Thinking Ahead
Part I: The Decision Problem

by

Patrick Bolton and Antoine Faure-Grimaud

[PRELIMINARY AND INCOMPLETE]

June 2004

Abstract: We propose a model of bounded rationality based on time-costs of deliberating current and future decisions. We model an individual decision maker's thinking process as a thought-experiment that takes time and let the decision maker 'think ahead' about future decision problems in yet unrealized states of nature. By formulating an intertemporal, state-contingent, planning problem, which may involve costly deliberation in every state of nature, and by letting the decision-maker deliberate ahead of the realization of a state, we attempt to capture the basic idea that individuals generally do not think through a complete action-plan. Instead, individuals prioritize their thinking and leave deliberations on less important decisions to the time or event when they arise.

Key Words: Decision Problem, Thinking Costs.

JEL Classification:

*We thank Oliver Hart, Roman Inderst, Nobuhiro Kiyotaki, Thomas Mariotti, John Moore, Soenje Reiche, Jean Tirole, David Webb and seminar participants at Amsterdam, Tilburg, Florence, Royal Holloway and the London School of Economics for helpful comments.

**Princeton University.

***Department of Economics and FMG at the London School of Economics, and CEPR.
This work was started when visiting Princeton University, he thanks this institution for its hospitality.

1 Introduction

This paper proposes a simple and tractable model of bounded rationality based on time-costs of deliberating current and future decisions. We introduce a deliberation technology based on the classical two-armed bandit problem (Gittins and Jones, 1974 and Rothschild, 1974) and model an individual decision maker's thinking process as a thought-experiment that takes time.

The basic idea we explore is that a boundedly rational decision-maker thinks through a simple decision, such as of which of two actions to take, by weighing in her mind the costs and benefits associated with each possible action through thought-experiments which take time. Eventually, following enough "thought-experimentation" the decision-maker (DM) becomes sufficiently confident about which action is best and takes a decision.

Although we rely on the powerful multi-armed bandit framework, we depart from the classical bandit problem in a fundamental way by introducing the notion of 'thinking ahead' about future decision problems in yet unrealized states of nature. By formulating an intertemporal, state-contingent, planning problem, which may involve costly deliberation in every state of nature, and by letting the decision-maker deliberate ahead of the realization of a state, we attempt to capture the basic idea that individuals generally do not think through a complete action-plan. Instead, individuals prioritize their thinking and first think through the decisions that seem most important to them. They also generally leave deliberations on less important decisions to the time or event when they arise.

Such behavior is understandable if one has in mind deliberation costs but, as Rubinstein (1998) has noted, it is irreconcileable with the textbook model of the rational DM with no thinking costs:

" In situations in which the decision maker anticipates obtaining information before taking an action, one can distinguish between two timings of decision making: 1. Ex ante decision making. A decision is made before the information is revealed, and it is contingent on the content of the information to be received. 2. Ex post decision making.

The decision maker waits until the information is received and then makes a decision. In standard decision problems, with fully rational decision makers, this distinction does not make any difference." [Rubinstein 1998, page 52]

There are at least three reasons for developing a model with costly deliberation. First, there is the obvious reason that the behavior captured by such a model is more descriptive of how individuals make decisions in reality. Second, as we shall explain, such a model can provide new foundations for two popular behavioral hypotheses: "satisficing" behavior (Simon, 1955 and Radner, 1975) and decision-making under time pressure that takes the form of "putting out fires" (Radner and Rothschild, 1975). The main motivation of the current paper is to show how a model of decision-making with costly deliberation can explain both "satisficing" behavior and a prioritization of decision problems akin to "putting out fires". Third, a model with costly deliberation can also provide a tractable framework to analyze long-term contracting between boundedly rational agents. We analyze this contracting problem in our companion paper on "satisficing contracts" (Bolton and Faure-Grimaud, 2004).

Our initial objective was mainly to formulate a tractable framework of contracting between boundedly rational agents in response to Oliver Hart's observation that: "*In reality, a great deal of contractual incompleteness is undoubtedly linked to the inability of parties not only to contract very carefully about the future, but also to think very carefully about the utility consequences of their actions. It would therefore be highly desirable to relax the assumption that parties are unboundedly rational.*" [Hart, 1995, p. 81] However, having formulated our deliberation technology for boundedly rational agents we found that the decision problem is of sufficient independent interest to be discussed in a separate paper.

In our model the decision-maker starts with some prior estimate of the payoff associated with each possible action choice in every state of nature. She can either take her prior as her best guess of her final payoff and determine her optimal action-plan associated with that prior, or she can 'think' further and run an experiment on one action. This experiment will allow her to update her estimate of the payoff associated with that action and possibly to improve her action choice.

At each point in time, DM, thus, faces the basic problem whether to explore further the payoff associated with a particular action, search further other parts of her optimization problem, or make a decision based on what she has learnt so far. Since 'thinking ahead' takes time, DM will generally decide to leave some future decisions that she is only likely to face in rare or distant states of nature to be thought through later.

As is well understood, thinking ahead and resolving future decisions allows a DM to make better current decisions only when current actions are partially or completely irreversible. Our decision problem, thus, involves irreversible actions, and thereby introduces a bias towards deliberation on future actions ahead of the realization of future states of nature.

To understand the underlying logic of our decision problem it is helpful to draw a parallel with the problem of irreversible investment under uncertainty involving a 'real option' (see Henry, 1974 and Dixit and Pyndick, 1996). In this problem the rational DM may choose to delay investment, even if it is known to generate an expected positive net present value, in an effort to maximize the value of the 'real option' of avoiding making investments that ex post turn out to have a negative net present value. If one interprets information acquisition through delayed investment as a form of deliberation on future actions, one is able to infer from the 'real options' literature that a boundedly rational DM with time-costs of deliberation is likely to behave in a similar way. That is, the boundedly rational DM will postpone taking a current irreversible action until she is sufficiently confident that this action yields the highest payoff. The interesting observation here is not so much that the boundedly rational DM will delay taking an action, but that she will eventually decide to act even if she has not fully resolved her entire future action-plan, provided that she is sufficiently confident that she is making the right choice.

As helpful as this parallel is in understanding the basic logic of our problem, it is not a perfect analogy. To reveal how our problem differs from a reformulation of a real options problem we specialize our framework to a decision problem where there is no real option value and there is no doubt that a particular current action (investment) is clearly preferable. Nevertheless in this problem a boundedly rational DM will generally choose to delay investment and think ahead about future decisions in some if not all future states of nature.

What is the reason for this delay if no option value is present? The answer is that, by thinking ahead about future decisions the boundedly rational DM can reduce the time-lag between the realization of a state of nature and the time when the DM takes a decision in that state. By reducing this time-lag the boundedly rational DM is able to reduce the overall expected lag between the time she makes a costly investment decision and the time when she recoups the returns from her investment. Because the DM discounts future payoffs, reducing this time-lag raises her payoff. In other words, our framework builds on the idea that the benefit of 'thinking ahead' is to be able to react more promptly to new events, but the cost is delayed current decisions. This is the main novel mechanism we study in this paper.

How does this framework provide a new foundation for satisficing behavior? In general it is optimal for the boundedly rational DM to engage in what we refer to as 'step-by-step' thinking. This involves singling out a subset of future decision problems and think these through first. If the thought-experiments on these problems reveal that the payoff from investing is appreciably higher than the DM initially thought then the DM will decide that she is satisfied with what she found and will choose to invest without engaging into further thinking on other future decisions she has not yet thought about. If, on the other hand, the thought-experiments reveal that the payoff from investing is no higher and possibly lower than initially thought then the DM will continue thinking about other decision problems in yet unexplored future states of nature. In other words, the boundedly rational DM will generally refrain from fully determining the optimal future action-plan and will settle on an incomplete plan which provides a satisfactory expected payoff. Note that, in our framework the satisficing threshold is determined endogenously, as the solution of an optimal stopping problem. Thus, our framework can address a basic criticism that has been voiced against the satisficing hypothesis, namely that the threshold is imposed exogenously.

In what way does the boundedly rational DM behave as if she were "putting out fires"? We show that quite generally the boundedly rational DM prioritizes her thinking by first choosing to think about the most important and urgent problems. It is in this sense that she behaves as if she were putting out fires. The original formulation of this behavioral hypothesis by Radner and Rothschild considered only very extreme situations, where the DM had no choice but to put out fires. Our framework highlights that the general idea underlying the notion of putting out fires, that a boundedly rational DM prioritizes her thinking by focusing first on the most important problems or those most likely to arise, extends far beyond the extreme high stress situation considered by Radner and Rothschild.

A number of other major insights emerge from our analysis. Thus, when the number of future decisions to think about is large so that the overall planning problem is overwhelming then it is best to act immediately by guessing which action is best, and to postpone all the thinking to the time when the problems arise. This result is quite intuitive and provides one answer to the well known 'how-to-decide-how-to-decide' paradox one faces when one introduces deliberation costs into a rational decision-making problem (see Lipman, 1995). In the presence of deliberation costs the DM faces a larger decision problem, as she has to decide how to economize on deliberation costs. Presumably this larger decision problem itself requires costly deliberation, which ought to be economized, etc. We suggest here that one way of resolving this paradox is to have the DM act on a best guess without any deliberation when the problem becomes overwhelming.

In contrast, when the number of future decision problems (or states of nature) is more manageable then step-by-step thinking is generally optimal. For an even lower number of future problems complete planning is optimal (in particular, when there is only one state of nature, and therefore only one future problem to think about, then thinking ahead and complete planning is always optimal).

There is obviously a large and rapidly growing literature on bounded rationality and the ideas we have discussed have been explored by others. The literature on bounded rationality that is most closely related to ours is the one on costly deliberation (see Conlisk, 1996), and within this sub-literature our work is closest in spirit to Conlisk (198.), who formulates a related problem where a boundedly rational agent thinks ahead about state-contingent problems. The main difference with Conlisk (198.) is that we formulate a different deliberation technology, with thinking modeled as thought experimentation and thinking costs taking the form of time thinking costs.

The remainder of our paper is organized as follows.

2 A Simple Model of "Bandit" Rationality

2.1 The general framework

The general decision problem we consider is an infinite horizon, discrete time problem involving an initial decision on which of n actions to take in the action set $A_0 = \{a_1, ..., a_n\}$ and future decision problems which are contingent on the initial action choice a_i and on the realized state of nature $\theta_j \in \Theta$, where Θ denotes the finite set of states $\Theta = \{\theta_1, ..., \theta_N\}$. The initial action may be taken at any time $t \geq 0$ and when an action is chosen at time t a state of nature is realized at time $t + \Delta_i$, where $\Delta_i \geq 1$. When some action a_i has been chosen the DM receives an immediate payoff $\omega(a_i)$. In addition, when action a_i has been chosen and a state θ_j is realized the DM must choose another action a_{ijk} from another (finite) set of actions $A_{ij} = \{a_{ij1}, ..., a_{ijm_{ij}}\}$, where m_{ij} denotes the number of actions available in state θ_j when DM has picked the initial action a_j. Again, the DM may take this action at any time $\tau \geq t + \Delta_i$. When the action a_{ijk} following the realization of the state θ_j has been taken, DM obtains another payoff $\pi(a_i, a_{ijk}, \theta_j)$. Future payoffs are discounted and the discount factor is given by $\delta < 1$. Thus, the present discounted payoff when DM chooses action a_i in period t and action a_{ijk} in period $\tau \geq t + \Delta_i$ is given by:

$$\delta^t \omega(a_i) + \delta^{t+\tau} \pi(a_i, a_{ijk}, \theta_j).$$

The probability that state θ_j arises when action a_i has been chosen is given by $\mu_{ij} \geq 0$. Thus, the expected present discounted payoff of an action plan, with initial action a_i chosen at date t and subsequent actions a_{ijk} chosen at dates $\tau_{ij} \geq t + \Delta_i$ is given by

$$\delta^t \omega(a_i) + \sum_{j=1}^{N} \delta^{t+\tau_{ij}} \pi(a_i, a_{ijk}, \theta_j) \mu_{ij}.$$

Although DM knows the true payoff $\omega(a_i)$ she does not know the true payoff $\pi(a_i, a_{ijk}, \theta_j)$. She starts out believing that the true payoff $\pi(a_i, a_{ijk}, \theta_j)$ can take any of the values $\pi(a_i, a_{ijk}, \theta_j, \eta_{ijh})$ in the finite set $\Pi_{ij} = \{\pi(a_i, a_{ijk}, \theta_j, \eta_{ij1}), ..., \pi(a_i, a_{ijk}, \theta_j, \eta_{ijy})\}$ and has a prior belief over those values given by $\nu_{ijk0h} = \Pr(\eta_{ijh})$. To side-step a fundamental conceptual issue, which is beyond the scope of this paper, we shall follow the literature on Bayesian learning and assume that the true payoff is given by $\pi(a_i, a_{ijk}, \theta_j, \eta_{ijh})$ for one of the η_{ijh}.

Before taking any action $a_{ijk} \in A_{ij}$, the DM can learn more about the true payoff associated with that or any other action by engaging in thought experimentation. We model this thought experimentation in an exactly analogous way as in the multi-armed bandit literature. That is, in any given period t DM can 'think' about an action a_{ijk} and obtain a signal σ_{ijk} which is correlated with the true payoff parameter η_{ijh}. Upon obtaining this signal, DM can then revise her belief on the payoff associated with the action a_{ijk} to $\nu_{ijkh} = \Pr(\eta_{ijh} \mid \sigma_{ijk})$.

Thus, at $t = 0$ DM' s decision problem is to decide whether to pick an action a_i in A_0 right away or whether to think ahead about one of the future decision problems. The DM faces this same problem, with possibly updated beliefs from earlier thought experimentation, as long as she has not picked an action a_i.

When she has chosen an action a_i a time Δ_i elapses until a state of nature is realized. Upon realization of a state θ_j, DM' s decision problem is again to decide whether to pick an action a_{ijk} in A_{ij} right away or whether to think about one of the actions in A_{ij}. Should the DM decide to think about the payoff associated with an action in A_{ij} then she faces this same decision problem next period, with updated beliefs.

This general framework is clearly restrictive in some respects: we only allow for two rounds of action choice, the action sets are finite, the state-space is finite and learning through thought-experimentation can only be done for one action at a time. Yet, the framework is sufficiently general and versatile to be able to capture many dynamic decision problems boundedly rational DM' s are likely to face in reality. Also, the framework, as described is clearly too general to be tractable. Accordingly, we shall specialize the framework as a first step and take the dynamic decision problem to be a problem of irreversible investment under uncertainty. In addition, we shall only allow DM to choose between two initial actions, *invest* and *don't invest*. Finally, we shall allow for at most two states of nature and only two actions in each state. We describe this simpler model in greater detail below. Before proceeding to the analysis of the simpler problem it is helpful to briefly describe the benchmark of a rational DM.

2.2 The Benchmark with no deliberation costs

We shall take it that the rational DM can at no cost determine the true value of the payoffs associated with each action $a_{ijk} \in A_{ij}$. The rational DM is then (weakly) better off resolving all future decision problems ahead of their occurence, choosing an initial action immediately, and choosing the follow-up action as soon as the state of nature is realized, provided that the expected present value under the optimal action plan

$$\max_{(a_i, a_{ijk})} \{\omega(a_i) + \sum_{j=1}^{N} \delta^{\Delta_i} \pi(a_i, a_{ijk}, \theta_j) \mu_{ij}\}.$$

is strictly positive. If the expected present value is non-positive then the rational DM should "procrastinate" forever.

2.3 A simple model

More specifically, our simple model has the following basic structure. It involves an initial investment decision with a set-up cost $I > 0$. If DM chooses to invest at date t, then at date $t+1$ the project ends up in one of two states: $\theta \in \{\theta_1, \theta_2\}$. We denote by μ the ex ante probability that state θ_1 occurs. When state θ_i is realized, investment returns are obtained only after the DM has chosen one of two possible actions: a "risky" and a "safe" action. The return of the risky decision, R, is unknown and may take two possible values, $R \in \{\underline{R}, \bar{R}\}$. The return of the safe decision is known and equal to $S \in (\underline{R}, \bar{R})$, so that there is some prior uncertainty about the efficient decision to take in state θ_i. We denote by ν the prior probability that the risky action is the efficient action in either state θ_i. Thus, in our simple model both states have the same payoff structure but payoffs are independently drawn across both states.

As we have described above, the DM may think about the decision to take in each state, which in our simple model means experimenting on the risky action. To keep the analysis as tractable as possible we formulate the simplest possible thought experimentation problem. We shall assume that as DM thinks in any given period, there is a probability λ that she finds out the true payoff associated with the risky strategy, and a probability $(1 - \lambda)$ that she learns nothing. In that case she must continue to think, or experiment, in the subsequent periods until she gets lucky if she wants to find out the true payoff of the risky action[1].

[1] This approach is to be linked with Conlisk's approach of optimization costs. In Conlisk, a decision

Critically, as we have explained, DM can think through what should be done in state θ_i before and/or after investing. She can think about what to do do in one of these states or in both. Of course, if he invests first and state θ_i is realized there is no need to think through what to do in state θ_j.

We shall make the following assumption on the payoff structure:

Assumption:

$$\begin{aligned} A_1 &: \nu\bar{R} + (1-\nu)\underline{R} > S \\ A_2 &: \delta S - I > 0 \end{aligned}$$

Assumption A_1 implies that the safe action is the best action for DM if she decides not to do any thinking. This is not a critical assumption for our main qualitative results. Assumption A_2, on the other hand, is more important. It ensures that the investment project has a positive net present value (NPV) when the safe action is always chosen. Under this assumption, any deliberation on future decisions, or knowledge of the payoff of the risky action, does not impact the initial decision whether to invest: there is always a way of generating a positive NPV, so that there is no ex ante real option value to be maximized. As we explained earlier this assumption is imposed mainly to highlight the difference of our simple model with the classical real options model.

2.4 Solving the Decision Problem

DM can decide to invest right away or to think first about what decision to take in future states of nature. If she were to invest right away she would not be prepared to act immediately following the realization of some state θ_i. She would then have to either act on incomplete prior knowledge, or first think through what to do and thus delay the time when she would reap the returns from her investment. Thus, the benefit of identifying the efficient decision in state θ_i ahead of time (before investing) is that DM will be able to act immediately following the realization of a state of nature and thus reduce the time gap between when she makes an investment outlay and when she reaps the returns from her investment. On the other hand, thinking before investing

maker whose optimal decision is some x^* draws a sequence of random variables $\tilde{x}$ to approximate x^*. This process is costly, e.g. it takes time to collect more draws and the longer the person thinks, the better the estimate of x^*. Our model can be viewed as a particular case of Conlisk's where the decision space is binary and where the draws are either fully informative or not at all.

may be of little use and could unnecessarily delay investment. This is the basic tradeoff DM faces that we now explore.

A lower bound for DM's payoff is the expected return obtained if DM acts immediately and does not think at all (we refer to this as the *no thinking* regime). Under assumption A_1 this expected return is given by:

$$V_{\emptyset} = -I + \delta \left[\nu \bar{R} + (1 - \nu) \underline{R} \right].$$

Consider now the payoff DM could obtain by engaging in some thought experimentation. A first thinking strategy for DM is to initiate the project right away at date 0 with an investment of I, learn which state prevails at date 1, and see whether some thinking is worthwhile once the state is realized.

We shall assume that once in state θ_i DM prefers to think before acting rather than act without thinking. Note that under our learning technology, where following each thought experiment, DM either learns the true payoff for sure (with probability λ) or learns nothing, if it is optimal to undertake one experiment, then it is optimal to continue experimenting until the true payoff is found. Indeed, suppose that when state θ_i is realized DM prefers to experiment. Then, since the decision problem at that point is stationary, she will also prefer to continue experimenting should she gain no additional knowledge from previous rounds of experimentation. Our experimentation problem has been deliberately set up to obtain this particularly simple optimal stopping solution.

If, following successful experimentation DM learns that the true payoff of the risky action is $\underline{R}$ then she will optimally chose the safe action given that $\underline{R} < S$. If, on the other hand, she learns that the true payoff is $\bar{R}$ then she chooses the risky action. Thus, the expected present discounted payoff from thinking in state θ_i is given by

$$\underbrace{\lambda\delta \left[\nu \bar{R} + (1 - \nu) S \right]}_{\substack{\text{present discounted payoff} \\ \text{from learning the true} \\ \text{payoff in the first round} \\ \text{of experimentation}}} + \underbrace{(1 - \lambda)\delta^2 \lambda \left[\nu \bar{R} + (1 - \nu) S \right]}_{\substack{\text{present discounted payoff} \\ \text{from learning the true} \\ \text{payoff in the second round}}}$$

$$+ (1 - \lambda)^2 \delta^3 \lambda \left[\nu \bar{R} + (1 - \nu) S \right] + \sum_{t=4}^{\infty} (1 - \lambda)^{t-1} \delta^t \lambda \left[\nu \bar{R} + (1 - \nu) S \right].$$

Or, letting $\hat{\lambda} = \frac{\lambda\delta}{1-(1-\lambda)\delta}$,the expected present discounted payoff from thinking in state

θ_i is:

$$\hat{\lambda}\left[\nu\bar{R}+(1-\nu)\,S\right].$$

To make our highly simplified problem at all interesting we shall assume that the payoff structure is such that, once in state θ_i DM prefers to think before acting, or:

Assumption:

$$\begin{aligned} A_3 &: \quad \nu\bar{R}+(1-\nu)\underline{R} \leq \hat{\lambda}\left[\nu\bar{R}+(1-\nu)\,S\right] \\ &\Leftrightarrow \\ \hat{\lambda} &\geq \quad \hat{\lambda}_L \equiv \frac{\nu\bar{R}+(1-\nu)\underline{R}}{\nu\bar{R}+(1-\nu)S}. \end{aligned}$$

Assumption A_3 essentially imposes a lower bound on DM' s thinking ability. If DM is a very slow thinker (λ is very small) then it obviously makes no sense to waste a huge amount of time thinking. Thus, for sufficiently high values of λ, DM will choose to "think on the spot" if she has not done any thinking prior to the realization of the state of nature. In that case, when DM chooses to first invests and then think on the spot she gets an ex ante payoff V_L equal to:

$$V_L = -I + \delta\hat{\lambda}\left(\nu\bar{R}+(1-\nu)\,S\right)$$

Notice that under this strategy DM only has to solve only one decision problem: the one she faces once the state of nature is realized.

It is convenient to introduce the following notation: $x \equiv \nu\bar{R}+(1-\nu)S$ and $y = \nu\bar{R}+(1-\nu)\underline{R}$. It is immediate to see that the strategy of "no thinking" dominates the strategy of "thinking on the spot" if and only if:

$$V_L \geq V_{\emptyset} \Leftrightarrow \hat{\lambda} \geq \hat{\lambda}_L = \frac{y}{x}.$$

Consider next a third strategy available to DM, which is to "think ahead" about one or both states of nature. Suppose to begin with that $\mu = \frac{1}{2}$ and that DM, being indifferent between which state to think about first, thinks first about state θ_1. Again, if it is optimal to begin thinking about state θ_1 and DM does not gain any new knowledge from the first thought-experiment then it is optimal to continue thinking until the true payoff of the risky action in state θ_1 is found.

Suppose now that DM has learned the true payoffs in state θ_1, under what conditions should she continue thinking about the other state before investing, instead of investing right away and gambling on the realization of state θ_1? If she decides to think about state θ_2 she will again continue to think until she has found the true payoffs in that state. If she learns that the return on the risky action in state θ_2 is $\bar{R}$, her continuation payoff is

$$V_r = -I + \delta \left[\frac{\pi_1}{2} + \frac{\bar{R}}{2} \right],$$

where $\pi_1 \in \{S, \bar{R}\}$ is DM' s payoff in state θ_1. Similarly, if she finds that the best action in state θ_2 is the safe action, her continuation payoff is:

$$V_s = -I + \delta \left[\frac{\pi_1}{2} + \frac{S}{2} \right].$$

Therefore, DM' s expected continuation payoff from thinking ahead about state θ_2, given that she has already thought through her decision in state θ_1 is:

$$V_E^1 = \hat{\lambda} \left(-I + \delta \left(\frac{\pi_1}{2} + \frac{x}{2} \right) \right).$$

If instead of thinking ahead about state θ_2, DM immediately invests once she learns the true payoffs in state θ_1 her continuation payoff is:

$$V_L^1 = -I + \delta \left(\frac{\pi_1}{2} + \frac{\max\{y, \hat{\lambda} x\}}{2} \right).$$

Thus, continuing to think about state θ_2 before investing, rather than investing right away is optimal if:

$$\Delta^1 \equiv V_E^1 - V_L^1 = -(1 - \hat{\lambda})(\delta \frac{\pi_1}{2} - I) + \frac{\delta}{2} \left[\hat{\lambda} x - \max\{y, \hat{\lambda} x\} \right] \geq 0.$$

From this equation it is easy to characterize the solution of DM's continuation decision problem when she is thinking ahead, once she knows her true payoff in state θ_1. We state it in the following lemma:

Lemma 1: *Suppose that DM is thinking ahead and has already solved her decision problem in state θ_1, then it is better to invest right away and possibly think on the spot in state θ_2 rather than continue thinking ahead about state θ_2 if $\delta \frac{\pi_1}{2} - I \geq 0$.*
If, on the other hand, $\delta \frac{\pi_1}{2} - I < 0$, then there exists a cut-off $\hat{\lambda}_E^1$ such that thinking ahead about state θ_2 is preferred if and only if $\hat{\lambda} \geq \hat{\lambda}_E^1$.

Proof. see the appendix. ■

As one might expect, the decision to continue thinking ahead about state θ_2 depends on what DM has learned before. The higher is π_1, the less keen DM is to continue thinking. When DM finds a good outcome in state θ_1 she want to reap the rewards from her discovery by accelerating investment. By thinking further about state θ_2 she delays investment and if she ends up in state θ_1 anyway her thinking will be wasted. The cost of these delays is captured by $\left(\delta\frac{\pi_1}{2} - I\right)$, the expected payoff in state θ_1. Note, in particular, that a thinking strategy such that DM stops thinking on a bad outcome where she learns that $\pi_1 = S$, but continues thinking on a good outcome, where $\pi_1 = \overline{R}$, is necessarily sub-optimal. This simple observation, we believe, highlights a basic mechanism behind satisficing behavior. Why do boundedly rational DM's settle with good but not necessarily optimal outcomes? Because they want to bring forward the time when they get the good reward, or as the saying goes, because "the best is the enemy of the good".

Having characterized this key intermediate step we are now in a position to determine when DM should start to think ahead at all and when DM should defer until later all of her thinking. Depending on the value of π_1, several cases have to be considered.

First, suppose that $I \leq \delta\frac{S}{2}$. In that case, DM will *not* think about state θ_2 ahead of time, irrespectively of the outcome of her optimization in state θ_1. DM will then think about at most one state. If she thinks ahead about one state her payoff is:

$$V_E = \hat{\lambda}\left[\nu\left(-I + \frac{\delta}{2}\overline{R} + \frac{\delta}{2}\max\{y, \hat{\lambda}x\}\right) + (1-\nu)\left(-I + \frac{\delta}{2}S + \frac{\delta}{2}\max\{y, \hat{\lambda}x\}\right)\right] =$$
$$\hat{\lambda}\left[-I + \frac{\delta}{2}x + \frac{\delta}{2}\max\{y, \hat{\lambda}x\}\right]$$

If she decides to invest without thinking her payoff is:

$$V_L = -I + \delta\max\{y, \hat{\lambda}x\}$$

Comparing the two payoffs we immediately observe that:

Lemma 2: *When* $I \leq \delta\frac{S}{2}$, *thinking ahead is always dominated and DM chooses:*
\- *no thinking if and only if* $\hat{\lambda} \leq \hat{\lambda}_L$, or
\- *thinking on the spot if and only if* $\hat{\lambda} \geq \hat{\lambda}_L$.

Proof. see the appendix. ■

Lemma 2 establishes another important observation. If DM knows that she will stop thinking ahead irrespectively of the outcome of her current thinking, then it is not worth thinking ahead about her current problem. In other words, it is only worth thinking ahead about some state θ_i if DM may want to continue thinking ahead about other states with positive probability; in particular, when the outcome of her current thinking is bad. In our simple model it is quite intuitive that DM would not want to do any thinking ahead if she can guarantee a high net return, which is the case when $\delta\frac{S}{2} \geq I$.

Second, suppose that $I \geq \delta\frac{\overline{R}}{2}$ and that $\hat{\lambda} \geq \hat{\lambda}_L$. In that case, DM wants to think about state θ_2 ahead of time, irrespectively of the outcome of her optimization in state θ_1. In other words, if DM does any thinking ahead, she will want to work out a complete plan of action before investing. Thus, if she thinks ahead her payoff is:

$$V_E = \hat{\lambda}^2 \left[\delta x - I\right]$$

while if she does not, she can expect to get:

$$V_L = -I + \delta\hat{\lambda}x$$

It is easy to check that in this case, $V_L < V_E$.

Third, suppose that $I \geq \delta\frac{\overline{R}}{2}$ but $\hat{\lambda} \leq \hat{\lambda}_L$. In this case DM either thinks ahead and works out a complete plan of action before investing, or DM prefers not to do any thinking ever (thinking on the spot is always dominated by no thinking, since $\hat{\lambda} \leq \hat{\lambda}_L$). If she does not think at all she gets $V_\emptyset$ and if she thinks ahead her payoff is $V_E = \hat{\lambda}^2 \left[\delta x - I\right]$. Comparing $V_\emptyset$ and V_E, we find:

Lemma 3: *When $\hat{\lambda} \leq \hat{\lambda}_L$ thinking on the spot is always dominated and when $I \geq \delta\frac{\overline{R}}{2}$ DM chooses:*
- *no thinking if $\hat{\lambda} \leq \hat{\lambda}_E = \sqrt{\frac{\delta y - I}{\delta x - I}}$, or*
- *thinking ahead (complete planning) if $\hat{\lambda} \geq \hat{\lambda}_E$.*

Proof. see the appendix. ■

In this situation DM is a slow thinker but the costs of thinking ahead are also low as investment costs are high. Thinking on the spot is dominated because once investment costs are sunk DM wants to reap the returns from investment as quickly as possible. On the other hand, as long as investment costs have not been incurred, DM is less concerned about getting a low net expected return quickly.

Fourth, suppose that $\delta\frac{\bar{R}}{2} > I > \delta\frac{S}{2}$. In this intermediate case, DM wants to continue thinking ahead about state θ_2 only if the outcome of her thinking on state θ_1 is bad. Thus, if DM thinks ahead, then with probability ν she learns that the risky action has a payoff $\bar{R}$ in state θ_1 and stops thinking further about state θ_2. Later, of course, if state θ_2 is realized, DM may decide to think on the spot about what to do. With probability $1-\nu$ instead, she learns that the return of the risky decision in state θ_1 is S and decides to continue thinking about state θ_2 before investing. Therefore, in this situation, her payoff if she thinks ahead is:

$$V_E = \hat{\lambda}\left[\nu\left(-I + \frac{\delta}{2}\overline{R} + \frac{\delta}{2}\max\{y, \hat{\lambda}x\}\right) + (1-\nu)\hat{\lambda}\left(-I + \frac{\delta}{2}S + \frac{\delta}{2}x\right)\right]$$

Having determined all the relevant payoffs we are in a position to derive the conditions under which DM prefers to think ahead. To limit the number of cases, and consistent with our focus on the choice between thinking ahead or deferring thinking, we narrow down our analysis to values of $\hat{\lambda}$ greater than $\hat{\lambda}_L$, so that no thinking is always dominated. We provide an analysis of the alternative situation in the Appendix.

Assumption: $A_4 : \hat{\lambda} \geq \frac{y}{x}$

Proposition 1: *Under assumptions A_1, A_2, A_3 and A_4, the solution to DM's decision problem is as follows: DM prefers to think on the spot if and only if:*

$$I \leq \frac{\hat{\lambda}\delta\left[x - \nu\frac{\overline{R}}{2}\right]}{1 + \hat{\lambda} - \nu\hat{\lambda}}.$$

Otherwise DM prefers to think ahead, either adopting:

- a "step by step" strategy where she first thinks about state θ_1 and continues her thinking about state θ_2 if she finds that the payoff in state θ_1 is S and if

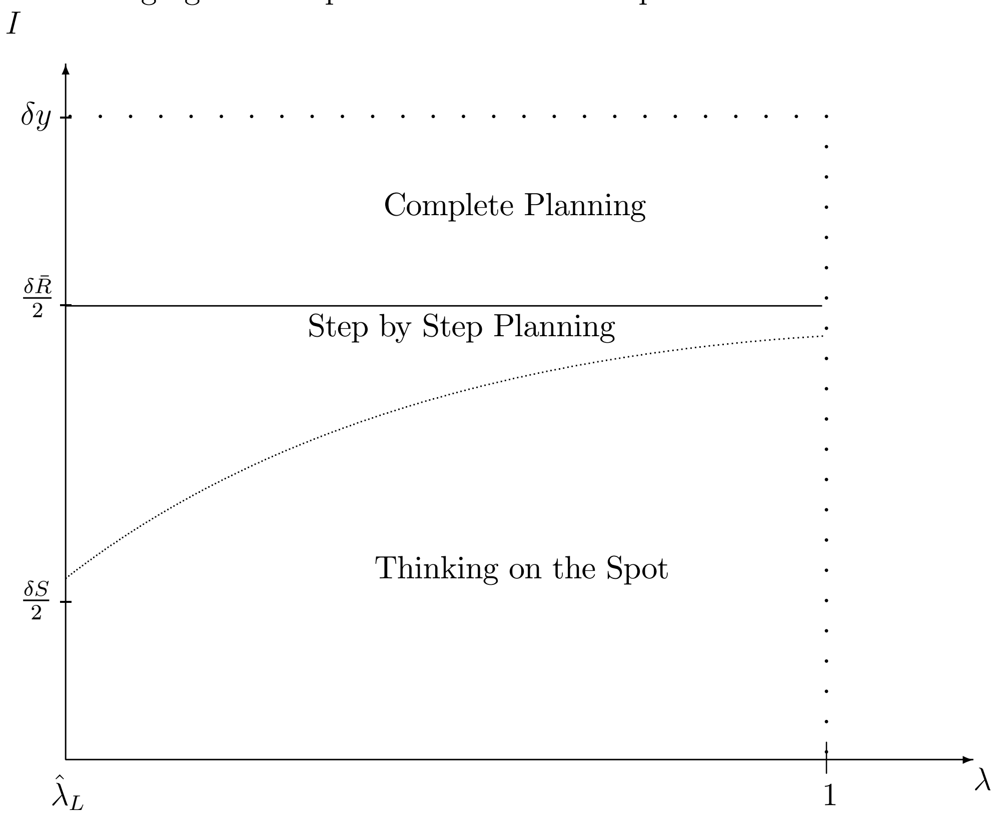
$$\frac{\hat{\lambda}\delta\left[x - \nu\frac{\overline{R}}{2}\right]}{1 + \hat{\lambda} - \nu\hat{\lambda}} \leq I \leq \frac{\delta\overline{R}}{2}.$$

Otherwise, she stops thinking ahead beyond state θ_1 and if state θ_2 is realized resumes her thinking then.

- a "complete planning" strategy where she thinks ahead about both states before investing if $\frac{\delta\overline{R}}{2} \leq I$.

Proof. see the appendix. ■

The following figure is helpful to understand Proposition 1.



The figure maps out the three different regions: i) "complete planning", where DM thinks everything through ahead of time; ii) "thinking on the spot", where DM defers all the thinking to when the state of nature is realized, iii) "step by step planning", where DM thinks first about state θ_1 before investing and maybe about state θ_2 as well. Recall that in this region, we could see DM thinking both before investing and after: she thinks about what to do in state θ_1, learns some good news about that state and invests, but unfortunately state θ_2 is realized. Then she thinks again before making a decision.

Intuition suggests that if DM is a faster thinker, she should do more thinking. This is borne out by the figure, as the area for which DM engages in step-by-step thinking is smaller the larger is λ. Intuition, however, is not a good guide to determine whether a faster thinker thinks more on the spot or plans her actions more. The figure again reveals that there are good reasons why one cannot determine intuitively whether faster thinkers think more on the spot. The answer depends on other factors.

Thus, for high investment costs, DM wants to plan ahead, whether she is a slow or fast thinker (as long as $\lambda \geq \hat{\lambda}_L$). As we have already explained, the reason is that this allows her to better align investment costs and monetary returns in time. When I is high, the NPV is smaller and the cost of thinking ahead goes down. in addition, the benefit of aligning cost and benefits goes up.

On the other hand, for intermediate values of investment costs faster thinkers do more thinking on the spot than slower thinkers, who engage in step-by-step thinking. The reason is that, for faster thinkers the time gap between investment and the realization of returns is smaller and therefore matters less than for slower thinkers. As a result, they are more likely to prefer to bring forward the positive NPV, by investing right away than slower thinkers, who are prepared to accept some delay in getting the positive NPV, in return for a higher NPV achieved by reducing the time lag between investment and recoupment.

Interestingly, the figure reveals that DM will not necessarily engage in complete planning as λ approaches 1. As long as $\lambda < 1$, what determines whether DM engages in complete planning is whether I is greater or smaller than $\frac{\delta\overline{R}}{2}$. Thus even for λ arbitrarily close to 1, DM will not engage in complete planning if $I < \frac{\delta\overline{R}}{2}$. It is only when there are no deliberation costs at all, so that $\lambda = 1$, that complete planning is always a rational behavior.

2.5 Comparative Statics in the simple model

Having characterized decision-making in the presence of positive deliberation costs in a simple symmetric example, we now explore how DM' s behavior is affected in a more general asymmetric setting, while continuing to retain our simple two-state structure.

2.5.1 The Role of State Uncertainty

In the symmetric example we have studied DM faces maximum uncertainty about which state will occur, as each state is equally likely. We now explore how DM' s decisions change when μ, the probability that state θ_1 is realized is larger than $\frac{1}{2}$, while keeping the other characteristics of the model unchanged.

It is immediate to check that the payoff of the "thinking on the spot" strategy is unchanged when $\mu > \frac{1}{2}$. Similarly, the payoff associated with the "complete planning" strategy remains the same. On the other hand, the payoff associated with "step-by-step thinking" is affected in an important way when state θ_1 becomes more likely. When

$\mu > \frac{1}{2}$ it is obvious that the best step-by-step thinking strategy for DM is to think first about state θ_1, the most likely state. The costs of thinking ahead about one state only are the same whichever state DM thinks about, but the expected benefit in terms of DM's ability to act quickly following the realization of a state is higher for state θ_1, as this is the most likely state. Note also that DM is more likely to think ahead about one state only when $\mu > \frac{1}{2}$, as the marginal expected payoff of thinking ahead about the other state is reduced.

The continuation payoff obtained by thinking ahead about state θ_2, once DM knows that the payoff in state θ_1 is π_1 is given by:

$$V_E^1(\mu) = \hat{\lambda}\left(-I + \delta\left(\mu\pi_1 + (1-\mu)x\right)\right)$$

Compare this to the continuation payoff obtained by stopping to think ahead at that point and investing:

$$V_L^1 = -I + \delta\left(\mu\pi_1 + (1-\mu)\hat{\lambda}x\right)$$

and observe that the difference in continuation payoffs is given by:

$$\Delta^1 = V_L^1 - V_E^1(\mu) = (1-\hat{\lambda})(\mu\delta\pi_1 - I)$$

Thus, an increase in μ, the likelihood of state θ_1, results in an increase in Δ^1. As a result, there are now more values of I for which DM stops to think ahead once she knows the payoff in state θ_1.

This has two implications:

First, DM is less likely to work out a complete action-plan before investing, so that the complete planning region in the figure is now smaller.

Second, as the payoff of step-by-step thinking increases with μ, this strategy becomes more attractive relative to not thinking ahead at all.

Consequently we have:

Proposition 2: *A reduction in the uncertainty about states of nature reduces the attractiveness of "complete planning" and "thinking on the spot" and favors a step by step approach.*

Proof. Obvious. ■

As μ approaches 1, it is clear that thinking ahead about state θ_1 only is the best strategy: it allows DM to reduce the time beween when she incurs investment costs and when she recoups her investment. In addition, thinking ahead is very unlikely to create unnecessary delays.

Vice-versa, as μ approaches $\frac{1}{2}$ and uncertainty about the state of nature increases, DM can respond by either working out a complete plan to deal with the greater uncertainty, or she can adopt a "wait-and-see" approach and defer all her thinking until after the uncertainty is resolved.

One may wonder whether this finding extends to changes in other forms of uncertainty. For instance, would DM change her behavior following a change in her prior belief ν_i in state θ_i that the risky action is efficient? It turns out that such a change in beliefs not only affects the perceived riskiness of the risky decision, but also changes the average payoff that DM can expect in state θ_i. However, if one considers a mean preserving change, where as ν_i is reduced $\overline{R}_i$ is increased to keep the mean return on the risky action unchanged, one can isolate the effect of this change in uncertainty. It turns out that such a change in uncertainty has no impact on DM' s deliberation policy. As long as x remains constant, thinking on the spot, complete planning and step-by-step thinking all have unchanged payoffs following a mean-preserving change in ν_i.

2.5.2 Variations in Problem Complexity

So far we have interpreted the parameter λ as a measure of DM' s thinking ability. But λ can also be seen as a measure of the difficulty of the problem to be solved, with a higher λ denoting an easier problem. If we take that interpretation, we can let λ vary with the problem at hand and ask whether DM thinks first about harder or easier problems. Thus, suppose that the decision problem in state θ_1 is easier, or less complex, than the one in state θ_2. That is, suppose that $\lambda_1 > \lambda_2$. Holding everything else constant, we now explore how variations in the complexity of decision problems affects DM' s payoff and behavior. It is immediate to verify that the payoff associated to the complete planning strategy is:

$$V_E = \hat{\lambda}_1 \hat{\lambda}_2 \left[\delta x - I\right]$$

and the payoff of "thinking on the spot" is:

$$V_L = -I + (\hat{\lambda}_2 + \hat{\lambda}_1)\frac{\delta x}{2}.$$

Comparing these two payoffs we can illustrate a first effect of variation in problem complexity. Take $\hat{\lambda}_1 = \hat{\lambda} + \varepsilon$ and $\hat{\lambda}_2 = \hat{\lambda} - \varepsilon$, where $\varepsilon > 0$. Note, first, that the payoff under the "thinking on the spot" strategy is the same whether $\varepsilon = 0$, as we have assumed before, or $\varepsilon > 0$. On the other hand, the payoff under complete planning, $(\hat{\lambda}^2 - \varepsilon^2)\,[\delta x - I]$, is lower the higher is ε. Thus, increasing the variance in problem-complexity across states, while keeping average complexity constant, does not affect the payoff under the thinking-on-the-spot strategy as DM then only incurs average thinking costs. In contrast, when DM attempts to work out a complete action-plan, thinking costs compound. That is why DM is worse off under a complete planning strategy when the variance in problem-complexity increases.

How do differences in problem-complexity across states affect DM' s step-by-step thinking strategy? Intuitively, it seems to make sense to first have a crack at the easier problem first. This intuition is, indeed, borne out in the formal analysis, but the reason why it makes sense to start first with the easier problem is somewhat subtle. Under the step-by-step thinking strategy, whenever DM ends up thinking about both states of nature, she compounds thinking costs in the same way as under complete planning, and whether she starts with state θ_1 or θ_2 is irrelevant.

Hence, the choice of which state to think about first only matters in the event when: i) the outcome of thinking is good (that is, if DM learns that the true payoff is $\bar{R}$), so that DM engages in only partial planning before investment, and; ii) when the state of nature which DM has thought about is realized. In this event, DM is better off thinking first about the easier problem since she then gets to realize the net return from investment sooner. Formally, the payoffs under the two alternative step-by-step strategies are given by:

$$V_{E1} = \hat{\lambda}_1 \left[\nu \left(-I + \frac{\delta}{2}\overline{R} + \frac{\delta}{2}\hat{\lambda}_2 x \right) + (1-\nu)\hat{\lambda}_2 \left(-I + \frac{\delta}{2}S + \frac{\delta}{2}x \right) \right]$$

$$V_{E2} = \hat{\lambda}_2 \left[\nu \left(-I + \frac{\delta}{2}\overline{R} + \frac{\delta}{2}\hat{\lambda}_1 x \right) + (1-\nu)\hat{\lambda}_1 \left(-I + \frac{\delta}{2}S + \frac{\delta}{2}x \right) \right]$$

Therefore, as $\left(-I + \frac{\delta}{2}\overline{R}\right) > 0$, it is best to think about the simple problem, with the higher $\hat{\lambda}_i$ first. We summarize our discussion in the proposition below:

Proposition 3: *Let* $\hat{\lambda}_1 = \hat{\lambda} + \varepsilon$ *and* $\hat{\lambda}_2 = \hat{\lambda} - \varepsilon$, *with* $\varepsilon > 0$.

Then, when DM chooses to think ahead, it is (weakly) optimal to think first about the easier problem, and the payoff associated with:

- the "complete planning" strategy is decreasing in the variance in problem-complexity across states,

- the "thinking on the spot" strategy is unchanged,

- the "step by step" strategy, where DM thinks first about the simple problem (in state θ_1), rises (resp. falls) with the variance in problem-complexity across states if $I < I_m$ (resp. $I > I_m$), for some cut-off $I_m > 0$.

The larger is the difference in complexity between the two decision problems (as measured by ε) the more attractive is the "step by step" approach.

Proof. See the appendix. ■

Thus, partial planning is more likely to take place in environments where problem-complexity varies a lot across states. Moreover, the simpler problems are solved first, while the harder ones are deferred to later in the hope that they won't have to be solved. These conclusions are quite intuitive and reassuring.

2.5.3 Variations in Expected Returns across States

If thinking first about more likely and easier problems makes sense, is it also desirable to think first about problems with a higher expected payoff? In our simple model expected returns may differ across states if either the probability of high returns on the risky action is different, or if returns themselves differ. We explore each variation in turn.

Suppose first that prior beliefs of high returns on the risky action in state θ_1 is higher than in state θ_2, but that the average belief across states remains unchanged : $\nu_1 = \nu + \varepsilon$, $\nu_2 = \nu - \varepsilon$, with $\varepsilon > 0$. In that case, the payoff of the "thinking on the spot", or "complete planning" strategies are unchanged. In contrast, the payoff of the step by step strategy is affected by this average-belief-preserving spread as follows. Whichever state θ_i DM thinks about first, DM will want to stop thinking further about the other state if she discovers that $\pi_i = \frac{\delta\overline{R}}{2}$, whenever $I \in [\frac{\delta S}{2}, \frac{\delta\overline{R}}{2}]$. It is then best for her to start thinking about state θ_1, the state with the higher prior belief ν_1. The reason is that she is then more likely to find that $\pi_1 = \frac{\delta\overline{R}}{2}$ and if state θ_1 arises, DM will be able to realize high returns relatively quickly. Note, therefore, that as more prior probability mass is shifted to the high return on the risky action in state θ_1, the step by step strategy also becomes relatively more attractive.

Second, suppose that returns themselves differ across the two states, and that returns in state θ_1 are higher than in state θ_2 : $S_1 = S + \varepsilon$ while $S_2 = S - \varepsilon$, and $\overline{R}_1 = \overline{R} + \varepsilon$

while $\overline{R}_2 = \overline{R} - \varepsilon$. We take $\varepsilon > 0$ to be small, so that even in state θ_2 DM will continue to think before acting if she does not know which decision is efficient. More formally, we take ε to be small enough that assumption A_3 remains valid.

This redistribution of returns across states leaves DM' s expected payoff unaffected as long as $I < \frac{\delta(S-\varepsilon)}{2}$ or $I > \frac{\delta(\overline{R}+\varepsilon)}{2}$. Indeed, in that case she chooses to either defer all of her thinking until the uncertainty about the state is resolved, or to work out a complete plan before investing. In both cases, her expected payoff only depends on average returns $\frac{x_1+x_2}{2}$ and is therefore unaffected by changes in ε.

But if $\frac{\delta(S+\varepsilon)}{2} < I < \frac{\delta(\overline{R}-\varepsilon)}{2}$, DM will engage in step-by-step thinking and will stop thinking ahead if she learns that the risky decision is efficient in the state she thinks through first. Once again, it is then best for her to think about the high payoff state first. The basic logic is the same as before: by thinking first about the high payoff state DM is able to bring forward in time the moment when she realizes the highest return.

Does this mean that when DM chooses to think ahead, she is always (weakly) better off thinking first about the high payoff state? The answer to this question is yes. There is one particular situation where conceivably this prescription might not hold. That is when $\frac{\delta(S-\varepsilon)}{2} < I < \frac{\delta(S+\varepsilon)}{2}$. In this case DM either thinks about the low return state first, or she defers all of her thinking to when the state of nature is realized. However, in the later case, her payoff is unaffected by changes in ε, while in the former it is decreasing in ε. What is more, for $\varepsilon = 0$ thinking on the spot dominates step by step thinking. Therefore, the latter strategy is then dominated by thinking on the spot for all $\varepsilon > 0$. Hence, it remains true that whenever thinking ahead pays it is best to think first about the problem with the highest return.

Finally, for $\frac{\delta(\overline{R}-\varepsilon)}{2} < I < \frac{\delta(\overline{R}+\varepsilon)}{2}$, DM either thinks first about the high return state or works out a complete plan. In the later case, her payoff is unaffected by ε. If she decides to think first about the high return state, her payoff goes up with ε so that eventually this strategy dominates complete planning. We then deduce:

Proposition 4: *Suppose that either* $\nu_1 = \nu + \varepsilon$, and $\nu_2 = \nu - \varepsilon$ or that $S_1 = S + \varepsilon$, $S_2 = S - \varepsilon$, while $\overline{R}_1 = \overline{R} + \varepsilon$ and $\overline{R}_2 = \overline{R} - \varepsilon$, *for* $\varepsilon > 0$ *but small.*

- *Whenever some thinking ahead takes place, it is best to start thinking about the high payoff state* (θ_1)*.*

- *The payoff associated with the "complete planning" or the "thinking on the spot" strategies is unaffected by changes in ε, however*

- *The payoff associated with the "step by step" strategy, where DM thinks first about the high return state, is increasing in ε.*

Proof. See the appendix. ■

It, thus, appears to be best for DM to think first about the most favorable state. Can this be seen as a rationalization of wishful thinking?

2.5.4 Thinking in Teams

Our ultimate goal is to analyze *satisficing contracts* between boundedly rational agents as modeled here. One step in that direction is to consider how a team of DMs, each facing positive deliberation costs, chooses to think ahead. Here, we only explore one simple angle of the team problem, which can be easily related to our comparative statics analysis.

Thus, suppose that there is now not one DM but two identical twins, each possibly investing the same amount of money and sharing equally the returns of the investment. Suppose that each team member has the same thinking ability given by λ_T. One key difference of deliberation by a team is that each team member can think about the same problem in parallel until one of the team members has discovered the solution. This parallel deliberation lets the team accelerate the process of discovery. Concretely, if both of team members think about, say, state θ_i they can expect to find the solution with probability

$$1 - (1 - \lambda_T)^2 = \lambda_T(2 - \lambda_T).$$

For the sake of argument, suppose that λ_T is such that

$$\lambda_T(2 - \lambda_T) = \lambda,$$

so that the team of twins, thinking about the same problem, is just as good as the single DM we have considered so far.

Then, it follows that the team payoff associated with the "thinking on the spot" strategy remains unchanged. The same is not true, however, for other strategies. Now, when thinking ahead, the team can learn faster when each member works on a different problem. This, avoids duplication and reduces compounding of delay: if one twin thinks about state θ_1 and the other about state θ_2, there is now a chance that they together learn the solution to their respective problems. This is clearly better so that:

Proposition 5: *When there are two identical agents working in a team, each with thinking ability λ_T given by $\lambda_T(2-\lambda_T)=\lambda$,*

- *The team payoff associated with:*
  - the "complete planning" and the "step by step" strategies is higher than with a single agent,
  - the "thinking on the spot" strategy is the same,

- *There is more planning ahead in the team than by the single DM.*

Proof. See the appendix. ■

By deferring thinking, the team cannot benefit as much from parallel thinking: once the state is realized, both team members must think on the same problem, thus partially duplicating their efforts. While planning ahead, duplication may be avoided, and therefore the benefits of some form of planning, with each agent thinking about a different state of the world are increased.

3 The Model with N states

One may wonder to what extent our insights into optimal dynamic decision and deliberation strategies derived in our tractable two state example extend to more general settings. We attempt a limited exploration into this question by partially analyzing the more general problem with a finite, arbitrary number, of states of nature, $N \geq 3$.

We begin by looking at the case of N equiprobable states, each with the same structure: a safe action with known payoff S, and a risky action with unknown payoff $R \in \{\underline{R}, \overline{R}\}$. As before, we assume that payoffs of the risky action are distributed independently across states and take DM' s prior belief to be the same in each state and given by $\Pr(R=\bar{R})=\nu$. We also continue to assume that assumption A_3 holds.

As there is a finite number of states to explore, the optimal deliberation policy cannot be stationary. As and when DM discovers the solution to future decision problems in some states of nature by thinking ahead, she is more likely to stop thinking further, given that the remaining number of unexplored states diminishes. We also know from our previous analysis that she may be more or less willing to continue her thought experimentation before investing, depending on whether she learns 'good' or 'bad' news about her payoff in state she is thinking about. Therefore there is no hope in identifying a simple optimal policy where, for instance, DM explores m^* out of N states and then stops. It is also unlikely that DM' s optimal policy would take the form of a simple stopping rule, such that DM would stop thinking further once she discovers that she can obtain a minimum average return in the states she has successfully explored. To identify DM' s optimal policy we therefore proceed in steps and characterize basic properties the optimal policy must statisfy.

One property we identified in our two state example is that if DM is sure to stop thinking and to invest after learning her payoff about the current state she thinks about, then she prefers not to think ahead about the current state. The next lemma establishes that this property holds in the more general model:

Lemma 4: *If DM prefers to stop thinking ahead and to invest, irrespectively of what she learns about state θ_i, then she also prefers not to think ahead about state θ_i.*

Proof. Suppose that DM has already deliberated about m out of N states, and found that z_m of these states have a payoff for the risky action of $\overline{R}$. If she deliberates about the $(m+1)^{th}$ state and knows that she will then invest no matter what, she can expect to get:

$$\widehat{\lambda}\left[-I+\frac{\delta}{N}\left(z_m\overline{R}+(m-z_m)S+x+(N-(m+1))\widehat{\lambda}x\right)\right]$$

Thus, suppose by contradiction that DM decides to deliberate on this $(m+1)^{th}$ state. This is in her best interest if the payoff above is larger than what she could get by investing right away :

$$-I+\frac{\delta}{N}\left(z_m\overline{R}+(m-z_m)S+(N-m)\widehat{\lambda}x\right)$$

Therefore, it must be the case that

$$C_1:-I+\frac{\delta}{N}\left(z_m\overline{R}+(m-z_m)S+(N-(m+1))\widehat{\lambda}x\right)<0$$

Now, if DM were sure to stop deliberating after the $(m+1)^{th}$ state, she must be better off stopping further deliberations even when she learns bad news about the m^{th} state. For that to be true, it must be the case that stopping even following bad news is better than continuing exploring just one more state $(m+2)$, before investing, or that:

$$-I+\frac{\delta}{N}\left(z_m\overline{R}+(m-z_m)S+S+(N-(m+1))\widehat{\lambda}x\right)\geq$$
$$\widehat{\lambda}\left[-I+\frac{\delta}{N}\left(z_m\overline{R}+(m-z_m)S+S+x+(N-(m+2))\widehat{\lambda}x\right)\right]$$

or:

$$-I+\frac{\delta}{N}\left(z_m\overline{R}+(m-z_m)S+S+(N-(m+2))\widehat{\lambda}x\right)\geq 0$$

or

$$-I+\frac{\delta}{N}\left(z_m\overline{R}+(m-z_m)S+(N-(m+1))\widehat{\lambda}x\right)-\frac{\delta}{N}\left(\widehat{\lambda}x-S\right)\geq 0$$

which implies that condition C_1 is violated, , as $\widehat{\lambda}x-S>0$. ■

One implication of this simple lemma is that if DM wants to do some thinking ahead, then she may also want to work out a complete plan of action with positive probability (e.g. in the event that she only learns bad news from her deliberations).

Recall that the reason why it may be in DM' s interest to do some thinking ahead is not to improve her current decision. In this respect, thinking ahead adds no value, as investment is guaranteed to yield a positive net present value. Rather, the reason why DM gains by thinking ahead is that she may be able to respond faster to the realization of a state of nature. So the intuition behind lemma 4 is that if the gain from a quick response in state $(m+1)$ does not justify the cost of delaying investment, and this even after learning that the expected net present value of the investment will be lower, then it cannot possibly be the case that this gain exceeds deliberation costs prior to learning the bad news.

This intuition immediately suggests that the following property of the optimal deliberation policy must be satisfied.

Lemma 5: *It is never optimal to stop thinking ahead on learning bad news and to continue thinking ahead on learning good news.*

Proof. Suppose, by contradiction that DM sometimes thinks ahead on good but not on bad news. Then there must exist a state m such that when DM learns that in this state the safe decision is optimal she invests right away while if the risky decision is optimal she continues exploring some more states.

Thus, suppose that after learning a payoff of $\overline{R}$ in state m, DM continues to explore $k(\sigma)$ states, where σ is a sequence of payoffs $(S, \overline{R})$ that DM uncovers in states $m+1, ..., m+k$. Denote by $v(\sigma)$ the total payoff in those states given the sequence σ. DM then prefers to continue deliberating after discovering that the payoff of the risky action in state m is $\overline{R}$, if:

$$-I + \frac{\delta}{N}\left(z_m\overline{R} + (m - z_m)\,S + \overline{R} + (N - (m+1))\,\widehat{\lambda}x\right) \leq$$
$$E_\sigma\left[\widehat{\lambda}^{\widetilde{k}(\sigma)}\left[-I + \frac{\delta}{N}\left(z_m\overline{R} + (m - z_m)\,S + \overline{R} + v(\sigma) + \left(N - \left(m + \widetilde{k}(\sigma) + 1\right)\right)\widehat{\lambda}x\right)\right]\right]$$

for some continuation policy $\widetilde{k}(\sigma)$, or:

$$-I + \frac{\delta}{N}\left(z_m\overline{R} + (m - z_m)\,S + (N - (m+1))\,\widehat{\lambda}x\right) \leq$$
$$E_\sigma\left[\widehat{\lambda}^{\widetilde{k}(\sigma)}\left[-I + \frac{\delta}{N}\left(z_m\overline{R} + (m - z_m)\,S + v(\sigma) + \left(N - \left(m + \widetilde{k}(\sigma) + 1\right)\right)\widehat{\lambda}x\right)\right]\right]$$
$$-\frac{\delta}{N}\overline{R}\left(1 - E_\sigma\left(\widehat{\lambda}^{\widetilde{k}(\sigma)}\right)\right)$$

The last term is negative so that if we now substitute $\overline{R}$ by S, we have:

$$-I + \frac{\delta}{N}\left(z_m\overline{R} + (m - z_m)\,S + S + (N - (m+1))\,\widehat{\lambda}x\right) \leq$$
$$E_\sigma\left[\widehat{\lambda}^{\widetilde{k}(\sigma)}\left[-I + \frac{\delta}{N}\left(z_m\overline{R} + (m - z_m)\,S + S + v(\sigma) + \left(N - \left(m + \widetilde{k}(\sigma) + 1\right)\right)\widehat{\lambda}x\right)\right]\right]$$

which means that if DM followed the policy that is optimal when discovering that the payoff of the risky action in state m is $\overline{R}$, when in fact it is $\underline{R}$, so that the best DM can hope to get in that state is S, then she would get a higher payoff by deliberating further than by investing right away. In addition, DM can re-optimize the continuation policy after learning that her payoff in state m is S. Therefore, she is better off continuing to explore before investing, a contradiction. ■

From these two simple observations we are able to infer that:

Theorem: *For N small enough, DM adopts a step by step strategy whereby she thinks ahead about some states, continues to do so upon learning bad news and invests only once she has accumulated enough good news.*

Proof. The two previous lemmata imply that DM never stops exploring upon receiving bad news. Indeed, if she did then she would not continue either when receiving good news. This in turn would imply that it would be best not to think ahead about that particular state. Therefore, when following a step by step strategy DM only stops on good news about the last state she explores.

Turning to the second part of the theorem, to see why it is best to do some thinking ahead when N is sufficiently small, it suffices to note that thinking ahead is obviously beneficial if $N = 1$. Also, the strategy of investing right away delivers a lower payoff than first thinking ahead about exactly one state (which, itself is a dominated strategy), if:

$$\underbrace{-I + \delta\widehat{\lambda}x}_{\text{invest immediately}} \leq \underbrace{\widehat{\lambda}\left[-I + \frac{\delta}{N}x + \frac{\delta}{N}(N-1)\widehat{\lambda}x\right]}_{\text{think ahead about one state}}$$

or,

$$\Leftrightarrow N \leq \frac{\delta\widehat{\lambda}x}{\delta\widehat{\lambda}x - I}$$

■

As stated, the theorem establishes that DM would want to do some thinking ahead if the number of possible states N is small enough. We also know from the analysis of the example with two states that even for $N = 2$, DM may prefer thinking ahead step-by-step rather than determine a complete action plan. However, what we cannot conclude from the theorem or our previous analysis is that DM would definitely not want to do any thinking ahead when N is sufficiently large. We establish this result in the next proposition.

Proposition 6: *Consider N equiprobable states. If $N \geq \frac{\delta S}{\delta S - I}$ the optimal thinking strategy is to invest right away and to think on the spot.*

Proof. Notice first that if $N \geq \frac{\delta S}{\delta S - I}$, DM will never think about the last state. That is, when $N \geq \frac{\delta S}{\delta S - I}$ then $m^* < N$. To see this, denote by Q_{N-1} the average payoff that DM expects from the $N-1$ other states. We then have:

$$\widehat{\lambda}\left[-I + \frac{\delta}{N}x + \frac{\delta}{N}(N-1)Q_{N-1}\right] \leq -I + \frac{\delta}{N}\widehat{\lambda}x + \frac{\delta}{N}(N-1)Q_{N-1}$$

$$\Leftrightarrow N(\delta Q_{N-1} - I) \geq \delta Q_{N-1}$$

as $Q_{N-1} \in (S, \overline{R})$ this is equivalent to $N \geq \frac{\delta Q_{N-1}}{\delta Q_{N-1} - I}$, which is true if $N \geq \frac{\delta S}{\delta S - I}$. Now from lemma 4 we know that if DM stops thinking ahead at the penultimate state irrespectively of what she learns then she also prefers not to think ahead about the penulatimate state, irrespectively of what she learned before. By backward induction, it then follows that DM prefers no to do any thinking ahead. ■

Proposition 6 is quite intuitive. There is little to be gained by working out a highly incomplete action-plan and if any reasonable plan is complex and will take a long time to work out then the only reasonable course of action is to just hope for the best and not do any planning at all.

In reality, people sometimes prefer to be faced with complex life situations where rational forethought makes little difference and the only course of action is to essentially "put their fate in God' s hands" so to speak. In such situations they are absolved of all responsibility for their fate and that makes them better off. In contrast, here our DM is always (weakly) worse off facing more uncertainty than less (as measured by the number of states N). Indeed, if one were to reduce the number of states below $\frac{\delta S}{\delta S - I}$, DM would want to do some thinking ahead and be better off as a result.

Proposition 7: *Consider N equiprobable states but differing in terms of expected payoffs with $x_1 > x_2 > ... > x_N$. The optimal thinking strategy is to think ahead about the best possible states first.*

Proof. To be completed. ■

Proposition 6: *Consider N states with probabilities of realization $\mu_1 > \mu_2 > ... > \mu_N$. The optimal thinking strategy is to think ahead about the most likely states first.*

Proof. To be completed. ■

4 Conclusion

...

5 References

Berry, D. A. and D. Fristedt (1985), *Bandit Problems: Sequential Allocation of Experiments*, London, Chapman Hall

Conlisk, John (1988), "Costly Optimization", *Journal of Economic Behavior and organization*, 9 (3), 213-28.

Conlisk, John. "Why bounded rationality?" *Journal of Economic Literature*, June 1996, 34 (2), pp.669-700.

Dixit, A.K. and Pindyck, (1994) *Investment under Uncertainty*, Princeton University Press

Gabaix, Xavier and Laibson, David. "A boundedly rational decision algorithm" *AER Papers and Proceedings*, May 200.

Gittins, J.C. and D. M. Jones (1974), "A Dynamic Allocation Index for the Sequential Design of Experiments.", in *Progress in Statistics*, Gani, Sarkadi and Vince (eds.), New York, North Holland

Gittins, J.C. (1979) "Bandit processes and dynamic allocation indices", *Journal of the Royal Statistical Society* B, 41, 2, p.148-77.

Henry, C. (1974a) "Option Values in the Economics of Irreplaceable Assets.", *Review of Economic Studies*, 41, 89-104

Henry, C. (1974b) "Investment Decisions under Uncertainty: The Irreversibility Effect", *American Economic Review*, 64, 1006–1012.

Jehiel, Philippe. "Limited horizon forecast in repeated alternate games," *Journal of Economic Theory*, 1995, 67: 497-519.

Kahneman, Daniel and Tversky, Amos. "Judgment Under Uncertainty: Heuristics and Biases", *Science*, 1974.

Macleod, W. Bentley. (1999) "Complexity, Bounded Rationality and Heuristic Search," mimeo, University of Southern California.

Payne, John W., James R. Bettman, and Eric J. Johnson, *The Adaptive Decision Maker*, 1993, Cambridge: Cambridge University Press.

Radner, R. (1975) "Satisficing", *Journal of Mathematical Economics*, 2 (2), 253-62.

Radner, R. and M. Rothschild (1975), "On the Allocation of Effort", *Journal of Economic Theory*, 10 (3), 358-76.

Rothschild, M. (1974) “A Two-Armed Bandit Theory of Market Pricing,” *Journal of Economic Theory*, 9, 185-202.

Rubinstein, Ariel. (1998), *Modeling Bounded Rationality.* MIT Press: Cambridge.

Simon, Herbert. “A behavioral model of rational choice,” *Quarterly Journal of Economics*, 1955, 69 (1).

Whittle, P. (1980) “Multi-armed Bandits and the Gittins Index”, *Journal of the Royal Statistical Society*, 42, 143-149

Whittle, P. (1982), *Optimization over time: Dynamic Programming and Stochastic Control*, John Wiley, New York

Winston, G. (1989) “Imperfectly Rational Choice: Rationality as the Result of a Costly Activity”, *Journal of Economic Behavior and organization*, 12 (1), 67-86.

APPENDIX

- *Proof of Lemma 1:*

It is immediate from the previous equation that thinking on the spot dominates if

$$\delta\frac{\pi_1}{2} - I \geq 0.$$

Suppose now that $\delta\frac{\pi_1}{2} - I < 0$.

- If $\hat{\lambda} \geq \hat{\lambda}_L$, then $\max\{y, \hat{\lambda}x\} = \hat{\lambda}x$ and $\Delta = (1-\hat{\lambda})(\delta\frac{\pi_1}{2} - I) < 0$ so that thinking ahead dominates.

- If $\hat{\lambda} \leq \hat{\lambda}_L$, then $\max\{y, \hat{\lambda}x\} = y$ and

$$\Delta = (1-\hat{\lambda})(\delta\frac{\pi_1}{2} - I) + \frac{\delta}{2}\left[y - \hat{\lambda}x\right]$$

therefore,

$$\begin{aligned} \Delta &\leq 0 \Leftrightarrow \\ -I + \delta\frac{\pi_1}{2} + \frac{\delta}{2}y &\leq \hat{\lambda}\left[-I + \delta\frac{\pi_1}{2} + \frac{\delta}{2}x\right] \\ &\Leftrightarrow \hat{\lambda} \geq \hat{\lambda}_E \equiv \frac{-I + \frac{\delta}{2}(\pi_1 + y)}{-I + \frac{\delta}{2}(\pi_1 + x)} \end{aligned}$$

- *Proof of Lemma 2*:

The difference $\Delta = V_E - V_L$ is always negative as:

$$\begin{aligned} \Delta &= \hat{\lambda}\left[-I + \frac{\delta}{2}x + \frac{\delta}{2}\max\{y, \hat{\lambda}x\}\right] - \left[-I + \delta\max\{y, \hat{\lambda}x\}\right] \\ &= -\left[-I + \delta\max\{y, \hat{\lambda}x\}\right](1-\hat{\lambda}) + \hat{\lambda}\frac{\delta}{2}\left[x - \max\{y, \hat{\lambda}x\}\right] \end{aligned}$$

Case 1: $\hat{\lambda} \leq \hat{\lambda}_L$.

$$\begin{aligned} \Delta &= -\left[-I + \delta y\right](1-\hat{\lambda}) + \hat{\lambda}\frac{\delta}{2}\left[x - y\right] \\ &= I - \delta y + \hat{\lambda}\left[-I + \delta y + \frac{\delta}{2}(x - y)\right] \end{aligned}$$

as the term in bracket is positive, Δ is at most equal to:

$$\begin{gathered}I-\delta y+\frac{y}{x}\left[-I+\delta y+\frac{\delta}{2}(x-y)\right]= \\ \frac{1}{x}(Ix-\delta yx-Iy+\delta y^2+\frac{\delta}{2}(x-y)y)= \\ \frac{x-y}{x}(I-\frac{\delta}{2}y)<0\end{gathered}$$

and therefore thinking on the spot dominates thinking ahead. As $\hat{\lambda} \leq \hat{\lambda}_L$, no thinking dominates thinking on the spot.

Case 2: $\hat{\lambda} \geq \hat{\lambda}_L$.

$$\begin{aligned}\Delta &= -\left[-I+\delta\hat{\lambda}x\right](1-\hat{\lambda})+\hat{\lambda}\frac{\delta}{2}\left[x-\hat{\lambda}x\right] \\ &= (1-\hat{\lambda})\left[I-\frac{\delta}{2}\hat{\lambda}x\right]<0\end{aligned}$$

and therefore thinking on the spot dominates thinking ahead. As $\hat{\lambda} \geq \hat{\lambda}_L$, thinking on the spot also dominates no thinking.

- *Proof of Lemma 3*:

As in the proof of Lemma 2, there are two cases to consider:

Case 1: $\hat{\lambda} \leq \hat{\lambda}_L$.

In that case, no thinking dominates thinking on the spot. Whether thinking ahead is best depends on the sign of

$$\Delta = \hat{\lambda}^2\left[\delta x - I\right] + I - \delta y$$

and this is positive for

$$I \geq \frac{\delta(y-\hat{\lambda}^2 x)}{1-\hat{\lambda}^2}$$

The right hand side of this inequality is decreasing in $\hat{\lambda}$. This is not true for $\hat{\lambda}=0$ but it is the case for $\hat{\lambda}=\frac{y}{x}$ as in that case, this inequality becomes:

$$I \geq \frac{\delta(y-\frac{y^2}{x^2}x)}{1-\frac{y^2}{x^2}} = \frac{\delta xy}{x+y}$$

but as $I \geq \delta\frac{\overline{R}}{2}$, then $I > \frac{\delta x}{2}$ and we remark that

$$\begin{aligned} \frac{\delta x}{2} &> \frac{\delta xy}{x+y} \\ &\Leftrightarrow x+y > 2y \end{aligned}$$

which is true as $x > y$. Therefore there exists a new threshold $\hat{\lambda}_E < \hat{\lambda}_L$ such that thinking ahead dominates no thinking if and only if $\hat{\lambda}$ exceeds that threshold.

Case 2: $\hat{\lambda} \geq \hat{\lambda}_L$.

In that case, no thinking is dominated by thinking on the spot. Whether thinking ahead is best depends on the sign of

$$\Delta = \hat{\lambda}^2 [\delta x - I] + I - \delta\hat{\lambda}x$$

and this is positive for

$$I \geq \frac{\delta\hat{\lambda}x}{1+\hat{\lambda}}$$

The right hand side of this inequality is increasing in $\hat{\lambda}$. It is enough to notice that this is true for $\hat{\lambda} = 1$ as in that case, this inequality becomes:

$$I > \frac{\delta x}{2}$$

which is true.

- *Proof of Proposition 1*:

First, lemmas 2 and 3 tell us that thinking ahead is dominated if $I \leq \frac{\delta S}{2}$ and that thinking on the spot is dominated if $I \geq \frac{\delta\overline{R}}{2}$.

Case 1: suppose first as we did in the text that $\hat{\lambda} \geq \hat{\lambda}_L$. Then surely for $I \geq \frac{\delta\overline{R}}{2}$, the best strategy is to think ahead as in that case it dominates thinking on the spot which itself dominates no thinking. Because I is so high, DM will think about both states (from lemma 1) before investing. Conversely for $I \leq \frac{\delta S}{2}$, deferring all thinking is best as it dominates both thinking ahead and no thinking.

We are left with the study of the intermediate state where $\frac{\delta S}{2} \leq I \leq \frac{\delta\overline{R}}{2}$, when the best alternative to planning ahead is to follow the strategy of thinking on the spot. We have:

$$\begin{aligned}
\Delta &\equiv V_E - V_L \\
&= \hat{\lambda}\left[\nu\left(-I+\frac{\delta}{2}\overline{R}+\frac{\delta}{2}\hat{\lambda}x\right)+(1-\nu)\hat{\lambda}\left(-I+\frac{\delta}{2}S+\frac{\delta}{2}x\right)\right] \\
&\quad -\left[-I+\delta\hat{\lambda}x\right] \\
&= \hat{\lambda}\left[\nu\left(-I+\frac{\delta}{2}\overline{R}\right)+\frac{\delta}{2}\hat{\lambda}x+(1-\nu)\hat{\lambda}\left(-I+\frac{\delta}{2}S\right)\right] \\
&\quad -\left[-I+\delta\hat{\lambda}x\right] \\
&= \hat{\lambda}\left[\hat{\lambda}\nu\left(-I+\frac{\delta}{2}\overline{R}\right)+(1-\hat{\lambda})\nu\left(-I+\frac{\delta}{2}\overline{R}\right)+\frac{\delta}{2}\hat{\lambda}x+(1-\nu)\hat{\lambda}\left(-I+\frac{\delta}{2}S\right)\right] \\
&\quad -\left[-I+\delta\hat{\lambda}x\right]
\end{aligned}$$

$$\begin{aligned}
&= \hat{\lambda}\left[(1-\hat{\lambda})\nu\left(-I+\frac{\delta}{2}\overline{R}\right)+\hat{\lambda}\left(\delta x-I\right)\right] \\
&\quad -\left[-I+\delta\hat{\lambda}x\right] \\
&= \delta x\left(\hat{\lambda}^2-\hat{\lambda}\right)-I\left(\hat{\lambda}^2-1\right)+\hat{\lambda}(1-\hat{\lambda})\nu\left(-I+\frac{\delta}{2}\overline{R}\right) \\
&= \left(\hat{\lambda}-1\right)\left[\hat{\lambda}\delta x-I\left(1+\hat{\lambda}\right)-\hat{\lambda}\nu\left(-I+\frac{\delta}{2}\overline{R}\right)\right]
\end{aligned}$$

and so we have $\Delta \geq 0$ iff:

$$\hat{\lambda}\delta x - I\left(1+\hat{\lambda}\right)-\hat{\lambda}\nu\left(-I+\frac{\delta}{2}\overline{R}\right)\leq 0$$

$$I \geq \frac{\hat{\lambda}\delta\left[x-\nu\frac{\overline{R}}{2}\right]}{1+\hat{\lambda}-\nu\hat{\lambda}}$$

We notice that the right hand side is increasing, concave in $\hat{\lambda}$. Moreover, we have:

$$\frac{\delta S}{2} < \frac{\frac{y}{x}\delta\left[x-\nu\frac{\overline{R}}{2}\right]}{1+\frac{y}{x}(1-\nu)}$$

and

$$\frac{\delta\overline{R}}{2} > \frac{\delta\left[x-\nu\frac{\overline{R}}{2}\right]}{2-\nu}$$

Case 2: suppose now that $\hat{\lambda} \leq \hat{\lambda}_L$. Then the best alternative to planning ahead is to follow the strategy of no thinking. From the previous lemmas, we check that for $I < \frac{\delta S}{2}$, the best strategy is to invest without any thinking taking place. Fo rintermediate values where $\frac{\delta S}{2} < I < \frac{\delta \overline{R}}{2}$, we have:

$$V_E = \hat{\lambda} \left[\nu \left(-I + \frac{\delta}{2} \overline{R} + \frac{\delta}{2} y \right) + (1-\nu) \hat{\lambda} \left(-I + \frac{\delta}{2} S + \frac{\delta}{2} x \right) \right]$$

and so

$$\begin{aligned} \Delta &= \hat{\lambda} \left[\nu \left(-I + \frac{\delta}{2} \overline{R} + \frac{\delta}{2} y \right) + (1-\nu) \hat{\lambda} \left(-I + \frac{\delta}{2} S + \frac{\delta}{2} x \right) \right] \\ & \quad - [-I + \delta y] \end{aligned}$$

Re-arranging terms as before, we obtain that $\Delta \geq 0$ if and only if:

$$I \geq \frac{\hat{\lambda}\delta \left[x - \nu \frac{\overline{R}}{2} \right]}{1 + \hat{\lambda} - \nu \hat{\lambda}} + \frac{\delta (y - \hat{\lambda} x)}{(1-\hat{\lambda})(1 + \hat{\lambda}(1-\nu))}$$

which may or may not be feasible given that we need $I \leq \frac{\delta}{2} \overline{R}$ and $\hat{\lambda} \leq \hat{\lambda}_L$. It is in particular impossible if $\overline{R} \leq \frac{2xy}{x+y}$, a sufficient condition for no thinking to dominate thinking ahead in that case. Otherwise thinking ahead may be best.

Similarly, when $\frac{\delta \overline{R}}{2} < I$, we have

$$\begin{aligned} \Delta &= \hat{\lambda}^2 [\delta x - I] - (\delta y - I) \\ &> 0 \Leftrightarrow \hat{\lambda} > \sqrt{\frac{\delta y - I}{\delta x - I}} \end{aligned}$$

which may or may not be compatible with $\hat{\lambda} \leq \hat{\lambda}_L$, and the condition on I. It requires in particular that $I > \frac{\delta xy}{x+y}$ so that it is needed that $S > \frac{2xy}{x+y}$. But if this is true, then thinking ahead dominates in this region, and that would also imply that there are some values for which thinking ahead also dominates when $\frac{\delta \overline{R}}{2} > I > \frac{\delta S}{2}$.