1, 2$. I apply backwards induction to obtain the optimal thresholds employed in each stage.

Because the comparison utility function is kinked at the origin, I derive $\overline{x}_i^{inc}$ under the assumption that r_i^{inc} is such that

$$(\frac{1}{1 + \epsilon_i})\overline{x}_i^{inc} - I_i \leq r_i^{inc} \leq (\frac{1 + 2\epsilon_i}{1 + \epsilon_i})\overline{x}_i^{inc} - I_i, \tag{7}$$

which is satisfied in equilibrium when expectations are rational. Thus his expected comparison

utility upon stopping at $\overline{x}_i^{inc}$ is

$$E[\psi(x_i - I_i - r_i^{inc})|\tilde{x}_i = \overline{x}_i^{inc}] = \frac{1}{2}\alpha\eta[(\frac{1}{1+\epsilon_i})\overline{x}_i^{inc} - I_i - r_i^{inc}] + \frac{1}{2}\eta[(\frac{1+2\epsilon_i}{1+\epsilon_i})\overline{x}_i^{inc} - I_i - r_i^{inc}]. \quad (8)$$

In the case of incremental goals, the current self evaluates current and future comparison utility the same way, since they involve only instantaneous utilities at the stopping time: $E[\psi(x_i - I_i - r_i^{inc})|\tilde{x}_i = \overline{x}_i^{inc}] = E[\Psi(x_i - I_i - r_i^{inc})|\tilde{x}_i = \overline{x}_i^{inc}]$. The agent employs the following stopping threshold $\overline{x}_i^{inc}$ in stage i:[13]

$$\overline{x}_i^{inc} = \frac{\overline{\gamma}[I_i + \frac{1}{2}\eta(\alpha + 1)(r_i^{inc} + I_i)]}{(\overline{\gamma} - 1)[1 + \frac{1}{2}\eta(\frac{1+\alpha+2\epsilon_i}{1+\epsilon_i})]}, \quad (9)$$

where $\overline{\gamma} \equiv \beta\gamma_1 + (1 - \beta)\gamma_2$, $\gamma_1 > 1$ is the positive root[14] of the quadratic equation

$$\frac{1}{2}\sigma^2\gamma_1^2 + (\mu - \frac{1}{2}\sigma^2)\gamma_1 - \rho = 0, \quad (10)$$

and $\gamma_2 \geq \gamma_1$ is the positive root[15] of the quadratic equation

$$\frac{1}{2}\sigma^2\gamma_2^2 + (\mu - \frac{1}{2}\sigma^2)\gamma_2 - (\rho + \lambda) = 0. \quad (11)$$

As in Hsiaw (2013), the parameter γ_1 reflects the fact that the agent discounts the future exponentially at the rate ρ, while the parameter γ_2 reflects the fact that each self's expected "lifetime" ends with hazard rate λ. The degree to which this feature affects behavior is determined by his degree of present-biasedness, measured by $1 - \beta$. Thus, the parameter $\overline{\gamma} = \beta\gamma_1 + (1 - \beta)\gamma_2$ serves as a sufficient statistic for measuring the agent's impulsiveness, which is determined by both β and λ. The self-control problem is absent when $\overline{\gamma} = \gamma_1$ and worsens as $\overline{\gamma}$ increases.

Proposition 1 *In a stationary equilibrium with incremental goals at each stage, the agent's stopping threshold in stage i, $\overline{x}_i^{inc}$, exhibits the following properties:*

---

[13]Unsurprisingly, in the absence of loss aversion ($\alpha = 1$), the threshold $\overline{x}_i^{inc}$ reduces to the threshold $\overline{x}^{SE}$ found in Hsiaw (2013), which describes the sophisticate agent's optimal stopping threshold in a single-stage stopping problem in the absence of loss aversion.

[14]The negative root is ruled out by the boundary condition for $x = 0$. Writing out γ_1 explicitly, we have $\gamma_1 = -\frac{\mu}{\sigma^2} + \frac{1}{2} + \sqrt{(\frac{\mu}{\sigma^2} - \frac{1}{2})^2 + \frac{2\rho}{\sigma^2}}$. To see that $\gamma_1 > 1$, note that $\sigma^2 > 0$ and the left-hand side of the quadratic is negative when evaluated at $\gamma_1 = 0$ and $\gamma_1 = 1$, implying that the negative root is strictly negative and the positive root is strictly greater than 1 if $\mu < \rho$.

[15]Again, the negative root is ruled out by the boundary condition for $x = 0$. It follows that $\gamma_2 \geq \gamma_1$ because $\lambda \geq 0$, with equality only if $\lambda = 0$.

1. *The stage i threshold is unaffected by stage j ex-ante uncertainty, where $j \neq i$: $\frac{\partial \overline{x}_i^{inc}}{\partial \epsilon_j} \geq 0$.*

2. *The stage i threshold increases with stage i ex-ante uncertainty if the agent is loss averse: $\frac{\partial \overline{x}_i^{inc}}{\partial \epsilon_i} \geq 0$, with equality only if $\alpha = 1$ or $\eta = 0$.*

3. *The stage i threshold increases with the degree of loss aversion: $\frac{\partial \overline{x}_i^{inc}}{\partial \alpha} > 0$.*

When the agent sets incremental goals for each stage, he employs thresholds in each as if they were separate stopping decisions, just as a standard agent without reference dependent preferences would. Note that the standard agent's stopping thresholds correspond to $x_i^{inc}(\eta = 0)$ for $i = 1, 2$. Once its option has been acquired, the stopping decision in stage 2 is completely independent of the first stage, since the incremental goal for stage 2 only pertains to the outcome of stage 2. Moreover, his incremental goal for stage 1 only pertains to the outcome of stage 1, so his stage 1 decision is unaffected by stage 2 uncertainty.

It is only in the presence of loss aversion ($\alpha > 1$) that the agent dislikes ex-ante, mean-zero outcome uncertainty, since it leads him to overweight the possibility of a loss. In this case, expected comparative disutility, given by Equation (8), increases with the degree of uncertainty, measured by ϵ_i, leading him to wait for a higher realization of the project value in order to compensate for the anticipated loss. Increasing loss aversion implies increasing weight on the possibility of a loss, leading the agent to wait longer to compensate for it.

The agent expects to receive a terminal payoff that is determined by stopping threshold $\overline{x}_i^{inc}$. Rational expectations imply that his incremental goal satisfies $r_i^{inc} = \overline{x}_i^{inc} - I_i$. Substituting this condition into Equation (9) yields the threshold $\overline{x}_i^{inc}$ when he sets incremental goals and stops sequentially:[16]

$$\overline{x}_i^{inc} = \frac{\overline{\gamma} I_i}{(\overline{\gamma} - 1)[1 - \frac{1}{2}\eta(\alpha - 1)(\frac{\epsilon_i}{1 + \epsilon_i})] - \frac{1}{2}\eta(\alpha + 1)}, \tag{12}$$

where $\frac{1}{2}\eta(\alpha + 1) < (\overline{\gamma} - 1)[1 - \frac{1}{2}\eta(\alpha - 1)(\frac{\epsilon_i}{1 + \epsilon_i})]$. In equilibrium, the thresholds $\overline{x}_2^{inc}$ and $\overline{x}_1^{inc}$ differ only in the fixed stopping cost I_i and degrees of noise ϵ_i where $i = 1, 2$. The expected equilibrium value functions, w_1^{inc} and v_1^{inc} in the first stage are provided in Appendix 10.2.

---

[16]Clearly, condition (7) is satisfied here.

6 Aggregate Goals

Now, consider the case in which the agent specifies an aggregate goal for the entire project, framing the problem broadly. He sets a goal regarding the total net payoff from the entire project, denoted r_2^{agg}, and evaluates the sum of net payoffs from both stages against r_2^{agg} upon completion of the entire project. When the agent sets an aggregate goal, he has no goal against which to evaluate himself in the first stage and derives no comparison utility upon its completion. Given that he sets an aggregate goal for himself, let $\overline{x}_i^{agg}$ be the stopping threshold that the agent employs to complete stage $i = 1, 2$.

In the standard case without reference dependent preferences (i.e., $\eta = 0$), the outcome of stage 1 is irrelevant to the decision in stage 2. Likewise, when the agent sets incremental goals for each stage, his behavior in stage 2 is unaffected by the outcome of stage 1. But when the agent sets an aggregate goal, information about the stage 1 outcome becomes relevant to his behavior in stage 2, because it enters into his evaluation relative to r_2^{agg}. Thus, there are two key differences between incremental and aggregate goals: the first is the nature of goal evaluation, and the second is the relevance of information regarding the outcome of the completed first stage.

6.1 Stage 2

Upon completing the project, it is the *sum* of the payoffs that the agent expects from both stages that he compares against his aggregate goal. When making this stopping decision, he knows that he has stopped stage 1 at the threshold $\overline{x}_1^{agg}$ and has observed the true realization $x_{1\overline{t}}$. Therefore, his second-stage stopping threshold will be contingent on this information. Let $\overline{x}_2^{aggL}$ be his stage 2 stopping threshold when he has learned that his stage 1 payoff was lower than expected: $x_{1\overline{t}_1} = (\frac{1}{1+\epsilon_1})\overline{x}_1^{agg}$. Let $\overline{x}_2^{aggH}$ be his stage 2 stopping threshold when he has learned that his stage 1 payoff was higher than expected: $x_{1\overline{t}_1} = (\frac{1+2\epsilon_1}{1+\epsilon_1})\overline{x}_1^{agg}$.

Because the comparison utility function is kinked at the origin, I derive $\overline{x}_2^{aggL}$ and $\overline{x}_2^{aggH}$ under the assumption the stage 2 payoff realization has a marginal effect on whether he ends up in the

gain or loss region. That is, the aggregate goal r_2^{agg} satisfies:

$$(\frac{1+2\epsilon_2}{1+\epsilon_2})\overline{x}_1^{agg} - I_1 + (\frac{1}{1+\epsilon_2})\overline{x}_2^{aggH} - I_2 \leq r_2^{agg} \leq (\frac{1+2\epsilon_2}{1+\epsilon_2})\overline{x}_1^{agg} - I_1 + (\frac{1+2\epsilon_2}{1+\epsilon_2})\overline{x}_2^{aggH} - I_2$$

$$(\frac{1}{1+\epsilon_2})\overline{x}_1^{agg} - I_1 + (\frac{1}{1+\epsilon_2})\overline{x}_2^{aggL} - I_2 \leq r_2^{agg} \leq (\frac{1}{1+\epsilon_2})\overline{x}_1^{agg} - I_1 + (\frac{1+2\epsilon_2}{1+\epsilon_2})\overline{x}_2^{aggL} - I_2.$$

These conditions are satisfied in equilibrium under rational expectations whenever $\epsilon_1 \leq \epsilon_2$, or as long as $\epsilon_1 - \epsilon_2 > 0$ is not too large. Then the contingent thresholds employed in stage 2 as a function of any given goal r_2^{agg} are:

$$\overline{x}_2^{aggL} = \frac{\overline{\gamma}[I_2 - \frac{1}{2}\eta(\alpha+1)(\frac{1}{1+\epsilon_1})\overline{x}_1^{agg} + \frac{1}{2}\eta(\alpha+1)(r_2^{agg} + I_1 + I_2)]}{(\overline{\gamma}-1)[1 + \frac{1}{2}\eta(\frac{1+\alpha+2\epsilon_2}{1+\epsilon_2})]} \tag{13}$$

$$\overline{x}_2^{aggH} = \frac{\overline{\gamma}[I_2 - \frac{1}{2}\eta(\alpha+1)(\frac{1+2\epsilon_1}{1+\epsilon_1})\overline{x}_1^{agg} + \frac{1}{2}\eta(\alpha+1)(r_2^{agg} + I_1 + I_2)]}{(\overline{\gamma}-1)[1 + \frac{1}{2}\eta(\frac{1+\alpha+2\epsilon_2}{1+\epsilon_2})]} \tag{14}$$

Because the agent compares the sum of net project payoffs from both stages against a given goal, Equations (13) and (14) are decreasing in $\overline{x}_1^{agg}$. Expecting to receive a larger payoff from the first stage brings the agent closer to a given aggregate goal and decreases the potential penalty from falling short of it for any $x_{2\overline{t}_2}$, weakening the agent's motivation to wait longer in the second stage. Since both $\overline{x}_2^{agg}$ and $\overline{x}_1^{agg}$ contribute to his comparison against a given goal r_2^{agg}, they act as motivational substitutes in the agent's stopping behavior across stages.

Imposing rational expectations, so that $r_2^{agg} = \overline{x}_1^{agg} - I_1 + \frac{1}{2}(\overline{x}_2^{aggL} + \overline{x}_2^{aggH}) - I_2$, the contingent stopping thresholds employed in stage 2 are

$$\overline{x}_2^{aggL} = \overline{x}_2^{inc} + \left(\frac{\frac{1}{2}\overline{\gamma}\eta(\alpha+1)(\frac{\epsilon_1}{1+\epsilon_1})}{(\overline{\gamma}-1)[1 - \frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_2}{1+\epsilon_2}) + \frac{1}{2}\eta(\alpha+1)]} \right) \overline{x}_1^{agg} \tag{15}$$

$$\overline{x}_2^{aggH} = \overline{x}_2^{inc} - \left(\frac{\frac{1}{2}\overline{\gamma}\eta(\alpha+1)(\frac{\epsilon_1}{1+\epsilon_1})}{(\overline{\gamma}-1)[1 - \frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_2}{1+\epsilon_2}) + \frac{1}{2}\eta(\alpha+1)]} \right) \overline{x}_1^{agg}. \tag{16}$$

Proposition 2 *When the agent is reference dependent ($\eta > 0$) and sets an aggregate goal, his stage 2 behavior responds to* unexpected *payoff-irrelevant gains and losses from stage 1. He exhibits the sunk cost fallacy by waiting longer in response to unexpected losses: $\frac{\partial \overline{x}_2^{aggL}}{\partial \epsilon_1} > 0$. In addition, he "coasts" by stopping earlier in response to unexpected gains: $\frac{\partial \overline{x}_2^{aggH}}{\partial \epsilon_1} < 0$.*

Equations (15) and (16) illustrate the effect of an aggregate goal on the agent's behavior. For

a standard agent without reference dependent preferences, information about the stage 1 outcome is not materially payoff-relevant in stage 2, so he would ignore such information when choosing his stage 2 stopping threshold. In contrast, the agent who sets an aggregate goal reacts to news about the stage 1 outcome, because it enters into his aggregate goal comparison.

The first terms of $\overline{x}_2^{aggL}$ and $\overline{x}_2^{aggH}$ are identical, since they indicate the agent's stage 2 stopping threshold if his realized stage 1 payoff were equal to his stage 1 stopping threshold (e.g., if $\epsilon_1 = 0$). If he knows he has neither fallen short of nor exceeded his expectation of the outcome of stage 1, this component of his goal exerts no influence on his decision in stage 2. As a result, he behaves as though the outcome of stage 1 is irrelevant to his decision, just as he does when he sets incremental goals for each stage and just as a standard agent without reference dependent preferences would. Thus, $\overline{x}_2^{aggL}(\epsilon_1 = 0) = \overline{x}_2^{aggH}(\epsilon_1 = 0) = \overline{x}_2^{inc}$.

The second terms of $\overline{x}_2^{aggL}$ and $\overline{x}_2^{aggH}$ indicate the agent's reaction to news that his stage 1 payoff has not met his expectation of it. If it has fallen short, he is motivated to "make up" for the shortfall by waiting for a higher payoff in stage 2 to meet his aggregate goal. Thus, he exhibits the sunk cost fallacy, but only in response to an *unexpected* loss. But if his stage 1 payoff has exceeded his expectation, he is closer to reaching his aggregate goal than expected, so he can settle for a lower payoff in stage 2 to meet his aggregate goal. That individuals often respond to sunk costs has been well documented (Kahneman and Tversky 1979, Arkes and Blumer 1985), and reference dependence has been proposed as a mechanism (Thaler 1980, Thaler 1999). While a distinction between planned and unexpected losses has not been made, many classic sunk cost examples actually involve reactions to interim surprises. For example, betting on longshots increases during the horseracing day (McGlothlin 1956, Ali 1977). While this behavior has been interpreted as poor adaptation to losses (Thaler 1980), it is also consistent with having an aggregate goal for the day's gambling profits, which is evaluated only at the end of the day.

In the psychology literature on goal-setting, there is extensive evidence that individuals are motivated when lagging behind on their goals (Locke and Latham 2002, Koo and Fishbach 2008).[17] While less studied, there is also evidence suggesting that individuals "coast" when their progress toward a goal is unexpectedly high. That is, they reduce subsequent effort toward accomplishing

---

[17]More precisely, individuals are motivated when lagging behind their goals, in the absence of uncertainty about their ability to accomplish them, as implicitly modeled here. When past behavior is a signal of willpower, Bénabou and Tirole (2004) show that lagging behind can be demotivating.

goals (Fulford, Johnson, Llabre and Carver 2010), or instead pursue unrelated or even conflicting goals (Louro, Pieters and Zeelenberg 2007, Fishbach and Dhar 2005).

Note that the agent's *expected* stage 2 stopping threshold, and therefore payoff, does not depend on ϵ_1: $E(\overline{x}^{agg}) = \overline{x}_2^{inc}$. Although loss aversion implies asymmetric utility over gains and losses, his *response* to good and bad news regarding the stage 1 outcome is not asymmetric when it is his *current* behavior that has a marginal effect on whether he gains or loses relative to the aggregate goal. Thus, when the agent sets an aggregate goal, the effect of stage 1 uncertainty is to introduce mean-zero variance into his stage 2 *behavior*.

Just as in the case of incremental goals, the aggregate goal leads the agent to wait longer as ϵ_2 increases: $\frac{\partial \overline{x}_2^{aggL}}{\partial \epsilon_2} > 0$ and $\frac{\partial \overline{x}_2^{aggL}}{\partial \epsilon_2} > 0$. Because he expects more disutility when he compares his payoffs against his aggregate goal, he waits for a higher realization of the stage 2 project value to compensate for the anticipated loss.

Corollary 1 *When the agent sets an aggregate goal, his reaction to news is amplified by the size of his payoff in stage 1 and his degree of uncertainty in either stage:* $\left| \frac{\partial \overline{x}_2^{aggL}}{\partial \epsilon_1} \right| = \left| \frac{\partial \overline{x}_2^{aggH}}{\partial \epsilon_1} \right| > 0$, $\left| \frac{\partial \overline{x}_2^{aggL}}{\partial \overline{x}_1^{agg}} \right| = \left| \frac{\partial \overline{x}_2^{aggH}}{\partial \overline{x}_1^{agg}} \right| > 0$, *and* $\left| \frac{\partial^2 \overline{x}_2^{aggL}}{\partial \epsilon_1 \epsilon_2} \right| = \left| \frac{\partial^2 \overline{x}_2^{aggH}}{\partial \epsilon_1 \epsilon_2} \right| > 0$.

Unsurprisingly, the extent of the agent's reaction to news about the stage 1 outcome increases with ϵ_1, which measures the size of the gap between outcome and expectation. The latter two properties of Corollary 1 result from the assumption that the observed payoff is proportional, rather than additive, to the true payoff. Because the magnitude of this gap is proportional to the payoff, the degree of his reaction also increases with $\overline{x}_1^{agg}$. Moreover, the magnitude of the agent's reaction to news is proportional to ϵ_2, because his *marginal* incentive to wait increases with anticipated comparative disutility when ex-ante uncertainty is multiplicative.

Note that loss aversion is not necessary for Proposition 2 to hold. Since the agent knows *ex post* whether his stage 1 payoff is than or less than he expected, his subsequent stage 2 behavior will be influenced by whether he is "ahead" or "behind" his prior expectation as long as he is reference dependent. In ex post situations, loss aversion serves to amplify his reaction to information regarding the gap between realized payoff and prior expectation.

The equilibrium value functions for the second stage, w_2^{agg} and v_2^{agg}, are provided in Appendix 10.2. They differ from those of incremental goals in the goal comparison that is being made and

the expected comparison utility given the realization of $x_{1\bar{t}}$.

6.2 Stage 1

In contrast to the case of incremental goals, the agent makes no direct evaluation against a goal upon completing the first stage. However, he *is* aware that he will be comparing the sum of net payoffs from *both* stages to the aggregate goal r_2^{agg} upon completion of the second stage. This knowledge is reflected in the option value of stage 2 that he obtains upon completion of the first stage. Thus, the current self's total utility upon stopping stage 1 is given by $E(\Phi_1^{agg}(x_1, r_2^{agg})|\tilde{x}_1)$, where

$$E(\Phi_1^{agg}(x_1, r_2^{agg})|\tilde{x}_1) = E(x_1 - I_1 + w_1^{agg}(x_2, r_2^{agg})|\tilde{x}_1). \tag{17}$$

His consideration of future selves' stopping utility for stage 1 is given by $E(\phi_1^{agg}(x_1, r_2^{agg})|\tilde{x}_1)$, where

$$E(\phi_1^{agg}(x_1, r_2^{agg}|\tilde{x}_1)) = E(x_1 - I_1 + v_1^{agg}(x_2, r_2^{agg})|\tilde{x}_1). \tag{18}$$

The stopping values $E(\Phi_1^{agg}(x_1, r_2^{agg})|\tilde{x}_1)$ and $E(\phi_1^{agg}(x_1, r_2^{agg}|\tilde{x}_1))$ differ in their last terms, since the agent considers the option to complete stage 2 differently depending on whether it is obtained by the current or future self.

Imposing the requirement that $r_2^{agg} = \overline{x}_1^{agg} - I_1 + \frac{1}{2}(\overline{x}_2^{aggL} + \overline{x}_2^{aggH}) - I_2$ yields the following nonlinear equation to describe the stationary threshold that the agent employs in stage 1, given an aggregate goal and rational expectations:

$$\overline{x}_1^{agg} = (\frac{\overline{\gamma}}{\overline{\gamma} - 1})I_1 + (\frac{\overline{x}_1^{agg}}{4})\eta(\alpha + 1)(\frac{1}{\overline{\gamma} - 1})^2 \left(\beta k^{\gamma_1}(\gamma_1 - 1) \left((\frac{1}{1 + \epsilon_1})\frac{\overline{x}_1^{agg}}{\overline{x}_2^{aggL}})^{\gamma_1} + (\frac{1 + 2\epsilon_1}{1 + \epsilon_1})(\frac{\overline{x}_1^{agg}}{\overline{x}_2^{aggH}})^{\gamma_1} \right) \right.$$

$$\left. + (1 - \beta)k^{\gamma_2}(\gamma_2 - 1) \left((\frac{1}{1 + \epsilon_1})(\frac{\overline{x}_1^{agg}}{\overline{x}_2^{aggL}})^{\gamma_2} + (\frac{1 + 2\epsilon_1}{1 + \epsilon_1})(\frac{\overline{x}_1^{agg}}{\overline{x}_2^{aggH}})^{\gamma_2} \right) \right). \tag{19}$$

Proposition 3 *Setting an aggregate goal is less effective at curbing impatience than setting incremental goals: $\overline{x}_1^{agg} < \overline{x}_1^{inc}$ and $E(\overline{x}_2^{agg}) = \overline{x}_2^{inc}$. However, the aggregate goal does induce patience even in the first stage, relative to its absence.*

Although the agent does not directly evaluate himself upon completion of the first stage, he anticipates that he will evaluate the sum of both stages' payoffs at the end of stage 2 when he sets

an aggregate goal. Because settling for a lower stage 1 payoff detrimentally affects his comparison utility in the future, the aggregate goal provides motivation for him to be more patient in the first stage as well. Recall that when he sets incremental goals, he expects to incur comparative disutility upon completing stage 1, due to evaluation of the goal r_1^{inc}. Consequently, he waits for a higher payoff to compensate for this disutility. But when he sets an aggregate goal, this disutility is absent from the first stage, so he does not have this immediate motivation to wait longer. Moreover, this disutility is reduced by the risk pooling effect of evaluating the outcomes of both stages together, and is incurred in the relatively distant future, so it is discounted more heavily in the first stage. In contrast, an incremental goal for the first stage is a source of larger, direct expected comparative disutility and is evaluated sooner on average, so it provides a stronger incentive to practice patience. Thus, the aggregate goal is less effective at curbing impatience than incremental goals.

Proposition 4 *In a stationary equilibrium with an endogenous, aggregate goal, the agent's stopping threshold in the first stage responds to ex-ante uncertainty differently in each stage when he stops sequentially: $\frac{\partial x_1^{agg}}{\partial \epsilon_2} \leq 0$ and $\frac{\partial x_1^{agg}}{\partial \epsilon_1} \geq 0$, with equality only if $\alpha = 1$ or $\eta = 0$.*

When the agent sets an aggregate goal and stops sequentially, his reaction to ex-ante uncertainty differs markedly from that of incremental goals in the first stage as well. Rather than being unresponsive, his stage 1 stopping threshold *decreases* as ϵ_2 increases. He anticipates that he will wait longer to complete stage 2 in order to compensate for greater anticipated comparative disutility, so the potential disutility from evaluating the aggregate goal is incurred even farther in the future, and thus it is discounted it more heavily in stage 1. For this reason, the agent chooses a *lower* stopping threshold in the first stage in reaction to more stage 2 uncertainty.

Just as with incremental goals, the agent's stage 1 threshold increases with ϵ_1 when he sets an aggregate goal. However, his motivation for doing so is quite different. While his *average* stage 2 stopping payoff does not change with ϵ_1 on average, its *variance* increases because ϵ_1 measures the degree to which he adjusts his stage 2 behavior to compensate for unexpected gains or losses. Thus, the parameter ϵ_1 introduces uncertainty over the agent's future behavior, so he waits for a higher payoff in stage 1 to compensate for the possibility of a lower stage 2 payoff. Thus, the mechanisms by which stage 1 uncertainty leads the agent to wait longer in stage 1 differ markedly depending on goal type. In the case of incremental goals, stage 1 uncertainty directly increases expected

comparative disutility from (incremental) evaluation upon stopping. In the case of aggregate goals, it increases uncertainty about future behavior and therefore the value of waiting.

In equilibrium, the agent's current and continuation value functions, denoted w_1^{agg} and v_1^{agg} respectively, are provided in Appendix 10.2. The first two terms of the expected current value function, given by Equation (45), in its wait region reflect the option value of stopping in the first stage, while the second two reflect that of stopping in the second. In contrast to the case of incremental goals, the disutility from ex-ante uncertainty is absent from the first two terms, as the agent does not directly evaluate himself against a goal in the first stage.

Since $\overline{x}_1^{agg} < \overline{x}_1^{inc}$ and $\overline{x}_2^{aggH} \leq \overline{x}_2^{inc} \leq \overline{x}_2^{aggL}$, the agent stops sequentially with either form of goal bracketing if he stops sequentially when setting an aggregate goal. This occurs whenever the following upper bound on k is satisfied:

$$k < \left(\frac{I_2}{I_1}\right)\left(\frac{(\overline{\gamma}-1)[1-\frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_1}{1+\epsilon_1})]-\frac{1}{2}\eta(\alpha+1)}{(\overline{\gamma}-1)[1-\frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_2}{1+\epsilon_2})]-\frac{1}{2}\eta(\alpha+1)}\right) - \frac{\frac{1}{2}\overline{\gamma}\eta(\alpha+1)(\frac{\epsilon_1}{1+\epsilon_1})}{(\overline{\gamma}-1)[1-\frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_2}{1+\epsilon_2})+\frac{1}{2}\eta(\alpha+1)]}.$$

$$(20)$$

Unsurprisingly, the second-stage process must start at a sufficiently low value for the agent not to invest in both stages simultaneously. The upper bound on the initial value of the second-stage process to allow sequential investment is increasing in I_2 and ϵ_2 because they increase his stage 2 stopping threshold, and decreasing in I_1 and ϵ_1 because they increase his stage 1 stopping threshold.

7   Optimal Bracketing

Having separately considered the agent's behavior given that he sets incremental or aggregate goals, I now determine the conditions under which each type of bracketing is optimal from an ex-ante perspective. Because he prefers that future selves behave in a time consistent manner, he chooses to bracket such that the *continuation* value, rather than the current value, of the project is maximized.

When the agent chooses how to bracket his goals and his goal level(s), he faces a trade-off between the benefits of additional motivation and the costs of additional comparative disutility from frequent goal evaluation. When decisions are made sequentially, the decision to aggregate goals becomes more complex. First, aggregating goals necessarily implies postponing interim evaluation,

and therefore delays the timing of evaluation. Second, aggregating goals introduces the ability to respond to interim news, and therefore introduces uncertainty over future behavior.

Proposition 5 *If the agent is sufficiently impulsive, he sets incremental goals whenever ϵ_1 is sufficiently low, and an aggregate goal when ϵ_1 is sufficiently high. That is, for $\overline{\gamma}$ sufficiently high, there exists a unique threshold $\tilde{\epsilon}_1 \geq 0$ such that $E[v_1^{inc}(x_1)|\tilde{x}_1] > E[v_1^{agg}(x_1)|\tilde{x}_1]$ when $\epsilon_1 < \tilde{\epsilon}_1$, and $E[v_1^{inc}(x_1)|\tilde{x}_1] < E[v_1^{agg}(x_1)|\tilde{x}_1]$ when $\epsilon_1 > \tilde{\epsilon}_1$.*

In the absence of stage 1 outcome uncertainty, there is no risk pooling benefit from setting an aggregate goal. Since incremental goals provide stronger motivation than aggregate goals, the agent prefers to set incremental goals rather than an aggregate goal if his self-control is sufficiently poor (i.e., $\overline{\gamma}$ is sufficiently high): $E[v_1^{agg}(x_1, r_2^{agg})|\tilde{x}_1] \leq E[v_1^{inc}(x_1, r_1^{inc})|\tilde{x}_1]$ when $\epsilon_1 = 0$. As stage 1 outcome uncertainty increases, the expected disutility incurred upon evaluation of the incremental goal increases, making it less desirable and the risk pooling benefit of the aggregate goal more attractive. Moreover, setting an aggregate goal becomes more attractive due to its effect on behavior in both stages. If the agent sets an aggregate goal, more stage 1 uncertainty leads to larger surprises, and therefore more ex-ante uncertainty regarding stage 2 behavior in reaction to such news. That is, setting an aggregate goal introduces a mean-preserving spread over the stage 2 payoff, relative to an incremental goal. This increases the option value of stage 2. Additionally, the fact that agent is motivated to wait longer in stage 1 to compensate for the possibility of a lower stage 2 payoff counteracts his impulsiveness, increasing the time value of his stage 1 option as well.

Proposition 6 *If the agent is sufficiently impulsive that he sets incremental goals given some $(\overline{\epsilon}_1, \overline{\epsilon}_2)$, then there exists no $\tilde{\epsilon}_2 \geq \overline{\epsilon}_2$ such that he sets an aggregate goal when $(\overline{\epsilon}_1, \tilde{\epsilon}_2)$.*

Surprisingly, setting an aggregate goal becomes relatively *less* appealing than setting incremental goals as stage 2 outcome uncertainty increases. This implies that if the agent sets incremental goals to counteract his impulsiveness given some environmental uncertainty, he would do so whenever stage 2 uncertainty is even greater. When stage 2 uncertainty determines whether the agent gains or loses relative to his aggregate goal, the risk pooling benefit of aggregation comes from shielding the agent from ϵ_1 shocks, rather than ϵ_2 shocks. In both cases, more stage 2 uncertainty leads the agent to wait longer on average in stage 2, to compensate for greater anticipated disutility
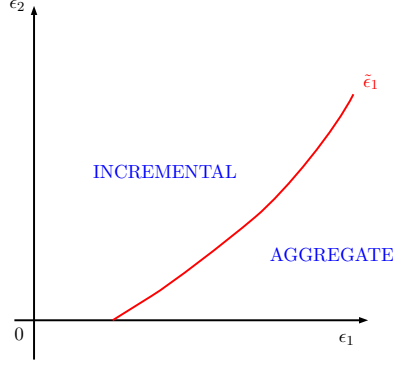
Figure 1: Optimal Bracketing

upon stopping and evaluating himself against his goal. Thus, more stage 2 uncertainty decreases the option value of stage 2 regardless of his bracketing choice, due to both waiting longer and incurring greater anticipated disutility. Because there is ex-ante uncertainty regarding stage 2 behavior in the case of aggregate goals, the (negative) marginal effect is greater under an aggregate goal than under incremental goals.[18] Moreover, when the agent has set an aggregate goal, waiting longer in stage 2 implies later goal evaluation and therefore weakens his motivation to be more patient in stage 1.

The results in Propositions 5 and 6 imply that optimal goal bracketing can be described by Figure 1 when the agent is sufficiently impulsive.

Thus, the intuition that risk pooling through an aggregate goal should become more appealing as ex-ante uncertainty increases is only partially correct when goals are evaluated sequentially rather than simultaneously. That gains or losses relative to the aggregate goal are determined by the stage 2 payoff implies that the risk pooling incentive is driven by the presence of uncertainty in stage 1. Moreover, the intertemporal separation of each stage means that setting aggregate goals implies both delayed evaluation of the stage 1 outcome and the ability to react to interim news about the stage 1 outcome. Both of these effects are anticipated ex-ante and therefore also affect stage 1 behavior and the optimal bracketing choice. Changes in ϵ_1 primarily affect the risk pooling incentive, which is amplified by these two effects of intertemporal separation. As ϵ_1 increases,

---

[18]There is an ambiguous second-order effect on the option value of stage 2. First, stopping earlier in stage 1 decreases the size of the stage 1 surprise. Because there is less uncertainty regarding the agent's future behavior, this decreases the option value of stage 2. Second, the multiplicative nature of uncertainty implies a stronger reaction to surprises, increasing the option value of stage 2. However, I show in the Appendix that the combination of these effects is always smaller than the first-order effects of waiting longer and incurring greater disutility.

delaying the expected disutility from evaluating the outcome of stage 1 and being able to react to interim surprises becomes more valuable, making the aggregate goal more attractive. Changes in ϵ_2 primarily affect the motivation incentive, which is again amplified by intertemporal separation. As ϵ_2 increases, further delayed goal evaluation weakens the motivation to wait in stage 1, making the aggregate goal less attractive.

Taken together, these results imply that we should observe aggregate goal-setting in contexts where ex-ante outcome uncertainty regarding earlier investment stages is relatively high, and ex-ante outcome uncertainty regarding the latter stages is relatively low. Otherwise, setting incremental goals is optimal when the self-control problem is sufficiently severe. That is, reactions to payoff-irrelevant news, like the sunk cost fallacy, are more likely to occur in settings where ex post realized values are more difficult to predict from ex-ante observable market conditions in the initial investment stages of a project in comparison to later investment stages. Setting an aggregate goal allows the agent to avoid the disutility from evaluating more uncertain outcomes in the initial stages and be able to react to large interim surprises without giving up much motivational power from goal setting. This implies that whether an agent reacts to sunk costs or coasts is dependent on the degree of environmental uncertainty that he faces, rather than his degree of reference dependence (and loss aversion) alone. For example, this predicts that an entrepreneur is more likely to exhibit the sunk cost fallacy when his business plan for a startup is to eventually open a second location, for which he will have likely learned enough from the first stage to better predict its payoff, than when his business plan is to eventually expand his product line, for which breaking into a new market is more unpredictable.

These results also have implications for contracting based on the nature of the project and the principal's objective. A principal whose agent is reference-dependent can expect the agent to exhibit the sunk cost fallacy or "coast" in certain projects whose outcome uncertainty is relatively high in the earlier stages and relatively low in the later stages. Therefore, an employer who wants to maximize worker output may be interested in designing a contract to counteract this behavior. On the other hand, consider a principal with the impulsive agent's welfare in mind, such as parent motivating a child. If the initial stages of the child's project are relatively difficult to predict at the stopping decision while the latter stages are not, the principal should set an aggregate goal. Otherwise, he should set incremental goals.

8 The Effect of Loss Aversion

While the detrimental effects of loss aversion have been demonstrated in contexts such as portfolio choice (Odean 1998) and contract renegotiation (Herweg and Schmidt 2015), this model implies that loss aversion can be beneficial for individual decision-making from an ex-ante perspective when the agent is sufficiently present-biased. In the presence of outcome uncertainty, the agent's degree of loss aversion has two countervailing effects on the agent's ex-ante value of the project. First, it is costly because it leads the agent to incur comparative disutility upon stopping. Second, it may be beneficial because it contributes to patient behavior for precisely this reason, as he waits longer to compensate for such expected disutility. Unsurprisingly, if the agent's degree of reference dependence is so high that it overcompensates for his self-control problem, loss aversion is unambiguously detrimental, since it both amplifies the effect of reference dependence and generates comparative disutility. But if reference dependence is not sufficient to counteract the self-control problem, some degree of loss aversion can be beneficial by counteracting impulsiveness, despite its comparative disutility. The threshold of outcome uncertainty that allows for beneficial loss aversion increases with $\overline{\gamma}$, which measures the degree of impulsiveness, and decreases with the level of reference dependence η, which amplifies the effects of loss aversion.

Proposition 7 *From an ex-ante perspective, a positive degree of loss aversion is beneficial to the agent when reference dependence (η) and outcome uncertainty (ϵ_1, ϵ_2) are sufficiently small.*

Some recent evidence suggests that agents may anticipate the beneficial effects of loss aversion if they are present-biased. In an experimental setting, Imas, Sadoff and Samek (forthcoming) find that subjects prefer performance incentives as upfront payments that they can lose over payoff-equivalent gain contracts. While this choice would be detrimental in the absence of a self-control problem, it is consistent with selection into loss contracts as a commitment device to improve performance when subjects anticipate loss aversion.

9 Conclusion

This paper develops a theoretical model of goal bracketing, which serves as a source of internal motivation to attenuate the self-control problem of a hyperbolic discounter, and offers testable

predictions of behavior in uncertain environments. When setting non-binding goals in a sequential stopping problem, an individual must decide how and when to evaluate himself against such goals. In particular, he can bracket broadly by setting an aggregate goal for the entire project, or he can bracket narrowly by setting incremental goals for individual stages. In the presence of loss aversion and uncertainty over outcomes, the intertemporal bracketing decision involves a trade-off between motivation and comparative disutility due to ex-ante uncertainty. When decisions are made sequentially, the effect of aggregating goals becomes more complex. First, aggregating goals necessarily implies postponing interim evaluation, and therefore delays the timing of evaluation. Second, it introduces the ability to respond to interim news, and therefore introduces uncertainty over future behavior. I find that when an agent sets an aggregate goal, he responds to payoff-irrelevant news, exhibiting behavior such as the sunk cost fallacy and "coasting." Either form of goal bracketing attenuates the self-control problem, but setting an aggregate goal is less effective at improving patience than setting an incremental goal. However, aggregating goals enables risk pooling and responses to interim news, so doing so becomes more appealing when initial ex-ante outcome uncertainty about a project is high. In contrast, greater ex-ante outcome uncertainty regarding later project stages primarily weakens the motivation to be patient by delaying aggregate goal evaluation, so setting incremental goals becomes more attractive when ex-ante uncertainty about later project stages is high. Thus, I find that narrow goal bracketing can be used as an instrument to counteract the self-control problem, and that broad goal bracketing can itself generate apparently erroneous behavior such as the sunk cost fallacy. Surprisingly, in contrast to a static model, the sequential nature of decision-making interacts with motivational power and risk pooling to introduce a differential reaction to outcome uncertainty based on its timing. I also show that loss aversion can, under some circumstances, be beneficial from an ex-ante perspective.

The results suggest that the prevalence of phenomena such as the sunk cost fallacy and "coasting" is quite dependent on both the nature and timing of ex-ante outcome uncertainty in the decision making process, not just an agent's degree of reference dependence and loss aversion. They also have potential ramifications for contracting, if the principal anticipates such behavior by agents in response to certain project types or is interested in goal-setting to motivate agents.

The model's predictions regarding the instrumental use of bracketing could also be tested in a controlled environment by varying the degree of ex-ante outcome uncertainty in each stage. In

particular, the comparative statics regarding the agent's response to environmental uncertainty in the first stage and whether he responds to payoff-irrelevant news differ with his bracketing choice.

Moreover, the model can be distinguished from an alternative specification, in which the agent's consumption utility is concave rather than linear and his comparison utility is either linear or non-existent. In this case, risk aversion in consumption utility implies that the agent dislikes mean-zero uncertainty, so he waits longer as uncertainty increases, just as if he has linear consumption utility but is loss averse and sets incremental goals. However, this alternative specification predicts that the bracketing choice does *not* vary in response to ex-ante uncertainty over outcomes, since there is no risk pooling benefit from aggregating anticipated comparative disutility. Thus, the agent's response to environmental uncertainty does not change in this case. In addition, an agent without comparison utility would never respond to payoff-irrelevant news.

The study of instrumental bracketing in intertemporal choice has been relatively unexplored by economists thus far. This paper offers several testable implications regarding the interaction between bracketing and ex-ante outcome uncertainty when the agent has a self-control problem due to present-biased preferences. In particular, it predicts the environments in which we can expect to observe phenomena like the sunk cost fallacy and "coasting." Experimental and empirical tests of the theory would greatly contribute to our understanding of how, why, and when individuals bracket decisions and set goals for themselves.

10 Appendix

10.1 Equilibrium Construction

To construct the continuation value function v_i^b, I suppose that all future selves inherit the goal $\hat{r}_i^b$ and employ the threshold $\hat{x}_i^b$ such that they wait if $\tilde{x}_i < \hat{x}_i^b$ and stop if $\tilde{x}_i \geq \hat{x}_i^b$. Because the geometric Brownian motion x_i, and thus $\tilde{x}_i$, changes continuously, I construct v_i^b by considering its behavior in the "wait" and "stop" regions separately, then joining them using the appropriate boundary conditions.

Because x_i is a geometric Brownian motion, $x_i = 0$ is an absorbing barrier. Clearly, the agent should never stop the process if $x_i = 0$. Moreover, the continuation value must be continuous everywhere, including at the threshold between waiting and stopping. Because there is no optimal decision embodied in the continuation value function, the smooth pasting condition does not apply to $v_i^b(x_i, \hat{r}_i^b)$ if the agent is present-biased. The stopping decision is never made by future selves, only by current selves. Thus, there are two relevant boundary conditions for v_i^b:

$$\text{Boundary: } E[v_i^b(x_i, \hat{r}_i^b)|\tilde{x}_i = 0] = 0, \tag{21}$$

$$\text{Value Matching: } E[v_i^b(x_i, \hat{r}_i^b)|\tilde{x}_i = \hat{x}_i^b] = E[\phi_i^b(x_i, \hat{r}_i^b)|\tilde{x}_i = \hat{x}_i^b]. \tag{22}$$

To construct the current value function w_i^b, I suppose that all current selves inherit the goal r_i^b and employ the threshold $\overline{x}_i^b$ such that they wait if $\tilde{x}_i < \overline{x}_i^b$ and stop if $\tilde{x}_i \geq \overline{x}_i^b$. The threshold strategy implies that the value of w_i^b in its "stop" region ($\tilde{x}_i \geq \overline{x}_i^b$) is given by $E[\Phi_i^b(x_i, r_i^b)|\tilde{x}_i]$. In its wait region ($\tilde{x}_i < \overline{x}_i^b$), standard results imply that w_i^b obeys the following linear differential equation:

$$\rho w_i^b(x_i, r_i^b) = \lambda(\beta v_i^b(x_i, \hat{r}_i^b) - w_i^b(x_i, r_i^b)) + \mu x_i \left(\frac{\partial w_i^b}{\partial x_i}\right) + \frac{1}{2}\sigma^2 x_i^2 \left(\frac{\partial^2 w_i^b}{\partial x_i^2}\right). \tag{23}$$

Comparing Equation (23) to Equation (5), the additional term $\lambda(\beta v_i^b(x_i, \hat{r}_i^b) - w_i^b(x_i, r_i^b))$ is the expected value of the change in the current value w_i^b that occurs through the stochastic arrival of a transition from the present to the future.

As with v_i^b, $x_i = 0$ is an absorbing barrier and w_i^b must be continuous everywhere. Since the optimal threshold is chosen to maximize the current value function by the current self, the smooth

pasting condition, that the marginal value of waiting equals that of stopping, must apply to w_i^b with respect to x_i. This yields the boundary conditions for w_i^b:

$$\text{Boundary: } E[w_i^b(x_i, r_i^b)|\tilde{x}_i = 0] = 0, \tag{24}$$

$$\text{Value Matching: } E[w_i^b(x_i, r_i^b)|\tilde{x}_i = \overline{x}_i^b] = E[\Phi_i^b(x_i, r_i^b)|\tilde{x}_i = \overline{x}_i^b], \tag{25}$$

$$\text{Smooth Pasting: } E[\frac{\partial w_i^b}{\partial x_i}(x_i, r_i^b)|\tilde{x}_i = \overline{x}_i^b] = E[\frac{\partial \Phi_i^b}{\partial x_i}(x_i, r_i^b)|\tilde{x}_i = \overline{x}_i^b]. \tag{26}$$

Applying conditions (21) and (22) to Equation (5) yields the solution to the continuation value function v_i^b. Under the assumption that $x_i \leq \hat{x}_i^b$, which the fixed point condition satisfies, it is the value of v_i^b in its wait region that applies to Equation (23). Combining v_i^b in its wait region with Equation (23), along with conditions (24), (25), (26), we obtain the solution to the optimal threshold $\overline{x}_i^b$ as a function of goal r_i^b and the conjectured future goals $\hat{r}_i^b$ and threshold $\hat{x}_i^b$. Moreover, stationarity implies that $\overline{x}_i^b = \hat{x}_i^b$ and $r_i^b = \hat{r}_i^b$, allowing us to obtain $\overline{x}_i^b$ as a function of the goal r_i^b.

10.2   Equilibrium Value Functions

This section collects the equilibrium current and continuation value functions that arise for each bracketing choice, with its corresponding goals, and each stage. For ease of reference, it reiterates the key features of each.

Incremental Goals: Stage 2

The value of Equation (32) in its wait region is the expected present value of the option to stop, given the current value of the project's payoff, $x < \overline{x}_2^{inc}$. This is essentially the weighted average of two time-consistent option values, where the first, weighted by β, uses the discount rate ρ, and the second, weighted by $1 - \beta$, uses the discount rate $\rho + \lambda$. Moreover, the expected present value of the stopping option reflects the comparative disutility that the agent expects to incur upon stopping and evaluating himself against his goal, which is $\frac{1}{2}\eta(\alpha - 1)(\frac{\epsilon}{1+\epsilon_2})\overline{x}_2^{inc}$ when $\tilde{x}_2 < \overline{x}_2^{inc}$. Because it reflects his preferences from an ex-ante perspective, the value of Equation (33) in its wait region is the expected present value of the option to stop, using only the discount rate ρ. In the second stage, the agent's problem is identical to a standard, single-stage optimal stopping problem. Because his goal in this stage only pertains to the outcome of stage 2, his behavior in

the first stage is irrelevant at this point.[19] Since he evaluates himself against the goal r_2^{inc} upon completion of stage 2, the current self's total utility upon stopping stage 2 is given by $E(\Phi_2^{inc}(x_2, r_2^{inc})|\tilde{x}_2)$:

$$E(\Phi_2^{inc}(x_2, r_2^{inc})|\tilde{x}_2) = E[x_2 - I_2 + \psi(x_2 - I_2 - r_2)|\tilde{x}_2], \tag{27}$$

which is simply the sum of his net terminal payoff and his expected comparison utility, and enters into Equation (6). Likewise, he anticipates that future selves obtain the same utility from stopping:

$$E(\phi_2^{inc}(x_2, r_2^{inc})|\tilde{x}_2) = E[x_2 - I_2 + \psi(x_2 - I_2 - r_2^{inc})|\tilde{x}_2], \tag{28}$$

which enters into Equation (5). Thus, the agent employs the following stopping threshold $\bar{x}_2^{inc}$ in the second stage:[20]

$$\bar{x}_2^{inc} = \frac{\bar{\gamma}[I_2 + \frac{1}{2}\eta(\alpha+1)(r_2^{inc} + I_2)]}{(\bar{\gamma}-1)[1 + \frac{1}{2}\eta(\frac{1+\alpha+2\epsilon_2}{1+\epsilon_2})]}, \tag{29}$$

where $\bar{\gamma} \equiv \beta\gamma_1 + (1-\beta)\gamma_2$, $\gamma_1 > 1$ is the positive root[21] of the quadratic equation

$$\frac{1}{2}\sigma^2\gamma_1^2 + (\mu - \frac{1}{2}\sigma^2)\gamma_1 - \rho = 0, \tag{30}$$

and $\gamma_2 \geq \gamma_1$ is the positive root[22] of the quadratic equation

$$\frac{1}{2}\sigma^2\gamma_2^2 + (\mu - \frac{1}{2}\sigma^2)\gamma_2 - (\rho + \lambda) = 0. \tag{31}$$

As in Hsiaw (2013), the parameter γ_1 reflects the fact that the agent discounts the future exponentially at the rate ρ, while the parameter γ_2 reflects the fact that each self's expected

---

[19]In the standard problem without reference dependent preferences, it is also the case that stage 1 behavior is irrelevant to the stage 2 decision. However, stage 1 behavior will *not* be irrelevant when the agent sets an aggregate goal.

[20]Unsurprisingly, in the absence of loss aversion ($\alpha = 1$), the threshold $\bar{x}_2^{inc}$ reduces to the threshold $\bar{x}^{SE}$ found in Hsiaw (2013), which describes the sophisticate agent's optimal stopping threshold in a single-stage stopping problem in the absence of loss aversion.

[21]The negative root is ruled out by the boundary condition for $x = 0$. Writing out γ_1 explicitly, we have $\gamma_1 = -\frac{\mu}{\sigma^2} + \frac{1}{2} + \sqrt{(\frac{\mu}{\sigma^2} - \frac{1}{2})^2 + \frac{2\rho}{\sigma^2}}$. To see that $\gamma_1 > 1$, note that $\sigma^2 > 0$ and the left-hand side of the quadratic is negative when evaluated at $\gamma_1 = 0$ and $\gamma_1 = 1$, implying that the negative root is strictly negative and the positive root is strictly greater than 1 if $\mu < \rho$.

[22]Again, the negative root is ruled out by the boundary condition for $x = 0$. It follows that $\gamma_2 \geq \gamma_1$ because $\lambda \geq 0$, with equality only if $\lambda = 0$.

"lifetime" ends with hazard rate λ. The degree to which this feature affects behavior is determined by his degree of present-biasedness, measured by $1 - \beta$. Thus, the parameter $\overline{\gamma} = \beta\gamma_1 + (1 - \beta)\gamma_2$ serves as a sufficient statistic for measuring the agent's impulsiveness, which is determined by both β and λ. The self-control problem is absent when $\overline{\gamma} = \gamma_1$ and worsens as $\overline{\gamma}$ increases. The equilibrium value functions for the second stage, w_2^{inc} and v_2^{inc}, are given by

$$
E[w_2^{inc}(x_2, r_2^{inc})|\tilde{x}_2] = \begin{cases} \frac{\overline{x}_2^{inc}\left(1 - \frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_2}{1+\epsilon_2}) + \frac{1}{2}\eta(\alpha+1)\right)}{\overline{\gamma}}\left(\beta(\frac{\tilde{x}_2}{\overline{x}_2^{inc}})^{\gamma_1} + (1 - \beta)(\frac{\tilde{x}_2}{\overline{x}_2^{inc}})^{\gamma_2}\right) & \text{if } \tilde{x}_2 < \overline{x}_2^{inc} \\ \tilde{x}_2 - I_2 + E[\psi(x_2 - I_2 - r_2^{inc})|\tilde{x}_2] & \text{if } \tilde{x}_2 \geq \overline{x}_2^{inc} \end{cases}
$$

$$(32)$$

$$
E[v_2^{inc}(x_2, r_2^{inc})|\tilde{x}_2] = \begin{cases} \frac{\overline{x}_2^{inc}\left(1 - \frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_2}{1+\epsilon_2}) + \frac{1}{2}\eta(\alpha+1)\right)}{\overline{\gamma}}(\frac{\tilde{x}_2}{\overline{x}_2^{inc}})^{\gamma_1} & \text{if } \tilde{x}_2 < \overline{x}_2^{inc} \\ \tilde{x}_2 - I_2 + E[\psi(x_2 - I_2 - r_2^{inc})|\tilde{x}_2] & \text{if } \tilde{x}_2 \geq \overline{x}_2^{inc}. \end{cases}
$$

$$(33)$$

Incremental Goals: Stage 1

In the first stage, the agent faces a problem very similar to that of the second, since his goal pertains only to the outcome of stage 1 and he evaluates himself upon its completion. The only difference is that in addition to receiving the project payoff $x_{1\overline{t}_1}$ upon completing stage 1 at time $\overline{t}_1$, he obtains the option to complete the second stage of the project. Thus, the current self's total utility upon stopping stage 1 is given by $E(\Phi_1^{inc}(x_1, r_1^{inc})|\tilde{x}_1)$, where

$$E(\Phi_1^{inc}(x_1, r_1^{inc})|\tilde{x}_1) = E(x_1 - I_1 + \psi(x_1 - I_1 - r_1^{inc}) + w_2^{inc}(x_2, r_2^{inc})|\tilde{x}_1), \qquad (34)$$

which enters into Equation (6) and its corresponding boundary conditions (24), (25), and (26). Note that $E(x_2|\tilde{x}_1) = kx_1$. Equation (34) only differs from the second-stage stopping utility described by (27) in its last term, the option to complete stage 2. When evaluating the possibility that a *future* self will complete the first stage, the agent considers the option to complete stage 2 by discounting it exponentially. Thus, his consideration of future selves' stopping utility for stage 1 is given by $E(\phi_1^{inc}(x_1, r_1^{inc})|\tilde{x}_1)$, where

$$E(\phi_1^{inc}(x_1, r_1^{inc})|\tilde{x}_1) = E(x_1 - I_1 + \psi(x_1 - I_1 - r_1^{inc}) + v_2^{inc}(x_2, r_2^{inc})|\tilde{x}_1), \qquad (35)$$

which enters into Equation (5) and its boundary conditions (21) and (22). The stopping values $\Phi_1^{inc}(x_1, r_1^{inc})$ and $\phi_1^{inc}(x_1, r_1^{inc})$ differ only in their last terms, since the agent values the option to complete stage 2 differently depending on whether it is obtained by the current or future self. The value of the option to complete stage 2 depends on whether it is optimal for him to complete it immediately upon completion of stage 1. Because I assume that he completes stages sequentially (without loss of generality), it is the wait regions of the stage 2 option values $w_2^{inc}(kx_1, r_2^{inc})$ and $v_2^{inc}(kx_1, r_2^{inc})$ that are applicable to Equations (34) and (35), respectively. Finally, imposing the requirement that $r_1^{inc} = \overline{x}_1^{inc} - I_1$ gives the optimal threshold when the goal is self-set and expectations are rational:

$$\overline{x}_1^{inc} = \frac{\overline{\gamma} I_1}{(\overline{\gamma} - 1)[1 - \frac{1}{2}\eta(\alpha - 1)(\frac{\epsilon_1}{1+\epsilon_1})] - \frac{1}{2}\eta(\alpha + 1)}, \tag{36}$$

where $\frac{1}{2}\eta(\alpha + 1) < (\overline{\gamma} - 1)[1 - \frac{1}{2}\eta(\alpha - 1)(\frac{\epsilon_1}{1+\epsilon_1})]$.

In the first stage, the equilibrium current and future value functions when the agent sets incremental goals are given by w_1^{inc} and v_1^{inc}, respectively and simplify to:

$$E[w_1^{inc}(x_1, r_1^{inc}, r_2^{inc})|\tilde{x}_1] = \begin{cases} \frac{\overline{x}_1^{inc}\left(1 - \frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_1}{1+\epsilon_1}) + \frac{1}{2}\eta(\alpha+1)\right)}{\overline{\gamma}}\left(\beta(\frac{\tilde{x}_1}{\overline{x}_1^{inc}})^{\gamma_1} + (1 - \beta)(\frac{\tilde{x}_1}{\overline{x}_1^{inc}})^{\gamma_2}\right) \\ \quad + k^{\gamma_1} E[w_2^{inc}(x_2, r_2^{inc})|\tilde{x}_1] & \text{if} \quad \tilde{x}_1 < \overline{x}_1^{inc} \\ \tilde{x}_1 - I_1 + E[\psi(x_1 - \overline{x}_1^{inc})|\tilde{x}_1] + k^{\gamma_1} E[w_2^{inc}(x_2, r_2^{inc})|\tilde{x}_1] & \text{if} \quad \tilde{x}_1 \geq \overline{x}_1^{inc}, \end{cases} \tag{37}$$

$$E[v_1^{inc}(x_1, r_1^{inc}, r_2^{inc})|\tilde{x}_1] = \begin{cases} \frac{\overline{x}_1^{inc}\left(1 - \frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_1}{1+\epsilon_1}) + \frac{1}{2}\eta(\alpha+1)\right)}{\overline{\gamma}}(\frac{\tilde{x}_1}{\overline{x}_1^{inc}})^{\gamma_1} + k^{\gamma_1} E[v_2^{inc}(x_2, r_2^{inc})|\tilde{x}_1] & \text{if} \quad \tilde{x}_1 < \overline{x}_1^{inc} \\ \tilde{x}_1 - I_1 + E[\psi(x_1 - \overline{x}_1^{inc})|\tilde{x}_1] + k^{\gamma_1} E[v_2^{inc}(x_2, r_2^{inc})|\tilde{x}_1] & \text{if} \quad \tilde{x}_1 \geq \overline{x}_1^{inc} \end{cases} \tag{38}$$

Because the agent completes the second stage strictly after the first, the stop regions of the current and continuation value functions, w_1^{inc} and v_1^{inc}, are composed of two regions. When the observed first-stage payoff $\tilde{x}_1$ is such that $k\tilde{x}_1 < \overline{x}_2^{inc}$, the option value of stage 2 is determined by the value of waiting to stop the process x_{2t}. When $\tilde{x}_1$ is sufficiently high that $k\tilde{x}_1 \geq \overline{x}_2^{inc}$, the option value of stage 2 is simply the value of stopping x_{2t} immediately. Since the agent evaluates himself against a goal at the end of each stage when he sets incremental goals, he expects to incur

comparative disutility at the end of each.

Aggregate Goals: Stage 2

In the second stage, the agent compares the sum of his net payoffs in each stage against his goal for it, rather than making a comparison regarding the outcome of stage 2 alone. At this point, he has stopped stage 1 at the threshold $\overline{x}_1^{agg}$ and observed the true realization $x_{1\overline{t}_1}$. Therefore, the current self's total expected utility upon stopping stage 2 is given by $E(\Phi_2^{agg}(x_2, r_2^{agg})|\tilde{x}_2)$:

$$E(\Phi_2^{agg}(x_2, r_2^{agg})|\tilde{x}_2) = E(x_2 - I_2 + \psi(x_{1\overline{t}_1} - I_1 + x_2 - I_2 - r_2^{agg})), \tag{39}$$

which enters into Equations (6) and its corresponding boundary conditions (24)-(26). Likewise, he anticipates that future selves obtain the same utility from stopping:

$$E(\phi_2^{agg}(x_2, r_2^{agg})|\tilde{x}_2) = E(x_2 - I_2 + \psi(x_{1\overline{t}_1} - I_1 + x_2 - I_2 - r_2^{agg})|\tilde{x}_2), \tag{40}$$

which enters into Equations (5) and its corresponding boundary conditions (21) and (22). In comparison to Equations (27) and (28), which describe the utility upon stopping stage 2 when he sets incremental goals, Equations (39) and (40) differ only in the goal evaluation that occurs upon completion of the project. Upon completing the project, it is the *sum* of the payoffs that the agent expects from both stages that he compares against his aggregate goal.

In the second stage, the equilibrium current and future value functions when the agent sets aggregate goals are given by w_2^{agg} and v_2^{agg}, respectively. They differ from those of incremental goals only in the goal comparison that is being made.

$$E[w_2^{aggL}(x_2, r_2^{agg})|\tilde{x}_2] = \begin{cases} \frac{\overline{x}_2^{aggL}\left(1 - \frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_2}{1+\epsilon_2}) + \frac{1}{2}\eta(\alpha+1)\right)}{\overline{\gamma}} \left(\beta(\frac{\tilde{x}_2}{\overline{x}_2^{aggL}})^{\gamma_1} + (1-\beta)(\frac{\tilde{x}_2}{\overline{x}_2^{aggL}})^{\gamma_2}\right) & \text{if} \quad \tilde{x}_2 < \overline{x}_2^{aggL} \\ \tilde{x}_2 - I_2 + E[\psi(x_1 + x_2 - \overline{x}_1^{agg} - \frac{1}{2}(\overline{x}_2^{aggL} + \overline{x}_2^{aggH}))|\tilde{x}_1] & \text{if} \quad \tilde{x}_2 \geq \overline{x}_2^{aggL} \end{cases}$$

$$\tag{41}$$

$$E[w_2^{aggH}(x_2, r_2^{agg})|\tilde{x}_2] = \begin{cases} \dfrac{\overline{x}_2^{aggH}\left(1-\frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_2}{1+\epsilon_2})+\frac{1}{2}\eta(\alpha+1)\right)}{\overline{\gamma}}\left(\beta(\frac{\tilde{x}_2}{\overline{x}_2^{aggH}})^{\gamma_1}+(1-\beta)(\frac{\tilde{x}_2}{\overline{x}_2^{aggH}})^{\gamma_2}\right) & \text{if} \quad \tilde{x}_2 < \overline{x}_2^{aggH} \\ \tilde{x}_2 - I_2 + E[\psi(x_1 + x_2 - \overline{x}_1^{agg} - \frac{1}{2}(\overline{x}_2^{aggL}+\overline{x}_2^{aggH}))|\tilde{x}_1] & \text{if} \quad \tilde{x}_2 \geq \overline{x}_2^{aggH} \end{cases}$$

(42)

$$E[v_2^{aggL}(\tilde{x}_2, r_2^{agg})|\tilde{x}_2] = \begin{cases} \dfrac{\overline{x}_2^{aggL}\left(1-\frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_2}{1+\epsilon_2})+\frac{1}{2}\eta(\alpha+1)\right)}{\overline{\gamma}}(\frac{\tilde{x}_2}{\overline{x}_2^{aggL}})^{\gamma_1} & \text{if} \quad \tilde{x}_2 < \overline{x}_2^{aggL} \\ \tilde{x}_2 - I_2 + E[\psi(x_1 + x_2 - \overline{x}_1^{agg} - \frac{1}{2}(\overline{x}_2^{aggL}+\overline{x}_2^{aggH}))|\tilde{x}_1] & \text{if} \quad \tilde{x}_2 \geq \overline{x}_2^{aggL}. \end{cases}$$

(43)

$$E[v_2^{aggH}(\tilde{x}_2, r_2^{agg})|\tilde{x}_2] = \begin{cases} \dfrac{\overline{x}_2^{aggH}\left(1-\frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_2}{1+\epsilon_2})+\frac{1}{2}\eta(\alpha+1)\right)}{\overline{\gamma}}(\frac{\tilde{x}_2}{\overline{x}_2^{aggH}})^{\gamma_1} & \text{if} \quad \tilde{x}_2 < \overline{x}_2^{aggH} \\ \tilde{x}_2 - I_2 + E[\psi(x_1 + x_2 - \overline{x}_1^{agg} - \frac{1}{2}(\overline{x}_2^{aggL}+\overline{x}_2^{aggH}))|\tilde{x}_1] & \text{if} \quad \tilde{x}_2 \geq \overline{x}_2^{aggH}. \end{cases}$$

(44)

Aggregate Goals: Stage 1

$$E[w_1^{agg}(x_1, r_2^{agg})|\tilde{x}_1] = \begin{cases} (\overline{x}_1^{agg} - I_1)\left(\beta(\frac{\tilde{x}_1}{\overline{x}_1^{agg}})^{\gamma_1}+(1-\beta)(\frac{\tilde{x}_1}{\overline{x}_1^{agg}})^{\gamma_2}\right) \\ \quad + \frac{k^{\gamma_1}}{2}\left(E[w_2^{aggL}(x_2, r_2^{agg})|\tilde{x}_1]+E[w_2^{aggH}(x_2, r_2^{agg})|\tilde{x}_1]\right) & \text{if} \quad \tilde{x}_1 < \overline{x}_1^{agg} \\ \tilde{x}_1 - I_1 + \frac{k^{\gamma_1}}{2}\left(E[w_2^{aggL}(x_2, r_2^{agg})|\tilde{x}_1]+E[w_2^{aggH}(x_2, r_2^{agg})|\tilde{x}_1]\right) & \text{if} \quad \leq \tilde{x}_1 \geq \overline{x}_1^{agg}, \end{cases}$$

(45)

$$E[v_1^{agg}(x_1, r_2^{agg})|\tilde{x}_1] = \begin{cases} (\overline{x}_1^{agg} - I_1)(\frac{\tilde{x}_1}{\overline{x}_1^{agg}})^{\gamma_1} + \frac{k^{\gamma_1}}{2}\left(E[w_2^{aggL}(x_2, r_2^{agg})|\tilde{x}_1]+E[w_2^{aggH}(x_2, r_2^{agg})|\tilde{x}_1]\right) & \text{if} \quad \tilde{x}_1 < \overline{x}_1^{agg} \\ \tilde{x}_1 - I_1 + \frac{k^{\gamma_1}}{2}\left(E[w_2^{aggL}(x_2, r_2^{agg})|\tilde{x}_1]+E[w_2^{aggH}(x_2, r_2^{agg})|\tilde{x}_1]\right) & \text{if} \quad \tilde{x}_1 \geq \overline{x}_1^{agg}. \end{cases}$$

(46)

The first two terms of the expected current value function, given by Equation (45), in its wait region reflect the option value of stopping in the first stage, while the second two reflect that of stopping in the second. In contrast to the case of incremental goals, the disutility from ex-ante uncertainty is absent from the first two terms, as the agent does not directly evaluate himself against a goal in the first stage. Likewise, the first term of the expected continuation value function, given by Equation (46), in its wait region reflects the option value of the first stage from an ex-ante perspective, while the second term reflects that of the second.

10.3 Proof of Proposition 3

Proof.

First, I show existence and uniqueness of $\overline{x}_1^{agg}$. Define the following function, $G(x)$:

$$G(x) = \overline{\gamma}I_1 + (\frac{x}{4})\eta(\alpha+1)(\frac{1}{\overline{\gamma}-1})\left(\beta k^{\gamma_1}(\gamma_1-1)\left((\frac{1}{1+\epsilon_1})(\frac{x}{\overline{x}_2^{aggL}})^{\gamma_1} + (\frac{1+2\epsilon_1}{1+\epsilon_1})(\frac{x}{\overline{x}_2^{aggH}})^{\gamma_1}\right)\right.$$

$$\left. +(1-\beta)k^{\gamma_2}(\gamma_2-1)\left((\frac{1}{1+\epsilon_1})(\frac{x}{\overline{x}_2^{aggL}})^{\gamma_2} + (\frac{1+2\epsilon_1}{1+\epsilon_1})(\frac{x}{\overline{x}_2^{aggH}})^{\gamma_2}\right)\right) - (\overline{\gamma}-1)x. \tag{47}$$

Clearly, $G(\overline{x}_1^{agg}) = 0$, so it is sufficient to verify that $G(x)$ has a unique root in the range $0 < \overline{x}_1^{agg} < \overline{x}_1^{inc}$. First, $G(0) = \overline{\gamma}I_1 > 0$. Second, G is clearly increasing in k. Therefore,

$$G(\overline{x}_1^{inc}) = \overline{\gamma}I_1 + (\frac{\overline{x}_1^{inc}}{4})\eta(\alpha+1)(\frac{1}{\overline{\gamma}-1})\left(\beta k^{\gamma_1}(\gamma_1-1)\left((\frac{1}{1+\epsilon_1})(\frac{\overline{x}_1^{inc}}{\overline{x}_2^{aggL}})^{\gamma_1} + (\frac{1+2\epsilon_1}{1+\epsilon_1})(\frac{\overline{x}_1^{inc}}{\overline{x}_2^{aggH}})^{\gamma_1}\right)\right.$$

$$\left. +(1-\beta)k^{\gamma_2}(\gamma_2-1)\left((\frac{1}{1+\epsilon_1})(\frac{\overline{x}_1^{inc}}{\overline{x}_2^{aggL}})^{\gamma_2} + (\frac{1+2\epsilon_1}{1+\epsilon_1})(\frac{\overline{x}_1^{inc}}{\overline{x}_2^{aggH}})^{\gamma_2}\right)\right) - (\overline{\gamma}-1)\overline{x}_1^{inc}$$

$$\leq \overline{\gamma}I_1 + (\frac{\overline{x}_1^{inc}}{4})\eta(\alpha+1)(\frac{1}{\overline{\gamma}-1})\left(\beta(\gamma_1-1)\left((\frac{1}{1+\epsilon_1}) + (\frac{1+2\epsilon_1}{1+\epsilon_1})\right)\right.$$

$$\left. +(1-\beta)(\gamma_2-1)\left((\frac{1}{1+\epsilon_1}) + (\frac{1+2\epsilon_1}{1+\epsilon_1})\right)\right) - (\overline{\gamma}-1)\overline{x}_1^{inc}$$

$$= \overline{\gamma}I_1 + (\frac{\overline{x}_1^{inc}}{2})\eta(\alpha+1)(\frac{1}{\overline{\gamma}-1})(\beta(\gamma_1-1)$$

$$+(1-\beta)(\gamma_2-1)) - (\overline{\gamma}-1)\overline{x}_1^{inc}$$

$$= \overline{\gamma}I_1 + \frac{1}{2}\eta(\alpha+1)\overline{x}_1^{inc} - (\overline{\gamma}-1)\overline{x}_1^{inc}$$

$$< \overline{\gamma}I_1 + \frac{1}{2}\eta(\alpha+1)\overline{x}_1^{inc} - (\overline{\gamma}-1)\overline{x}_1^{inc} + \frac{1}{2}\eta(\alpha-1)(\overline{\gamma}-1)(\frac{\epsilon}{1+\epsilon})\overline{x}_1^{inc}$$

$$= \overline{\gamma}(I_1 - I_I) = 0,$$

where the first inequality holds if $\overline{x}_2^{aggH} \geq k\overline{x}_1^{inc}$ and the last line follows from the definition of $\overline{x}_1^{inc}$. Since $G(0) > 0$ and $G(\overline{x}_1^{inc}) < 0$, then there exists at least one root in this range. Moreover, we can verify that $\frac{d^2 G(x)}{dx^2} > 0$ for all $x > 0$. Therefore, there exists a unique $\overline{x}_1^{agg} \in (0, \overline{x}_1^{inc})$. We can verify that our assumption for sequential stopping is satisfied: Since $\overline{x}_1^{agg}$ is bounded above by $\overline{x}_1^{inc}$, then $\overline{x}_2^{aggH} \geq k\overline{x}_1^{inc} \geq k\overline{x}_1^{agg}$.

Thus, $\overline{x}_1^{agg}$ exists and is unique, and $\overline{x}_1^{agg} < \overline{x}_1^{inc}$. The fact that $\overline{x}_1^{agg} > (\frac{\overline{\gamma}}{\overline{\gamma}-1})I_1$, where the right-hand side of the inequality is the agent's stopping threshold in the absence of reference

dependence ($\eta = 0$), is obvious by inspection of Equation (19). ∎

10.4   Proof of Proposition 4

To find $\frac{\partial \overline{x}_1^{agg}}{\partial \epsilon_i}$, we apply the implicit function theorem to Equation (47). We can verify that $\frac{\partial G}{\partial \epsilon_2} \leq 0$, with equality only when $\eta = 0$ or $\alpha = 1$. We can also verify that $\frac{\partial G}{\partial \epsilon_1} \geq 0$, with equality only when $\eta = 0$ or $\alpha = 1$. In the preceding proof of the existence of $\overline{x}_1^{agg}$, I have shown that $\frac{\partial G}{\partial x} < 0$ when $0 < x < \overline{x}_1^{inc}$. By the implicit function theorem,

$$\frac{\partial \overline{x}_1^{agg}}{\partial \epsilon_i} = -\frac{\frac{\partial G}{\partial \epsilon_i}}{\frac{\partial G}{\partial x}}.$$

Thus, $\frac{\partial \overline{x}_1^{agg}}{\partial \epsilon_2} \leq 0$ and $\frac{\partial \overline{x}_1^{agg}}{\partial \epsilon_1} \geq 0$, with equality only when $\eta = 0$ or $\alpha = 1$.

10.5   Proof of Proposition 5

To find the optimal bracketing choice as ϵ_1 varies, I compare the expectations of v_1^{agg} and v_1^{inc}. If Equation (20) is satisfied, then the agent completes each stage sequentially regardless of how he brackets.

Proof.

Direct differentiation of $E[v_1^{inc}(x_1, r_1^{inc}, r_2^{inc})|\tilde{x}_1]$ yields

$$\frac{\partial}{\partial \epsilon_1}(E[v_1^{inc}(x_1, r_1^{inc})|\tilde{x}_1]) = -\frac{I_1}{2}\eta(\alpha - 1)(\frac{1}{1+\epsilon_1})^2 \left(\frac{\tilde{x}_1}{\overline{x}_1^{inc}}\right)^{\gamma_1}$$
$$\left(\frac{(\overline{\gamma} - 1)[1 - \frac{1}{2}\eta(\alpha - 1)(\frac{\epsilon_1}{1+\epsilon_1})] - \frac{1}{2}\eta(\alpha + 1) + (\overline{\gamma} - 1)(\gamma_1 - 1)(1 + \frac{1}{2}\eta(\frac{1+2\epsilon_1+\alpha}{1+\epsilon_1}))}{[(\overline{\gamma} - 1)[1 - \frac{1}{2}\eta(\alpha - 1)(\frac{\epsilon_1}{1+\epsilon_1})] - \frac{1}{2}\eta(\alpha + 1)]^2}\right) < 0.$$

Direct differentiation of $E[v_1^{agg}(x_1, r_2^{agg})|\tilde{x}_1]$ yields

$$\frac{\partial}{\partial \epsilon_1}(E[v_1^{agg}(x_1, r_2^{agg})|\tilde{x}_1]) = (\tilde{x}_1)^{\gamma_1} \left[\left(\frac{\partial \overline{x}_1^{agg}}{\partial \epsilon_1}\right)\left(\frac{1}{\overline{x}_1^{agg}}\right)^{\gamma_1+1}[\gamma_1 I_1 - \overline{x}_1^{agg}(\gamma_1 - 1)] \right.$$
$$\left. + \frac{k^{\gamma_1}}{2}\left(\frac{\partial}{\partial \epsilon_1}\left(\frac{\epsilon_1}{1+\epsilon_1}x_1^{agg}\right)\right)\frac{1}{2}\eta(\alpha + 1)\left(\frac{\gamma_1 - 1}{\overline{\gamma} - 1}\right)\left(\left(\frac{1}{\overline{x}_2^{aggH}}\right)^{\gamma_1} - \left(\frac{1}{\overline{x}_2^{aggL}}\right)^{\gamma_1}\right)\right] > 0.$$

The first term is positive whenever Equation (48) holds, since $\frac{\partial \overline{x}_1^{agg}}{\partial \epsilon_1} > 0$. The second term is positive since $\frac{\epsilon_1}{1+\epsilon_1}$ and x_1^{agg} increase in ϵ_1, and $\overline{x}_2^{aggH} < \overline{x}_2^{aggL}$. Since $\frac{\partial}{\partial \epsilon_1}(E[v_1^{inc}]) < 0$ and

39

$\frac{\partial}{\partial \epsilon_1}(E[v_1^{agg}]) > 0$, uniqueness and existence of $\tilde{\epsilon}_1$ are shown when $E[v_1^{inc}(\epsilon_1 = 0)] > E[v_1^{agg}(\epsilon_1 = 0)]$ and $\lim_{\epsilon_1 \to \infty} E[v_1^{inc}] < \lim_{\epsilon_1 \to \infty} E[v_1^{agg}]$.

First, $E[v_1^{inc}(\epsilon_1 = 0)] > E[v_1^{agg}(\epsilon_1 = 0)]$ whenever the agent does not wait longer than the ex-ante optimum when he sets incremental goals, even when uncertainty is greatest. Note that the first-best stage 1 threshold given an incremental goal is given by:

$$\overline{x}_1^{inc*} = \frac{\gamma_1 I_1}{(\gamma_1 - 1)[1 - \frac{1}{2}\eta(\alpha - 1)(\frac{\epsilon_1}{1+\epsilon_1})]},$$

where $\overline{x}_1^{inc*}$ is increasing in ϵ_1. In particular, $\overline{x}_1^{inc*} \geq (\frac{\gamma_1}{\gamma_1 - 1})I_1$, where $E[v_1^{agg}(\epsilon_1 = 0)]$ is maximized when $\overline{x}_1^{agg} = (\frac{\gamma_1}{\gamma_1 - 1})I_1$. Since $\overline{x}_1^{agg} < \overline{x}_1^{inc}$ and the second-stage stopping threshold is the same regardless of how he brackets when $\epsilon_1 = 0$, then $E[v_1^{inc}(\epsilon_1 = 0)] > E[v_1^{agg}(\epsilon_1 = 0)]$ whenever $\lim_{\epsilon_1 \to \infty} \overline{x}_1^{inc} \leq (\frac{\gamma_1}{\gamma_1 - 1})I_1$, which holds when the following condition is satisfied:

$$0 \leq \overline{\gamma} - \gamma_1 - \gamma_1[1 + \frac{1}{2}\eta(\alpha - 1)\gamma_1 + \frac{1}{2}\eta(\alpha + 1)]. \tag{48}$$

The right-hand side of Equation (48) is decreasing in η and γ, and increasing in $\overline{\gamma}$ when $\frac{1}{2}\eta(\alpha - 1) < \frac{1}{\gamma_1}$. That is, the agent's self-control is sufficiently poor (i.e., $\overline{\gamma}$ is sufficiently high) relative to his reference dependence and loss aversion that incremental goals are preferred to the aggregate goal in the absence of outcome uncertainty.

I now consider conditions such that $E[v_1^{agg}(x_1, r_2^{agg})|\tilde{x}_1] > E[v_1^{inc}(x_1, r_1^{inc})|\tilde{x}_1]$ when $\epsilon_1 \to \infty$. Since $E[v_2^{agg}] > E[v_2^{inc}]$ whenever $\epsilon_1 > 0$, then this is certainly satisfied whenever the value of the first-stage threshold under the aggregate goal exceeds the value of the first-stage threshold under incremental goals:

$$\lim_{\epsilon_1 \to \infty} E[v_1^{inc}(x_1, r_1^{inc}, r_2^{inc}) - v_1^{agg}(x_1, r_2^{agg})|\tilde{x}_1] = [\overline{x}_1^{inc}(1 - \frac{1}{2}\eta(\alpha - 1)) - I_1](\frac{\tilde{x}_1}{\overline{x}_1^{inc}})^{\gamma_1} - [\overline{x}_1^{agg} - I_1](\frac{\tilde{x}_1}{\overline{x}_1^{agg}})^{\gamma_1}. \tag{49}$$

Although we do not have an explicit expression for $\overline{x}_1^{agg}$, we know that it is bounded below: $\overline{x}_1^{agg} > (\frac{\overline{\gamma}}{\overline{\gamma} - 1})I_1 \equiv \overline{x}_h$. Then the second term in Equation (49) is bounded below by

$$[\overline{x}_h - I_1](\frac{\tilde{x}_1}{\overline{x}_h})^{\gamma_1}.$$

Thus, any (η, α) such that

$$0 \geq [\overline{x}_1^{inc}(1 - \frac{1}{2}\eta(\alpha - 1)) - I_1](\frac{\tilde{x}_1}{\overline{x}_1^{inc}})_1^{\gamma} - [\overline{x}_h - I_1](\frac{\tilde{x}_1}{\overline{x}_h})^{\gamma_1}$$

will also satisfy the condition that Equation (49) is negative. Since the left-hand term is strictly decreasing in α and η, then Equation (49) will be negative for any (η, α) sufficiently high that this inequality is satisfied. This inequality can be reduced to the following:

$$M(\eta, \alpha) \equiv (\overline{\gamma} - 1)^{\gamma_1 - 1} - (1 + \eta)\left((\overline{\gamma} - 1)[1 - \frac{1}{2}\eta(\alpha - 1)] - \frac{1}{2}\eta(\alpha + 1)\right)^{\gamma_1 - 1} \geq 0,$$

where $M(0, 1) = 0$. Also, consider the maximum permissible combination(s) $(\overline{\eta}, \overline{\alpha})$, which satisfies

$$(\overline{\gamma} - 1)[1 - \frac{1}{2}\overline{\eta}(\overline{\alpha} - 1)(\frac{\epsilon_1}{1 + \epsilon_1})] - \frac{1}{2}\overline{\eta}(\overline{\alpha} + 1) = 0,$$

and where $\overline{\eta} \leq \overline{\gamma} - 1$. Evaluating M at such a point, we must have that $M(\overline{\eta}, \overline{\alpha}) > 0$ since $\epsilon_1 \geq 0$. Furthermore, $\frac{\partial M}{\partial \alpha} > 0$ for $\eta > 0$. Differentiating M with respect to η, we have that M is increasing if

$$\overline{\gamma}(\alpha + 1) - 2\gamma_1(1 + \eta) - \gamma_1\overline{\gamma}(\alpha - 1)(1 + \eta) \leq 0, \tag{50}$$

and decreasing otherwise. Since the left-hand side of Equation (50) is decreasing in both η and α, we require (η, α) sufficiently large so that Equation (50) is satisfied in order for Equation (49) to be negative. Since $M(0, 1) = 0$, $M(\overline{\eta}, \overline{\alpha}) > 0$, $\frac{\partial M}{\partial \alpha} > 0$ when $\eta > 0$, and M is initially decreasing (and therefore negative) before increasing thereafter in η, then there exist some combinations (η', α') such that $M(\eta', \alpha') = 0$ and $M(\eta, \alpha) > 0$ for all $\eta' \leq \eta \leq \overline{\eta}$ and $\alpha' \leq \alpha \leq \overline{\alpha}$, where $\eta' > 0$ and $\alpha' > 1$.

To verify that there exist values of (η, α) that satisfy both (48) and $M(\eta, \alpha) > 0$, let us consider whether (50) can be satisfied when Equation (48) holds with equality. If not, then such values (η, α) do not exist and such a threshold $\tilde{\epsilon}_1$ does not exist. Suppose that α is sufficiently high (denoted $\hat{\alpha}$), given some η, that Equation (48) holds with equality. Does there exist some range of

41

η such that $M(\eta, \alpha) > 0$ is still satisfied? When Equation (48) holds with equality, M becomes

$$M(\eta, \hat{\alpha}) = (\overline{\gamma} - 1)^{\gamma_1 - 1} - (1 + \eta)(\frac{\overline{\gamma}(\gamma_1 - 1)}{\gamma_1})^{\gamma_1 - 1}. \tag{51}$$

Since the second term is strictly positive and less than $\overline{\gamma} - 1$ when $\eta = 0$, then $M(0, \hat{\alpha})$ is strictly positive whenever $\eta = 0$. We can also verify that $\frac{\partial M}{\partial \eta}|_{(0, \hat{\alpha})} < 0$ and $\frac{\partial M}{\partial \alpha} > 0$ for $\eta > 0$. Therefore, there exists some range of η such that $M(\eta, \hat{\alpha})$ is satisfied when (48) holds with equality. Thus, for $\overline{\gamma}$ sufficiently high, there exists some threshold $\tilde{\epsilon}_1$ such that $E[v_1^{agg}(x_1, r_2^{agg})|\tilde{x}_1] < E[v_1^{inc}(x_1, r_1^{inc}, r_2^{inc})|\tilde{x}_1]$ when $\epsilon_1 < \tilde{\epsilon}_1$, and $E[v_1^{agg}(x_1, r_2^{agg})|\tilde{x}_1] > E[v_1^{inc}(x_1, r_1^{inc}, r_2^{inc})|\tilde{x}_1]$ when $\epsilon_1 > \tilde{\epsilon}_1$. In particular, $\tilde{\epsilon}_1$ exists for any (η, α) that satisfy both (48) and $M(\eta, \alpha) > 0$.

■

10.6   Proof of Proposition 6

To find the optimal bracketing choice as ϵ_2 varies, I compare the expectations of v_1^{agg} and v_1^{inc}. If Equation (20) is satisfied, then the agent completes each stage sequentially regardless of how he brackets.

That there exist some combinations of $(\overline{\epsilon}_1, \overline{\epsilon}_2)$ such that $E[v_1^{agg}(x_1, r_2^{agg})|\tilde{x}_1] < E[v_1^{inc}(x_1, r_1^{inc}, r_2^{inc})|\tilde{x}_1]$ when $\overline{\gamma}$ is sufficiently high follows from the Proof of Proposition 5.

Direct differentiation of $E[v_1^{agg}(x_1, r_2^{agg}) - v_1^{inc}(x_1, r_1^{inc}, r_2^{inc})]$ yields

$$\frac{\partial}{\partial \epsilon_2}(E[v_1^{agg}(x_1, r_2^{agg}) - v_1^{inc}(x_1, r_1^{inc}, r_2^{inc})]) =$$

$$(\tilde{x}_1)^{\gamma_1}\left(\left(\frac{\partial \overline{x}_1^{agg}}{\partial \epsilon_2}\right)\left(\frac{1}{\overline{x}_1^{agg}}\right)^{\gamma_1 + 1}(\gamma_1 I_1 - \overline{x}_1^{agg}(\gamma_1 - 1))\right.$$

$$+ \frac{k^{\gamma_1}}{\overline{\gamma}}\left[-\frac{1}{2}\eta(\alpha - 1)(\frac{1}{1 + \epsilon_2})^2\left(\frac{1}{2}\left(\left(\frac{1}{\overline{x}_2^{aggH}}\right)^{\gamma_1 - 1} + \left(\frac{1}{\overline{x}_2^{aggL}}\right)^{\gamma_1 - 1}\right) - \left(\frac{1}{\overline{x}_2^{inc}}\right)^{\gamma_1 - 1}\right)\right.$$

$$\left.\left. - (\gamma_1 - 1)\left(1 - \frac{1}{2}\eta(\alpha - 1)(\frac{\epsilon_2}{1 + \epsilon_2}) + \frac{1}{2}\eta(\alpha - 1)\right)(A(\epsilon_2))\right]\right).$$

Under aggregate goals, $\frac{\partial \overline{x}_1^{agg}}{\partial \epsilon_2} < 0$, so the option value of stage 1 decreases with ϵ_2 (i.e., the first

term is negative) when Equation (49) is satisfied. The second term is negative, since the function $(1/x)^c$ where $c > 0$ is convex. I now show that the third term is also negative since $A(\epsilon_2)$ is positive.

$$A(\epsilon_2) = \left(\frac{\frac{1}{2}\eta(\alpha-1)(\frac{1}{1+\epsilon_2})^2\overline{\gamma}(\overline{\gamma}-1)I_2}{[(\overline{\gamma}-1)(1-\frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_2}{1+\epsilon_2}))-\frac{1}{2}\eta(\alpha+1)]^2} \right) \left(\frac{1}{2}\left(\left(\frac{1}{\overline{x}_2^{aggH}}\right)^{\gamma_1} + \left(\frac{1}{\overline{x}_2^{aggL}}\right)^{\gamma_1} \right) - \left(\frac{1}{\overline{x}_2^{inc}}\right)^{\gamma_1} \right)$$
$$-\frac{1}{2}\left(\frac{\frac{1}{2}\eta(\alpha+1)\frac{1}{2}\eta(\alpha-1)(\frac{1}{1+\epsilon_2})^2\overline{\gamma}(\frac{\epsilon_1}{1+\epsilon_1})x_1^{agg}}{(\overline{\gamma}-1)[1-\frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_2}{1+\epsilon_2})+\frac{1}{2}\eta(\alpha+1)]^2} + \frac{\frac{1}{2}\overline{\gamma}\eta(\alpha+1)(\frac{\epsilon_1}{1+\epsilon_1})}{(\overline{\gamma}-1)[1-\frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_2}{1+\epsilon_2})+\frac{1}{2}\eta(\alpha+1)]}(\frac{\partial x_1^{agg}}{\partial \epsilon_2}) \right)$$
$$\left(\left(\frac{1}{\overline{x}_2^{aggH}}\right)^{\gamma_1} - \left(\frac{1}{\overline{x}_2^{aggL}}\right)^{\gamma_1} \right).$$

The first term is positive. The second term is bounded below, since $\frac{\partial x_1^{agg}}{\partial \epsilon_2} < 0$, yielding

$$A(\epsilon_2) \geq \left(\frac{\frac{1}{2}\eta(\alpha-1)(\frac{1}{1+\epsilon_2})^2\overline{\gamma}(\overline{\gamma}-1)I_2}{[(\overline{\gamma}-1)(1-\frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_2}{1+\epsilon_2}))-\frac{1}{2}\eta(\alpha+1)]^2} \right) \left(\frac{1}{2}\left(\left(\frac{1}{\overline{x}_2^{aggH}}\right)^{\gamma_1} + \left(\frac{1}{\overline{x}_2^{aggL}}\right)^{\gamma_1} \right) - \left(\frac{1}{\overline{x}_2^{inc}}\right)^{\gamma_1} \right)$$
$$-\frac{1}{2}\left(\frac{\frac{1}{2}\eta(\alpha+1)\frac{1}{2}\eta(\alpha-1)(\frac{1}{1+\epsilon_2})^2\overline{\gamma}(\frac{\epsilon_1}{1+\epsilon_1})x_1^{agg}}{(\overline{\gamma}-1)[1-\frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_2}{1+\epsilon_2})+\frac{1}{2}\eta(\alpha+1)]^2} \right) \left(\left(\frac{1}{\overline{x}_2^{aggH}}\right)^{\gamma_1} - \left(\frac{1}{\overline{x}_2^{aggL}}\right)^{\gamma_1} \right).$$
$$=\frac{1}{2}\eta(\alpha-1)(\frac{1}{1+\epsilon_2})^2 \left[\left(\frac{x_2^{inc}(\overline{\gamma}-1)}{(\overline{\gamma}-1)(1-\frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_2}{1+\epsilon_2}))-\frac{1}{2}\eta(\alpha+1)} \right) \left(\frac{1}{2}\left(\left(\frac{1}{\overline{x}_2^{aggH}}\right)^{\gamma_1} + \left(\frac{1}{\overline{x}_2^{aggL}}\right)^{\gamma_1} \right) - \left(\frac{1}{\overline{x}_2^{inc}}\right)^{\gamma_1} \right) \right.$$
$$\left. -\frac{1}{2}\left(\frac{H(\epsilon_2)x_1^{agg}}{1-\frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_2}{1+\epsilon_2})+\frac{1}{2}\eta(\alpha+1)} \right) \left(\left(\frac{1}{\overline{x}_2^{aggH}}\right)^{\gamma_1} - \left(\frac{1}{\overline{x}_2^{aggL}}\right)^{\gamma_1} \right) \right],$$

where the equality follows from $x_2^{aggH} = x_2^{inc} - H(\epsilon_2)x_1^{agg}$ and $x_2^{aggL} = x_2^{inc} + H(\epsilon_2)x_1^{agg}$, where $H(\epsilon_2) = \frac{\frac{1}{2}\overline{\gamma}\eta(\alpha+1)(\frac{\epsilon_1}{1+\epsilon_1})}{(\overline{\gamma}-1)[1-\frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_2}{1+\epsilon_2})+\frac{1}{2}\eta(\alpha+1)]}$. Combining $x_2^{aggH} = x_2^{inc} - H(\epsilon_2)x_1^{agg}$ and $x_2^{aggL} = x_2^{inc} + H(\epsilon_2)x_1^{agg}$ with the fact that

$$\frac{1}{2}\left(\left(\frac{1}{\overline{x}_2^{aggH}}\right)^{\gamma_1-1} + \left(\frac{1}{\overline{x}_2^{aggL}}\right)^{\gamma_1-1} \right) - \left(\frac{1}{\overline{x}_2^{inc}}\right)^{\gamma_1-1} \geq 0$$

implies that

$$\frac{1}{2}H(\epsilon_2)x_1^{agg}\left(\left(\frac{1}{\overline{x}_2^{aggH}}\right)^{\gamma_1} - \left(\frac{1}{\overline{x}_2^{aggL}}\right)^{\gamma_1} \right) \leq x_2^{inc}\left(\frac{1}{2}\left(\left(\frac{1}{\overline{x}_2^{aggH}}\right)^{\gamma_1-1} + \left(\frac{1}{\overline{x}_2^{aggL}}\right)^{\gamma_1-1} \right) - \left(\frac{1}{\overline{x}_2^{inc}}\right)^{\gamma_1-1} \right).$$

Moreover,

$$\frac{1}{1 - \frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_2}{1+\epsilon_2}) + \frac{1}{2}\eta(\alpha+1)} < \frac{\overline{\gamma}-1}{(\overline{\gamma}-1)(1 - \frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_2}{1+\epsilon_2})) - \frac{1}{2}\eta(\alpha+1)}.$$

Thus, $A(\epsilon_2) \geq 0$ so $\frac{\partial}{\partial\epsilon_2}(E[v_1^{agg}(x_1, r_2^{agg}) - v_1^{inc}(x_1, r_1^{inc}, r_2^{inc})]) < 0$.

10.7   Proof of Proposition 7

To show Proposition 7, it is sufficient to show that $\frac{\partial}{\partial\alpha}(E[v_1^{inc}(x_1)|\tilde{x}_1])|_{\alpha=1} > 0$ for some

combination of $\eta \geq 0$ and $\epsilon_i > 0$ for $i = 1, 2$.

If the agent sets incremental goals,

$$\frac{\partial}{\partial\alpha}(E(v_1^{inc}(x_1, r_1^{inc}, r_2^{inc}))) = \left(\frac{\tilde{x}_1}{\overline{x}_1^{inc}}\right)^{\gamma_1} (\frac{\frac{1}{2}\eta\overline{x}_1^{inc}}{\overline{\gamma}}) \left(\frac{1}{1+\epsilon_1} - \frac{\gamma_1[(\overline{\gamma}-1)(\frac{\epsilon_1}{1+\epsilon_1})+1][1 - \frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_1}{1+\epsilon_1} + \frac{1}{2}\eta(\alpha+1)]}{(\overline{\gamma}-1)(1 - \frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_1}{1+\epsilon_1})) - \frac{1}{2}\eta(\alpha+1)}\right)$$
$$+ \left(\frac{k\tilde{x}_1}{\overline{x}_2^{inc}}\right)^{\gamma_1} (\frac{\frac{1}{2}\eta\overline{x}_2^{inc}}{\overline{\gamma}}) \left(\frac{1}{1+\epsilon_2} - \frac{\gamma_1[(\overline{\gamma}-1)(\frac{\epsilon_2}{1+\epsilon_2})+1][1 - \frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_2}{1+\epsilon_2} + \frac{1}{2}\eta(\alpha+1)]}{(\overline{\gamma}-1)(1 - \frac{1}{2}\eta(\alpha-1)(\frac{\epsilon_2}{1+\epsilon_2})) - \frac{1}{2}\eta(\alpha+1)}\right)$$

$$(52)$$

Evaluating at $\alpha = 1$, a necessary condition for $\frac{\partial}{\partial\alpha}(E[v_1^{inc}(x_1)|\tilde{x}_1])|_{\alpha=1} > 0$ is

$$\eta < \frac{\overline{\gamma} - \gamma_1}{\gamma_1},$$

where the right-hand side is clearly non-negative and only equals zero when the agent is not

impulsive. In addition, a sufficient[23] condition for $\frac{\partial}{\partial\alpha}(E[v_1^{inc}(x_1)|\tilde{x}_1])|_{\alpha=1} > 0$ is that ex-ante

outcome uncertainty must be sufficiently small in each stage:

$$\epsilon_i < \frac{\overline{\gamma} - \gamma_1(1+\eta)}{\overline{\gamma}(\gamma_1 - 1)(1+\eta)} \quad \text{for } i = 1, 2.$$

---

[23]This condition is stricter than necessary and guarantees that both terms of Equation (52) are positive. Clearly, the necessary condition is that this upper bound be satisfied for at least one stage.

If the agent sets an aggregate goal, the analogous result holds.

$$\frac{\partial}{\partial \alpha}(E[v_1^{agg}(x_1, r_2^{agg})]) =$$

$$(\tilde{x}_1)^{\gamma_1} \left(\left(\frac{\partial \overline{x}_1^{agg}}{\partial \epsilon_2} \right) \left(\frac{1}{\overline{x}_1^{agg}} \right)^{\gamma_1+1} (\gamma_1 I_1 - \overline{x}_1^{agg}(\gamma_1 - 1)) \right)$$

$$+ \frac{k^{\gamma_1}}{\overline{\gamma}} \left[\left(\frac{\frac{1}{2}\eta \overline{x}_2^{inc}}{2} \left(\left(\frac{1}{x_2^{aggH}} \right)^{\gamma_1} + \left(\frac{1}{x_2^{aggL}} \right)^{\gamma_1} \right) \right) B(\alpha) \right.$$

$$+ D(\alpha) \left(\frac{H(\epsilon_2)\overline{x}_1^{agg}}{2} \left(\left(\frac{1}{x_2^{aggH}} \right)^{\gamma_1} - \left(\frac{1}{x_2^{aggL}} \right)^{\gamma_1} \right) \right)$$

$$\left. + \left(\frac{H(\epsilon_2)}{2} \left(\left(\frac{1}{x_2^{aggH}} \right)^{\gamma_1} - \left(\frac{1}{x_2^{aggL}} \right)^{\gamma_1} \right) \right) (\gamma_1 - 1)(1 - \frac{1}{2}\eta(\alpha - 1)(\frac{\epsilon_2}{1+\epsilon_2}) + \frac{1}{2}\eta(\alpha + 1))(\frac{\partial \overline{x}_1^{agg}}{\partial \alpha}) \right]$$

where

$$B(\alpha) = \frac{1}{1 + \epsilon_2} - \frac{((\overline{\gamma} - 1)(\frac{\epsilon_2}{1+\epsilon_2}) + 1)(\gamma_1 - 1)(1 - \frac{1}{2}\eta(\alpha - 1)(\frac{\epsilon_2}{1+\epsilon_2}) + \frac{1}{2}\eta(\alpha + 1))}{(\overline{\gamma} - 1)(1 - \frac{1}{2}\eta(\alpha - 1)(\frac{\epsilon_2}{1+\epsilon_2})) - \frac{1}{2}\eta(\alpha + 1)}$$

$$D(\alpha) = -\frac{1}{2}\eta(\frac{1}{1 + \epsilon_2}) + (\gamma_1 - 1)(1 - \frac{1}{2}\eta(\alpha - 1)(\frac{\epsilon_2}{1 + \epsilon_2}) + \frac{1}{2}\eta(\alpha + 1)(\frac{\epsilon_2}{1 + \epsilon_2}))$$

Evaluating at $\alpha = 1$, a sufficient condition for $\frac{\partial}{\partial \alpha}(E[v_2^{agg}(x_1)|\tilde{x}_1])|_{\alpha=1} > 0$ is

$$\eta < \frac{\overline{\gamma} - \gamma_1}{\gamma_1},$$

and

$$\epsilon_2 < \frac{\overline{\gamma} - \gamma_1(1 + \eta)}{\overline{\gamma}(\gamma_1 - 1)(1 + \eta)}.$$

The first term of $\frac{\partial}{\partial \alpha}(E[v_1^{agg}(x_1, r_2^{agg})])|_{\alpha=1}$ is positive whenever $\frac{\partial x_1^{agg}}{\partial \alpha} > 0$, which is guaranteed[24]

whenever

$$I_1 k(\overline{\gamma} - 1 - \eta)^2 - I_2 \gamma_2 \eta(\overline{\gamma} - 1) > 0 \tag{53}$$

and

$$\epsilon_2 < \frac{I_1 k(\overline{\gamma} - 1 - \eta)^2 - I_2 \gamma_2 \eta(\overline{\gamma} - 1)}{I_2 \gamma_2 \eta \overline{\gamma}(\overline{\gamma} - 1) - I_1 k(\overline{\gamma} - 1 - \eta)^2}. \tag{54}$$

Note that Equation (53) is decreasing in η, and is positive when $\eta = 0$, so there exists a range of

---

[24]These conditions are stronger than required, to establish existence.

$\eta > 0$ sufficiently small that satisfies this requirement, and likewise ϵ_2 in Equation (54). Thus, $\frac{\partial}{\partial \alpha}(E[v_1^{agg}(x_1, r_2^{agg})])|_{\alpha=1} > 0$ whenever η and ϵ_2 are sufficiently small, implying that some degree of loss aversion is beneficial to the agent from an ex-ante perspective.

References

Abeler, Johannes, Armin Falk, Lorenz Goette, and David Huffman, "Reference Points and Effort Provision," *American Economic Review*, April 2011, *101* (2), 470–492.

Ali, Mukhtar M, "Probability and utility estimates for racetrack bettors," *The Journal of Political Economy*, 1977, pp. 803–815.

Allen, Eric J., Patricia M. Dechow, Devin G. Pope, and George Wu, "Reference-Dependent Preferences: Evidence from Marathon Runners," July 2014. Working Paper.

Arkes, Hal R and Catherine Blumer, "The psychology of sunk cost," *Organizational behavior and human decision processes*, 1985, *35* (1), 124–140.

Bandura, Albert, "Self-Regulation of Motivation and Action Through Internal Standards and Goal Systems," in Lawrence A. Pervin, ed., *Goal Concepts in Personality and Social Psychology*, Lawrence Erlbaum Associates, Publishers, 1989, pp. 19–63.

___ **and Karen M. Simon**, "The role of proximal intentions in self-regulation of refractory behavior," *Cognitive Therapy and Research*, September 1977, *1* (3), 177–193.

Barberis, Nicholas C. and Wei Xiong, "Realization Utility," *Journal of Financial Economics*, May 2012, *104* (2), 251–271.

Barberis, Nicholas, Ming Huang, and Richard H. Thaler, "Individual Preferences, Monetary Gambles, and Stock Market Participation: A Case for Narrow Framing," *The American Economic Review*, September 2006, *96* (4), 1069–1090.

___ , ___ , **and Tano Santos**, "Prospect Theory and Asset Prices," *Quarterly Journal of Economics*, 2001, *116* (1), 1–53.

Bénabou, Roland and Jean Tirole, "Willpower and Personal Rules," *Journal of Political Economy*, August 2004, *112*, 848–886.

Benartzi, Shlomo and Richard H. Thaler, "Myopic Loss Aversion and the Equity Premium Puzzle," *The Quarterly Journal of Economics*, 1995, *110* (1), 73–92.

Brocas, Isabelle and Juan D. Carrillo, "A theory of haste," *Journal of Economic Behavior & Organization*, January 2005, *56* (1), 1–23.

Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler, "Labor Supply of New York City Cabdrivers: One Day at a Time," *The Quarterly Journal of Economics*, May 1997, *112* (2), 407–41.

Crawford, Vincent P. and Juanjuan Meng, "New York City Cabdrivers' Labor Supply Revisited: Reference-Dependence Preferences with Rational-Expectations Targets for Hours and Income," UCSD Discussion Paper 2008-03, University of California, San Diego July 2008.

Dillenberger, David, "Preferences for One-Shot Resolution of Uncertainty and Allais-Type Behavior," *Econometrica*, November 2010, *78* (6), 1973–2004.

Ericson, Keith M. and Andreas Fuster, "Expectations as Endowments: Evidence on Reference-Dependent Preferences from Exchange and Valuation Experiments," *Quarterly Journal of Economics*, November 2011, *126* (4), 1879–1907.

Farber, Henry S., "Is Tomorrow Another Day? The Labor Supply of New York City Cabdrivers," *Journal of Political Economy*, 2005, *113* (1), 46–82.

——, "Reference-Dependent Preferences and Labor Supply: The Case of New York City Taxi Drivers," *American Economic Review*, June 2008, *98* (3), 1069–1082.

Fishbach, Ayelet and Ravi Dhar, "Goals as excuses or guides: The liberating effect of perceived goal progress on choice," *Journal of Consumer Research*, 2005, *32* (3), 370–377.

Fulford, Daniel, Sheri L. Johnson, Maria M. Llabre, and Charles S. Carver, "Pushing and coasting in dynamic goal pursuit: coasting is attenuated in bipolar disorder," *Psychological Science*, 2010, *21* (7), 1021–1027.

Gneezy, Uri and Jan Potters, "An Experiment on Risk Taking and Evaluation Periods," *Quarterly Journal of Economics*, 1997, *112* (2), 631–645.

Goette, Lorenz F. and David Huffman, "Affect as a Source of Motivation in the Workplace: A New Model of Labor Supply, and New Field Evidence on Income Targeting and the Goal Gradient," IZA Discussion Paper Series 1890, IZA December 2005.

Grenadier, Steven and Neng Wang, "Investment Under Uncertainty and Time-Inconsistent Preferences," *Journal of Financial Economics*, 2007, *84*, 2–39.

Gul, Faruk, "A Theory of Disappointment Aversion," *Econometrica*, 1991, *59* (3), 667–686.

Harris, Christopher and David Laibson, "Instantaneous Gratification," *Quarterly Journal of Economics*, 2013, *128* (1), 205–248.

Heath, Chip and Jack B. Soll, "Mental Budgeting and Consumer Decisions," *The Journal of Consumer Research*, 1996, *23* (1), 40–52.

⎯⎯ , **Richard P. Larrick, and George Wu**, "Goals as Reference Points," *Cognitive Psychology*, 1999, *38*, 79–109.

Herweg, Fabian and Klaus M Schmidt, "Loss aversion and inefficient renegotiation," *The Review of Economic Studies*, 2015, *82* (1), 297–332.

Hsiaw, Alice, "Goal-Setting and Self-Control," *Journal of Economic Theory*, March 2013, *148* (2), 601–626.

Imas, Alex, "The Realization Effect: Risk-Taking After Realized Versus Paper Losses," August 2015.

⎯⎯ , **Sally Sadoff, and Anya Samek**, "Do People Anticipate Loss Aversion?," *Management Science*, forthcoming.

Kahneman, Daniel and Amos Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 1979, *47* (2), 263–291.

Kőszegi, Botond and Matthew Rabin, "A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics*, November 2006, *121* (4), 1133–1165.

⎯⎯ **and** ⎯⎯ , "Reference-Dependent Consumption Plans," *American Economic Review*, June 2009, *99* (3), 909–936.

Kirschenbaum, Daniel, "Proximity and specificity of planning: A position paper," *Cognitive Therapy and Research*, October 1985, *9* (5), 486–506.

Koch, Alexander K. and Julia Nafziger, "Self-Regulation through Goal Setting," *Scandinavian Journal of Economics*, 2011, *113* (1), 212–227.

_____ **and** _____ , "Goals and Mental Accounting," August 2014. Working Paper.

_____ , _____ , **Anton Suvorov, and Jeroen van de Ven**, "Self-Rewards and Personal Motivation," *European Economic Review*, May 2014, *68*, 151–167.

Koo, Minjung and Ayelet Fishbach, "Dynamics of self-regulation: How (un) accomplished goal actions affect motivation.," *Journal of personality and social psychology*, 2008, *94* (2), 183.

Latham, Gary P. and Edwin A. Locke, "Self-Regulation through Goal Setting," *Organizational Behavior and Human Decision Processes*, 1991, *50* (2), 212–247.

Locke, Edwin A. and Gary P. Latham, "Building a practically useful theory of goal setting and task motivation: A 35-year odyssey.," *American Psychologist*, September 2002, *57* (9), 705–717.

Loewenstein, George, "Because It Is There: The Challenge of Mountaineering . . . for Utility Theory," *Kyklos*, 2007, *52* (3), 315–343.

Louro, Maria J, Rik Pieters, and Marcel Zeelenberg, "Dynamics of multiple-goal pursuit," *Journal of Personality and Social Psychology*, 2007, *93* (2), 174.

McGlothlin, William H, "Stability of choices among uncertain alternatives," *The American Journal of Psychology*, 1956, pp. 604–615.

Miao, Jianjun, "Option exercise with temptation," *Economic Theory*, March 2008, *34* (3), 473–501.

Odean, Terrance, "Are Investors Reluctant to Realize Their Losses?," *The Journal of Finance*, 1998, *53* (5), 1775–1798.

Palacios-Huerta, Ignacio, "The Aversion to the Sequential Resolution of Uncertainty," *Journal of Risk and Uncertainty*, 1999, *18*, 249–269.

Pope, Devin G. and Maurice E. Schweitzer, "Is Tiger Woods Loss Averse? Persistent Bias in the Face of Experience, Competition, and High Stakes," *American Economic Review*, 2011, *101* (1), 129–159.

Rabin, Matthew and Georg Weizsäcker, "Narrow bracketing and dominated choices," *American Economic Review*, 2009, pp. 1508–1543.

Read, Daniel, George Loewenstein, and Matthew Rabin, "Choice Bracketing," *Journal of Risk and Uncertainty*, December 1999, *19* (1), 171–197.

Sprenger, Charles, "An Endowment Effect for Risk: Experimental Tests of Stochastic Reference Points," *Journal of Political Economy*, in press.

Suvorov, Anton and Jeroen van de Ven, "Goal Setting as a Self-Regulation Mechanism," October 2008. Working Paper.

Thaler, Richard, "Toward a positive theory of consumer choice," *Journal of Economic Behavior & Organization*, 1980, *1* (1), 39–60.

Thaler, Richard H., "Mental accounting matters," *Journal of Behavioral Decision Making*, 1999, *12* (3), 183–206.

Tversky, Amos and Daniel Kahneman, "The Framing of Decisions and the Psychology of Choice," *Science*, 1981, *211* (4481), 453–458.