

# Experimentation and Approval Mechanisms

Andrew McClellan

April 2, 2017

## Abstract

We study the design of approval rules when experimentation must be delegated to an agent with misaligned preferences. Our motivating example looks at how the FDA can design approval rules a function of the outcome clinical trials. In these clinical trials, the agent (the drug company) must pay a cost for experimentation and may have information about the likelihood that the state is high (the drug is good). We study this question in a dynamic learning framework and look at how the level of commitment the regulator can place on the agent changes the structure of the optimal approval rule, both when the agent has private information about the payoff relevant state and when he does not. When the mechanism must satisfy only ex-ante participation constraints, the optimal approval rule becomes a stationary threshold (similar to the problem with no agency concerns). However, when the mechanism must satisfy interim, participation constraints, the approval threshold will no longer be stationary change over time. We find the optimal approval rule and show that it moves downward monotonically. Surprisingly, the approval threshold only moves downward as a function of the minimum of the regulator's beliefs. When the agent possess private information about the state, we find that the agent with high information may receive a fast-track: his approval threshold is initially low (the fast-track) but takes a jump up if the regulator's beliefs fall too low. These dynamics are to our knowledge new and help us understand how approval rules change over time.

---

I would especially like to thank Debraj Ray and Ennio Stachhetti for their support, guidance and insightful suggestions. I would also like to thank Dhruva Bhaskar, Boyan Jovanovic, Alessandro Lizzeri and Jacopo Perego for helpful conversations and suggestions as well as seminar participants at NYU.

# 1 Introduction

In many real world economic situations, decision makers face a tradeoff between making a decision and waiting for more evidence. For example, when deciding whether or not to approve a drug, the FDA can mandate that companies conduct clinical trials to determine the efficacy and any side-effects the drug may have. Typically, learning in these trials is slow, as the drug may affect the patients health only with a great deal of noise. In deciding when to approve a drug, the FDA must trade off the need for haste (in order to alleviate the suffering of those currently afflicted) and the need to discretion (in order to prevent the use of harmful drugs). Finding the optimal balance between these two has large welfare implications.

Given their importance, the design of clinical trials has developed an extensive literature. One observation, which is missing in the existing literature, is that there is often a misalignment of incentives between those with decision making authority and those with private prior information which may inform the decision. For example, a company may spend a long time developing a drug prior to the start of a clinical trial and will possess a more informed prior about whether the drug is good or not. It is natural to think that the FDA would like to elicit this information from the company. However, the misalignment of incentives will prevent straightforward elicitation: the company, which doesn't internalize the externality imposed by the approval of bad drugs, wants to get its drug approved while the FDA only wants to approve the drug if it is good. This brings us to the mechanism design question: is it possible for the FDA to set an approval rule which incentivizes the companies to truthfully reveal their information?

In this paper, we will study the design of approval mechanisms in a dynamic framework with learning. We look at how a regulator can design stopping and approval decision rules (without monetary transfers) which incentivize an agent to truthfully reveal his private information about some payoff relevant state of nature. The players have misaligned incentives: the regulator only wishes to approve if the state is high and the agent wants approval regardless of the state and must pay the cost of experimentation. Thus the regulator has additional incentives to consider when designing his optimal stopping rule: in addition to the tradeoff between haste and discretion, the regulator must also consider the preservation of the agent's incentives to truthfully reveal information. We also examine how these mechanisms change under different levels of commitment on the part of the agent. We look at two-sided commitment (where the agent can commit to keep experimenting as long as the regulator directs him to) and one-sided commitment (where the agent may quit experimenting and take his outside option at any time). We find that the two commitment structures lead to qualitatively different mechanisms.

Under two-sided commitment, the optimal stopping mechanism takes the form of static threshold rules: approve if the evidence reaches an upper threshold and reject if the evidence reaches a lower threshold. These static thresholds qualitatively mirror the standard single agent optimal stopping problem. We find that the agency problem can easily be turned into a single decision-maker problem (introducing distortion terms for incentive constraints) and solved accordingly.

With one-sided commitment, the story is much different. In addition to satisfying ex-ante participation and incentive constraints, the regulator must also satisfy interim participation constraints after every history. We show that one-sided commitment, threshold rules (as in the case with two-sided commitment) are not optimal: once the agent is about to quit, the regulator has an incentive to lower the approval threshold to keep the agent experimenting. Given the complexity of determining the stopping rule that the agent will use in response to the stopping rule used by the regulator, finding the optimal mechanism can seem daunting. However, we are able to solve for the optimal mechanism by identifying regions over which the regulator uses threshold rules. We find that the stopping rule used by the regulator is highly path dependent but still tractable and is, to our knowledge, novel in the optimal stopping literature.

We first derive the optimal mechanism when there is no private information. Even in this setting, threshold rules are not optimal. More complex stopping rules allow the regulator to incentivize the agent to continue experimenting more than when the regulator uses simple threshold rules. We show how the optimal mechanism can be written as a function of the current belief and the minimum over the realized path of beliefs up to the current time. We show how the approval threshold takes a stochastic, but monotonic, path downwards. Interestingly, this implies that, as a function of the approval time, the probability of Type I error is not constant, as it is in the single decision maker problems. We also are able to show that the assumption of commitment by  $R$  is not crucial: this mechanism can be implemented as an equilibrium even when  $R$  cannot commit to a stopping rule.

When we give private information to the agents, we find that the optimal mechanism takes the form of a “fast-track” menu option. Low types are offered a mechanism which is qualitatively similar to the case with no adverse-selection: the approval threshold is monotonically decreasing. However, high types are given a very different mechanism. They are offered an initial approval threshold which is lower than the initial approval threshold offered to low types. However, they also face a “failure” threshold. If the failure threshold is reached, the project is not rejected but the approval threshold takes a discrete jump upward. This result shows how adverse selection creates a backloading of costly distortion for the high type. By introducing a higher approval threshold, the regulator hurts both his

and the agents payoffs. However, a deviating low type will view this distortion as more likely, thereby creating a wedge in the effect of the distortion on payoffs. This wedge allows the regulator to create separating contracts even without transfers.

In Section 2 we will discuss related literature and then introduce the model in Section 3. Section 4 will cover the optimal mechanism where there are no information asymmetries while Section 5 will derive the optimal mechanism when there are information asymmetries.

## 2 Literature

The setting of our paper ties into a large literature on the problem of hypothesis testing. Wald (1947) is the seminal work on the study of sequential testing and began a rich literature in mathematics and statistics. Peskir and Shiryaev (2006) provide a textbook summary and history of the problem. Moscaroni and Smith (2001) also examine a similar framework but they look at the optimal policy in a large class of sampling strategies. Unlike our paper, this literature focuses on the problem of a single-decision maker. While some papers study the optimal stopping problem under constraints, the participation constraints our problem will impose are new and yield very different solutions.

Our paper is also related to the bandit experimentation literature. Bolton and Harris (1999), Keller, Rady and Cripps (2005), Keller and Rady (2010,2015), Strulovici (2010), Chan et al. (2015) and many others have analyzed the strategic interaction among experimenting agents. Typically, they focus on equilibrium experimentation levels and often find equilibrium strategies in cutoff rules. In our paper, we will endow one player (the regulator) with commitment power, which will drive the optimality of more complex stopping rules.

A recent literature has developed around the incentivization of experimentation in bandit problems. Garfagnini (2011) studies equilibrium levels of experimentation when a principal must delegate experimentation to an agent. Guo (2016), one of the closest papers to our own, looks at a bandit problem in a principal-agent model when the agent possesses private information about the probability that the bandit is “good.” Like our model, Guo finds optimal mechanisms when monetary transfers are infeasible and the agent has private information about a payoff-relevant state of the world. Besides the technical differences between our settings, (Guo examines the optimal mechanism for eliciting information in a bandit model while we consider the optimal mechanism in a stopping problem, and in our model the misalignment between principal and agent preferences is more severe), we consider the case in which the agent has the ability at any time to quit experimenting. Our two-sided commitment model corresponds to Guo’s commitment structure, and like her model, we find that threshold rules are optimal with two-sided commitment. Grenadier

et al. (2015) model a situation in which a principal must elicit an agent's information about the optimal exercise time of an option. Like our model, they study the case when the principal cannot make monetary transfers and the agent has private information (in their case, his payoff to exercising the option).

Kruse and Strack (2015) look at an optimal stopping problem in a principal-agent framework in which the principal sets transfers in order to incentivize an agent to use particular stopping rules. They find that, under some conditions, transfers which only depend on the stopping decision implement cut-off rules and all cut-off rules are implementable by such transfers. Maddsen (2016) also studies a principal-agent stopping problem with transfers in the case of the quickest detection problem.

Orlov et al. (2016) also look at the dynamic revelation of information between a regulator and agent. Unlike our model, they look at the nature of equilibrium when, on top of a public news process, the agent has the ability to design information structures to reveal some private information ala Kamenica and Gentzkow (2011).

Liu, Halac and Kartik (2016a, 2016b) also look at different ways of incentivizing experimentation, both in the framework of a contest and a contract. Our paper differs in that we are not allowing for monetary transfers, and instead look at how the probability of future approval can be used to incentivize agents. The incentivization of experimentation using monetary transfers from a moral hazard viewpoint has also been studied by Bergemann and Hege (1998,2005) and Horner and Samuelson (2013).

The study of the FDA approval process has also been studied theoretically and empirically. Carpenter and Ting (2007) looks at a theoretical model of drug approval when the drug companies are better informed about the state of their drug. They study the resulting equilibria of a discrete time model. They find that the length of experimentation determines the comparative static on the effect of firm size on the amount of Type I and Type II errors. Carpenter (2004) also studied the effect of firm size on regulatory decisions.

Henry and Ottaviani (2013,2015) also look at a model of regulatory approval when learning takes place through a publicly observed Brownian motion. In their model, both the regulator and the agent possess a common prior about the state. They study the deconstruction of the approval process, when the regulator has the ability to approve and the agent has the ability to quit. They find that varying the level of commitment and the possession of authority changes the expected amount of experimentation.

### 3 Model

#### 3.1 Environment

Following our motivating example, we will set the model up as the dynamic interaction between a (female) regulator  $R$  and a (male) agent  $A$ . A project which is up for approval is either good ( $\omega = H$ ) or bad ( $\omega = L$ ). The regulator wants to approve only good projects. The benefit to approving a good project is  $B$  and the loss to approving a bad project is  $K$  (for simplicity, we will take  $B = 1, K = -1$ ). The tension in our model comes from the fact that there is a misalignment of preferences: the agent wants to have the project approved, regardless of the state (e.g., the drug company only cares about whether or not its drug is approved). The agent's payoff to having the project approved is normalized to 1. The game takes place in continuous time and both  $R$  and  $A$  share a common discount factor  $r > 0$ .

	R's Payoffs			A's Payoffs	
	H	L		H	L
Approve	1	-1	Approve	1	1
Reject	0	0	Reject	0	0

We will assume that  $A$  pays a constant flow cost  $c_A$  until the game ends and  $R$  pays a flow cost of  $c_R$ . For simplicity, we assume that  $c_A = c > 0 = c_R$  (none of the results will rely on  $c_R = 0$ , but it makes the analysis a bit simpler).

We model the experiments as the observation of a Brownian motion whose drift depends on whether or not the project is good. While the experimentation is ongoing, the regulator observes outcomes

$$X_t = \mu_\omega t + \sigma W_t$$

where  $W = \{W_t, \mathcal{F}_t, 0 \leq t < \infty\}$  is a standard one-dimensional Brownian motion<sup>1</sup> on the state space  $(\Omega, \mathcal{F}, P)$  and  $\mu_L = -\mu < 0 < \mu = \mu_H$ . For example, if the project is a new drug, we can think of  $X_t$  as the health of the patients during a clinical trial. If the state is a worker's skill, then we can think of  $X_t$  as the worker's cumulative output.

By observing  $X_t$ ,  $R$  can update her belief about the state. After observing  $X_t$ , the player's update their beliefs by Bayes rule to

---

<sup>1</sup>Which implies that  $W_0 = 0$  so  $X_0 = 0$ .

$$\pi_t = \frac{\pi_0 f_t^H(X_t)}{\pi_0 f_t^H(X_t) + (1 - \pi_0) f_t^L(X_t)}$$

To simplify the belief updating procedure, we note that we can write the beliefs in terms of log-likelihoods<sup>2</sup> i.e.

$$Z_t = \log\left(\frac{\pi_t}{1 - \pi_t}\right)$$

Putting in our terms for  $\pi_t$ , we have

$$\begin{aligned} Z_t &= \log\left(\frac{\pi_0}{1 - \pi_0}\right) + \log\left(\frac{f_t^H(X_t)}{f_t^L(X_t)}\right) \\ &= Z_0 + \frac{\phi}{\sigma} X_t \end{aligned}$$

where  $\phi = \frac{2\mu}{\sigma}$ , which is the signal-to-noise ratio (this describes how informative the signals are). Since beliefs and the evidence level are isomorphic, we will use them interchangeably in the following sections. The change in  $Z_t$  is then given by

$$dZ_t = \frac{\phi}{\sigma} dX_t$$

We define  $\mathcal{F}_t^X = \sigma((X_s, Y_0) : 0 \leq s \leq t)$  (where  $Y_0 \sim U[0, 1]$  time 0 is used simply to allow for randomization) to be the augmented natural filtration. A history  $h_t = \omega|_{[0, t]}$  is the realization of a path of  $X_t$  and  $Y_0$ .

Note that  $R$  receives positive utility from approving at belief  $Z_t$  if and only if  $Z_t \geq 0$ . We will refer to  $Z_t = 0$  as  $R$ 's myopic cutoff point-i.e., when she would approve if she were myopic. Where this myopic cutoff point is determined by the specific payoff of approving and the state. None of the results will depend on the specific payoffs. Additionally, we could introduce lower constraints on the approval rule (such as approval can only happen at  $Z_t > Z_c$  for some  $Z_c$ ) without affecting the main results. Situations such as this may happen when  $R$  faces political constraints that must ensure it maintains a certain standard of safety.

### 3.1.1 Remarks

We make several simplifying assumptions in the model, which we motivate below

---

<sup>2</sup>We subsequently abuse notation by referring to  $Z_t$  as beliefs

- **Slow Learning:** We choose to model the news process as Brownian motion for both tractability and its similarity to real-world applications. In our motivating example where  $X_t$  corresponds to patient's health during a clinical trial, Brownian motion reflects the gradual nature of learning and the noisiness of health. Even when administered good drugs, patients health will still sometimes decline. However, the drift of patients health should be positive for good drugs (i.e.  $\mu_H > 0$ ). Additionally, the use of Brownian motion ties into a rich statistics literature on the design of adaptive clinical trials and hypothesis testing (e.g. Peskir and Sharyaev (2006) for a textbook treatment).
- **Public News:** We assume that the signal is publicly observable to both  $R$  and  $A$ . This assumption is satisfied in many situations. For example, the FDA can require companies to publicly register and continuously report the outcome of the trial. Assuming the news process is public allows to avoid  $R$  and  $A$ 's beliefs diverging, which would make the model intractable.
- **Payoffs:** We assume that only  $A$  pays a flow cost. This might correspond to the cost of administering the trial (e.g., producing drugs, paying doctors to administer the drugs), which are not small and are important economic determinants of companies' testing decisions see (DiMasi (2014)).

### 3.2 Mechanism

We now move on the mechanism that  $R$  must design. We will assume that transfers are infeasible. For example, the FDA cannot make transfers to incentivize companies to reveal their private information. Instead, they have the ability to set approval standards. Our question of interest is to understand how they can optimally design approval standards to elicit the private information of the firm. More formally, we allow  $R$  to design a stopping mechanism:

**Definition 1.** A *stopping mechanism* is a pair  $(\tau, d_\tau)$  such that  $\tau$  is an  $\mathcal{F}_t^X$ -measurable stopping rule and  $d_\tau$  is an  $\mathcal{F}_t^X$  measurable decision rule which takes value 0 or 1.

We will denote  $\mathbb{T}$  to be the set of  $\mathcal{F}_t^X$ -measurable stopping rules and  $\mathbb{D}$  to be the set of  $\mathcal{F}_t^X$ -measurable decision rules.

We will endow  $R$  with perfect commitment power, allowing us to focus on direct revelation stopping mechanism. The decision to approve or reject is irrevocable<sup>3</sup>. The

---

<sup>3</sup>This simplifies the model. In terms of real world applicability, we note that continued monitoring post-approval by the FDA is typically very weak and rarely results in approval revocation (see Carpenter (20??))



utility of  $R$  is given by

$$J(\tau, d_\tau, Z_0) = \mathbb{E}[e^{-r\tau}(\pi_\tau - (1 - \pi_\tau))d_\tau | Z_0] = \mathbb{E}[e^{-r\tau} \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} d_\tau | Z_0]$$

and the utility of  $A$  is given by

$$V(\tau, d_\tau, Z_0) = \mathbb{E}[e^{-r\tau} d_\tau - \int_0^\tau e^{-rt} c dt | Z_0] = \mathbb{E}[e^{-r\tau} d_\tau - \frac{1 - e^{-r\tau}}{r} c | Z_0]$$

Before moving on the general analysis, we first define some notation that will be useful in the following analysis. We begin with a salient subclass of mechanisms, in which the regulator approves if her beliefs ever reach  $B$  and rejects in her beliefs ever reach  $b$ .

**Definition 2.** A *static threshold* is a pair number  $(B, b) \in \mathbb{R}^2$  such that  $b < B$  and  $\tau = \inf\{t : X_t \notin (b, B)\}$ .

We will refer to  $B$  as the static approval threshold and  $b$  as the static rejection threshold. We will use the notation  $\tau(c) = \inf\{t : X_t = c\}$  where  $c \in \mathbb{R}$ . Then a static threshold mechanism  $\tau$  can be written  $\tau = \tau(B) \wedge \tau(b)$ .

This is a focal class of stopping rules, are tractable and easily implemented. Suppose that  $R$  approves the drug if the evidence  $Z_t$  ever reaches  $B$  and rejects the drug if ever reach  $b$ . An important determinant of the utility from these threshold mechanisms is the expected discounted probability that beliefs cross the threshold  $B$  before crossing  $b$ . The formula for this discounted probability (see Stokey (2009)) when the state is good is given by

$$\Psi(B, b, Z) = \mathbb{E}[e^{-r\tau} d_\tau | \omega = H, Z_0 = Z] = \frac{e^{-R_1(Z-b)} - e^{-R_2(Z-b)}}{e^{-R_1(B-b)} - e^{-R_2(B-b)}}$$

and the discounted probability that the beliefs cross  $b$  before ever crossing  $B$  if  $\omega = H$  is

$$\psi(B, b, Z) = \mathbb{E}[e^{-r\tau} (1 - d_\tau) | \omega = H] = \frac{e^{R_2(B-Z)} - e^{R_1(B-Z)}}{e^{R_2(B-b)} - e^{R_1(B-b)}}$$

where  $R_1 = \frac{1}{2}(1 - \sqrt{1 + \frac{8r}{\mu'}})$  and  $R_2 = \frac{1}{2}(1 + \sqrt{1 + \frac{8r}{\mu'}})$  for  $\mu' = \frac{2\mu}{\sigma^2}$ .

We will generally drop the dependence of  $\Psi, \psi$  on  $B, b, X$  when the choice of  $B, b, X$  is clear. Notationally, we will define  $\Psi_i := \Psi(B(Z_i), b(Z_i), X_i)$ ,  $\psi_i := \psi(B(Z_i), b(Z_i), X_i)$ . For derivatives, we will also use the notation  $\Psi_b := \frac{\partial \Psi}{\partial b}$  and  $\Psi_B := \frac{\partial \Psi}{\partial B}$  (with similar notation for the derivatives of  $\psi$  and the derivatives of  $\Psi, \psi$  with respect to  $B, b$ ).

Doing a bit of algebra (see Henry and Ottaviani (2015)) allows us to show that the discounted probability that  $B$  is crossed before  $b$  if  $\omega = L$  is

$$\Psi(B, b, Z)e^{Z-B}$$

and the the discounted probability that  $b$  is crossed before  $B$  if  $\omega = L$  is

$$\psi(B, b, Z)e^{Z-b}$$

This allows us to rewrite the utility of  $R, A$  when  $(\tau, d_\tau)$  takes a threshold form when  $X_0 = 0$

$$\begin{aligned} J(\tau, d_\tau, Z_0) &= \frac{1}{1 + e^{Z_0}} (\Psi(B, b, Z_0)(e^{Z_0} - e^{-B})) \\ V(\tau, d_\tau, Z_0) &= -\frac{c}{r} + \frac{1}{1 + e^{Z_0}} (\Psi(B, b, Z_0)(e^{Z_0} + e^{-B})(1 + \frac{c}{r}) + \frac{c}{r} \psi(B, b, Z_0)(e^{Z_0} + e^{0-b})) \end{aligned}$$

In general, we allow for a wide range of stopping mechanisms which may consist of continuation regions which may be very complex and will depend on the history of  $X_t$  up until time  $t$ . We define a dynamic analogue of the static threshold mechanism below.

**Definition 3.** A *dynamic threshold* is a pair of mappings  $(B, b) : \mathcal{F}_t^X \rightarrow \mathbb{R}^2$  such that after history  $h_t$ ,  $b(h_t) < B(h_t)$  the process is stopped if  $X_t \notin (b(h_t), B(h_t))$ .

A dynamic threshold is a general stopping policy which can depend on the history of realized beliefs in many different ways. While static thresholds are also dynamic thresholds, there are many dynamic thresholds which are not static thresholds-e.g., a deadline policy (in which  $R$  waits until time  $T$  and approves if  $X_T \geq B_T$  and rejects otherwise) is a dynamic threshold policy but not a static one.

## 4 Symmetric Information

We begin the analysis by studying what the optimal mechanism looks like when there is symmetric information-i.e., both  $A$  and  $R$  share the same prior when the news process begins. Studying the symmetric information case will be useful both for finding the optimal mechanism with asymmetric information and is of independent interest<sup>4</sup>. It extends the canonical hypothesis testing model, which is well-studied in single decision-maker problems, into a mechanism design framework. This new multi-agent framework will yields dynamics far different from that of single decision maker problems. How different depends on the level of commitment we put in the model. We look at the tradeoff between waiting for evidence and incentivizing  $A$  to keep experimenting and derive a novel and tractable optimal mechanism.

---

<sup>4</sup>Henry and Ottaviani (2015) study a model with the same payoff structure as ours, but restrict attention to the class of static threshold mechanisms.

## 4.1 First-Best

We begin by solving for the first best solution for  $R$ . Note that  $R$  has no experimentation costs. Therefore he will never find it optimal to reject: since the news process will never lead  $R$  to know for sure that the drug is bad (i.e.,  $Z_t$  can never reach  $-\infty$ ), then the option value of continuing to experiment is always strictly positive. We can also note that  $R$ 's preferences are time-consistent, her first-best policy must be a threshold rule with  $b = -\infty$ . Clearly she must approve an some interior  $B < \infty$ . If we write out his utility for a fixed  $B$ , we can see that

$$\lim_{b \rightarrow -\infty} \Psi \frac{e^Z - e^{Z-B}}{1 + e^Z} = e^{R_1 B} \frac{e^Z - e^{Z-B}}{1 + e^Z}$$

Taking the derivative with respect to  $B$ , we get a first-order condition

$$\begin{aligned} 0 &= R_1(1 - e^{-B}) + e^{-B} \\ \Rightarrow B &= -\log\left(\frac{R_1}{-R_2}\right) \end{aligned}$$

Which as we should expect implies that the optimal threshold choice of  $R$  is invariant to the current belief. This threshold mechanism  $(B, b) = (-\log(\frac{R_1}{-R_2}), -\infty)$  is the first-best mechanism for  $R$  (and will be optimal in the space of all stopping mechanisms).

## 4.2 Two-Sided Commitment

We now move to the case when  $R$  must consider participation and incentive constraints for  $A$ . We begin by assuming that once  $A$  has selected a stopping rule, then experimentation takes place until  $R$  either approves or rejects. This means that  $R$  can make  $A$  commit to continue experimentation and  $A$  cannot quit early-i.e. both sides can sign a binding contract that is enforceable.

**Definition 4.** A mechanism has *two-sided commitment* if once  $A$  has selected  $(\tau, d_\tau)$ , experimentation continues until  $\tau$ .

The mechanism design problem faced by  $R$  is given by

$$\begin{aligned} \sup_{(\tau, d_\tau)} \quad & \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_0] \\ \text{subject to} \quad & \\ P(\tau, d_\tau) \quad & \mathbb{E}[e^{-r\tau} (d_\tau + \frac{c}{r}) | Z_0] - \frac{c}{r} \geq 0 \end{aligned}$$

where  $P(\tau, d_\tau)$  is a participation constraint that ensures that the agent finds it optimal to agree to mechanism.

Our problem takes the form of a constrained optimal stopping problem. A robust finding from the single decision-maker problem is the optimality of static-threshold rules (e.g. Wald (1947), Moscaroni and Smith (2001)). However, it is not ex-ante obvious that static-threshold rules will be optimal here, our first result establishes that  $R$  will offer a stationary approval and rejection threshold.

**Proposition 1.** *The solution to the symmetric information problem with two-sided commitment takes the form of a static-threshold policy. If  $b \neq -\infty$ , then the optimal approval and rejection thresholds  $(B, b)$  are the solution to the following equations:*

$$\frac{\Psi_B(e^{Z_0} - e^{-B}) + e^{-B}\Psi}{\Psi_B(e^{Z_0} + e^{-B}) - e^{-B}\Psi + \frac{c}{r+c}\psi_B(e^{Z_0} + e^{-b})} = \frac{\Psi_b(e^{Z_0} - e^{-B})}{\Psi_b(e^{Z_0} + e^{-B}) - \frac{c}{r+c}\psi e^{-b} + \frac{c}{r+c}\psi_b(e^{Z_0} + e^{-b})}$$

$$\Psi(e^{Z_0} + e^{-B}) + \frac{c}{r+c}\psi(e^{Z_0} + e^{-b}) = \frac{c}{r+c}(1 + e^{Z_0})$$

if  $b = -\infty$ , then  $B = \log(\frac{R_2}{R_1})$ .

*Proof.* See Appendix. □

This result establishes that the solution under two-sided commitment is qualitatively the same as in the single decision-maker.

One question of interest is how the approval policies differ across the levels of commitment. Our first result on this question establishes that approval takes place at a lower level in two-sided commitment.

**Corollary 1.** *If  $B_T^*$  is the solution to the optimal two-sided commitment problem and  $B_F^*$  is the first-best approval level for  $R$ , then  $B_F^* \geq B_T^*$ .*

Because there is rejection in the two-sided commitment case but not in the first-best solution, this means that the option value of experimentation is lower in the two-sided commitment case. Since experimentation will continue as long, the value of experimentation decreases relative to approval. This decrease in the option value leads to a lower approval standard.

Interestingly, the choice of  $B, b$  will depend on the initial  $Z_0$ . This differs from the single-decision maker problem and the one-sided commitment solution. Unlike in the single-decision maker problem, the participation constraint at  $t = 0$  means that time zero beliefs affect the possible selection of mechanisms.

### 4.3 One-Sided Commitment

In many applications, the assumption of two-sided commitment is unreasonable. While the FDA can commit to approval standards, they do not possess the authority to force a company to continuing experimentation. If over the course of the trial the company becomes pessimistic that the drug will ever be approved, the company may decide to cut their losses and end the trial early. The assumption that the FDA can force them to continue experimenting is beyond the scope of the agency's authority. We can think of this as the analogue of a "no forced service" assumption in a standard principal-agent model, in which the principal can commit to a contract but the agent cannot be prevented from taking an outside option. We call this commitment structure one-sided commitment.

**Definition 5.** A mechanism has **one-sided commitment** if once  $A$  has selected  $(\tau, d_\tau)$ , after any history  $h_t$ ,  $A$  can quit experimenting and take an outside payoff  $Q$ .

For simplicity, we will assume that  $Q = 0$  (since quitting experimentation and rejection both lead to non-approval, it seems natural that the payoffs will be the same). Since  $A$  has the ability at any time to take an outside option, we must reformulate what participation constraints mean in the environment with one-sided commitment. Under two-sided commitment, we only had to ensure that the expected payoff at time  $t = 0$  was weakly positive. With one-sided commitment, we must ensure that  $A$ 's continuation payoff is weakly positive at all  $t$  and histories  $h_t$  until  $R$  ends experimentation. We call the constraints this property imposes to be dynamic participation constraints.

**Definition 6.** A mechanism  $(\tau, d_\tau)$  satisfies **dynamic participation constraints** if  $A$  after any history  $h_t$ , the expected continuation to  $A$  from  $(\tau, d_\tau)$  is non-negative:

$$\forall h_t, \quad \mathbb{E}[e^{-r(\tau-t)}(d_\tau + \frac{c}{r})|Z_t, h_t] - \frac{c}{r} \geq 0$$

Because there is a participation constraint for each history, writing out all the constraints is infeasible. Another way of stating the dynamic participation constraint is to say that  $A$  never finds it strictly optimal to quit. Following this idea, we can write  $R$ 's problem of choosing a mechanism which satisfies dynamic participation constraints as

$$\begin{aligned} (SM) \quad & \sup_{(\tau, d_\tau)} \quad \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_0] \\ & \text{subject to} \\ DP : \quad & \sup_{\tau' \in \mathbb{T}} \mathbb{E}[e^{-r(\tau \wedge \tau')} (d_\tau \mathbb{1}_{\tau' > \tau} + \frac{c}{r}) | Z_0] \leq \mathbb{E}[e^{-r\tau} (d_\tau + \frac{c}{r}) | Z_0] \end{aligned}$$

where  $DP$  is the dynamic participation constraint. We refer to  $\tau'$  as  $A$ 's quitting rule.  $DP$  implies that for any quitting rule that  $A$  might use, the payoff to potentially quitting early is weakly less than letting  $R$  decide when to end experimentation. We can think of this as an obedience constraint:  $A$  must find it optimal to follow  $R$ 's recommendation to keep experimenting.

This specification of  $DP$  is just a rewritten version of the definition of dynamic participation constraints from an ex-ante perspective. By looking at the supremum over all  $\tau' \in \mathbb{T}$  on the left-hand side, we are implicitly including the dynamic participation constraints. To see that the two are equivalent, note that if there was such a history that  $A$  had a strictly negative expected continuation payoff, then the quitting rule

$$\tau' = \inf\{t : \mathbb{E}[e^{-r(\tau-t)}(d_\tau + \frac{c}{r})|Z_t, h_t] = 0\}$$

would lead to a strictly higher payoff, violating  $DP$ . Similarly, if  $(\tau, d_\tau)$  satisfies dynamic participation constraints, then it also must satisfy  $DP$  since  $A$  would never be better off quitting when he has a non-negative continuation payoff.

Note also that we are specifying in  $DP$  that  $A$  never quits early. Because  $R$  has full commitment power, this is without loss of generality: if a mechanism allows  $A$  to quit after any history  $h_t$ , then we could specify another mechanism in which  $R$  rejects at the same moment that  $A$  does. This will not change any incentives for  $A$  to quit earlier than time  $t$  and hence the expected payoff to  $R$  from the two mechanisms will be the same.

**Lemma 1.** *For any stopping mechanism  $(\tau, d_\tau)$  which satisfies dynamic participation constraints, there exists another stopping mechanism  $(\tilde{\tau}, \tilde{d}_\tau)$  that delivers the same payoff to  $R$  and under which  $A$  never quits early.*

Given the previous result that static threshold mechanisms are optimal under two-sided commitment, they seem to be a natural guess for the form that the optimal mechanism will take. However, we will show that the optimal mechanism will never take this form. We begin by presenting a simple example of the non-optimality of static threshold rules.

Suppose that  $R$  is using a static approval threshold of  $B_1 > 0$ . Since  $R$  would always like to keep experimenting, he will never reject the project before the agent decides to quit. Let  $b_Z^*(B)$  be the value at which  $A$  will choose to quit experimenting<sup>5</sup> when  $R$  uses a static threshold of  $B$  and the current beliefs are  $Z$ .

$$b_Z^*(B) = \operatorname{argmax}_b V(B, b, Z) = \Psi(B, b, Z) \frac{e^Z(1 + e^{-B})}{1 + e^Z} \left(1 + \frac{c}{r}\right) + \frac{c}{r} \psi(B, b, Z) \frac{e^Z(1 + e^{-b})}{1 + e^Z}$$

---

<sup>5</sup>We show in Lemma 4 that  $A$  will find it optimal to quit using a static-threshold strategy when the approval strategy is also a static-threshold strategy

Note that  $b_Z^*$  doesn't depend on  $Z$ ; this is a result of the time consistency of  $V$  so that the optimal choice of a point at which to quit doesn't depend on the current level of  $Z$ . We can show that  $\frac{db_Z^*(B)}{dB} > 0$ -i.e.,  $A$  will want to quit earlier when the approval threshold is higher. This is intuitive since the probability of reaching the approval threshold is lower and the expected costs are higher.

Let Mechanism 1 by a static-threshold mechanism  $(B_1, b_Z^*(B_1))$ . The expected payoff to  $R$  from this mechanism will be

$$\Psi(B_1, b_Z^*(B_1), Z_0) \frac{e^{Z_0}(1 - e^{-B_1})}{1 + e^{Z_0}}$$

Now consider Mechanism 2, in which  $R$  uses an approval threshold of  $B_1$  until either  $Z_t = B_1$  or  $Z_t = b_Z^*(B_1)$ . If  $Z_t$  reaches  $b_Z^*(B_1)$ , then instead of rejecting,  $R$  lowers the approval threshold to  $\alpha B_1$  for some  $\alpha < 1$  such that  $b_Z^*(B_1) < \alpha B_1$ . Note that  $A$  will now only quit experimenting if the evidence reaches  $b_Z^*(\alpha B_1)$  (where  $Z_1 = Z_0 + \frac{\phi}{\sigma} b_Z^*(B_1)$  are the beliefs when the evidence reaches  $b_Z^*(B_1)$ ), since the lowering of the approval threshold strictly incentivizes  $A$  to keep experimenting. Under this new policy, the expected payoff to  $R$  from Mechanism 2 is

$$\Psi(B_1, b_Z^*(B_1), Z_0) \frac{e^{Z_0}(1 - e^{-B_1})}{1 + e^{Z_0}} + \psi(B_1, b_Z^*(B_1), Z_0) \frac{e^{Z_0}(1 + e^{b_Z^*(B_1)})}{1 + e^{Z_0}} \Psi(\alpha B_1, b_Z^*(\alpha B_1), b_Z^*(B_1)) \frac{e^{Z_1}(1 + e^{b_Z^*(B_1) - \alpha B_1})}{1 + e^{Z_1}}$$

Breaking down the above payoff, the expected payoff if  $B_1$  is reached before  $b_Z^*(B_1)$  is the same as in the original policy i.e.

$$\Psi(B_1, b_Z^*(B_1), Z_0) \frac{e^{Z_0} - e^{-B_1}}{1 + e^{Z_0}}$$

However, because she doesn't reject yet in Mechanism 2,  $R$  receives an additional payoff conditional on the evidence reaching  $b_Z^*(B_1)$  before  $B_1$ , which is given by

$$\Psi(\alpha B_1, b_Z^*(\alpha B_1), b_Z^*(B_1)) \frac{e^{Z_1}(1 - e^{b_Z^*(B_1) - \alpha B_1})}{1 + e^{Z_1}}$$

This is multiplied by  $\psi(B_1, b_Z^*(B_1), Z_0) \frac{e^{Z_0}(1 + e^{b_Z^*(B_1; Z_0)})}{1 + e^{Z_0}}$  (the discounted probability that beliefs hit  $b_Z^*(B_1)$  before  $B_1$ ). Note that  $\Psi(\alpha B_1, b_Z^*(\alpha B_1)) \frac{e^{Z_1}(1 - e^{b_Z^*(\alpha B_1)})}{1 + e^{b_Z^*(B_1)}}$  is strictly positive. Therefore Mechanism 2 yields a higher payoff for  $R$ . Since the choice of  $B_1$  was arbitrary, we can see that static-threshold mechanisms are not optimal.

Heuristically,  $R$  is being too stubborn by sticking to the static threshold  $B_1$ . Once the evidence has gone low enough,  $R$  would be better off by decreasing his approval threshold

a bit in order to keep  $A$  experimenting: conditional on the evidence reaching  $b_Z^*(B_1)$ ,  $R$  can achieve a positive continuation value by “cutting some slack” and lowering the approval threshold some, thereby increasing  $A$ ’s value of experimentation. An important fact to note is that since  $A$  can only make a single irreversible choice to quit, lowering the approval threshold if the evidence gets to  $b_Z^*(B_1)$  doesn’t change  $A$ ’s incentives to quit when the evidence is still above  $b_Z^*(B_1)$ .

Once we have moved out of the realm of threshold rules, conjecturing the form that the optimal policy will take is difficult. Because the space of stopping rules is so large, it is not possible to immediately identify a class of mechanisms that the optimal policy will lie in or know if the optimal policy is feasible to derive. The key tradeoff will be between keeping a discerning approval threshold and providing incentives for  $A$  to keep experimenting. Our first main result establishes what the optimal mechanism is and shows that it takes a relatively simple form. In the interest of keeping notation consistent, we will describe the mechanism in terms of  $X_t$  rather than  $Z_t$ . We define an equivalent version of  $b_Z^*$  for the process measured in terms of  $X_t$  as

$$b^*(B) := \operatorname{argmax}_b V(\tau(B), \tau(b), 0)$$

We get that the optimal mechanism turns out to depend on the realized path of  $X_t$  only through the current minimum of the evidence path  $M_t^X := \min\{X_s : s \in [0, t]\}$  and consists of two regimes:

- **Stationary Regime:** The mechanism begins with a static approval threshold  $B_1$  which lasts until  $X_t$  reaches  $B_1$  or  $b^*(B_1)$ .
- **Incentivization Regime:** Once  $X_t$  first hits  $b^*(B_1)$ , the stopping rule is given by a dynamic approval threshold  $B(M_t^X)$  which decreases as  $M_t^X$  decreases in order to incentivize  $A$  to keep experimenting when beliefs get too low.

As in the example above,  $R$  decreases the current threshold in order to incentivize  $A$  to keep experimenting; the decrease is gradual, just enough to keep  $A$  from quitting.

Define

$$\underline{B}(X) := \min\{B : b^*(B) = X\}$$

i.e.  $\underline{B}(X)$  is the lowest static approval threshold such that  $A$  would quit optimally when the evidence reaches  $X$ . Formally, the result is given in the following theorem.

**Theorem 1.** *The optimal stopping mechanism under symmetric information is given by the stopping rule  $\tau = \tau(B(M_t^X)) \wedge \tau(b^*(0))$  and  $d_\tau = \mathbb{1}(X_\tau = B(M_t^X))$  where  $B(M_t^X)$  is defined as*



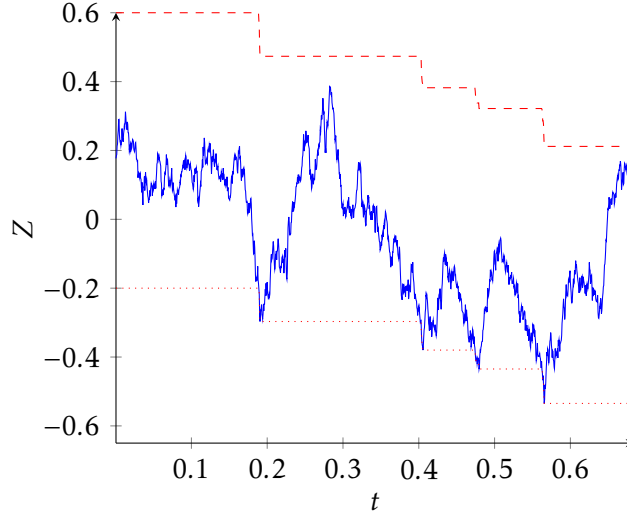


Figure 2: The graph above corresponds to the changing approval threshold for a particular realization of  $X_t$ . The upper dashed line corresponds to the current approval threshold. This approval threshold will stay at the same level until  $X$  crosses the current minimum of the process, which is given by the bottom dashed line.

$$B(M_t^X) = \begin{cases} B^1 & M_t^X \in [b^*(B^1), 0] \\ \underline{B}(M_t^X) & M_t^X \in [\underline{b}(0), b^*(B^1)) \end{cases}$$

where  $B^1 := \operatorname{argmax}_B J(B, \underline{b}, X_0)$ ,  $\underline{b}(Z_0) := -Z_0 \frac{\sigma}{\phi}$

When the  $M_t^X > b^*(B^1)$ , we will say that the mechanism is in a stationary regime (more generally, we will call any area in which the approval threshold is locally static a stationary regime). When  $M_t^X < b^*(B^1)$ , we will say the mechanism is in a dynamic regime. We note several features of the optimal mechanism:

- **Monotonicity:** The approval threshold only drifts downward: the approval threshold only changes in order to provide incentives to keep  $A$  from quitting, which can only happen when the approval threshold decreases. The times at which the current approval threshold decreases are stochastic (since they are a function of  $M_t$ ).

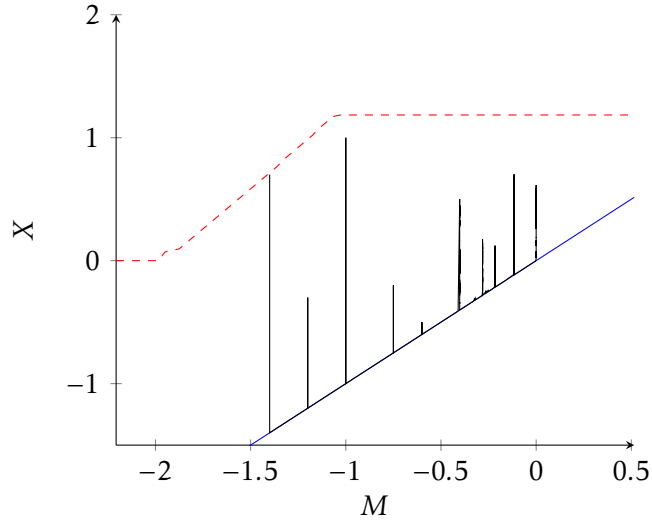


Figure 3: The dashed line gives the approval threshold as a function of  $M$  and the solid line marks the 45 degree line. The dashed line is initially constant in  $M$  during the stationary regime while it decreases in  $M$  for the incentivization regime. The lines coming up from the 45 degree line illustrate a sample path of  $X$  which is approved when  $X = 0.7$ .

- **Agent Indifference:** Whenever the evidence level is at  $X_t = M_t^X$ , the agent will be indifferent between quitting and continuing.  $R$  would like to keep the approval threshold from decreasing and will thus wait until  $A$  is indifferent between quitting and continuing, which occurs at  $X_t = M_t^X$ .
- **Starting Belief Invariance:** Because the level of evidence is isomorphic to beliefs, we can alternatively write the approval threshold in terms of what beliefs  $R$  approves at. If we do this transformation, then the optimal mechanism is invariant to what initial beliefs  $Z_0$  are. This property, which is common in single-decision maker problems, is absent if we were to restrict attention to static threshold rules (see Henry and Ottaviani (2015)).

The rest of this section will be devoted to sketching out the ideas of the proof (a full proof can be found in the Appendix). The optimal mechanism describes the best way to balance out the desire for more learning and the need to incentivize  $A$  to experiment longer. The problem of agent incentivization in continuous time has developed a growing literature

in recent years. Our approach differs from the standard continuous-time approach (e.g., Sannikov (2007) when transfers are feasible and Fong (2007) where transfers are not feasible) where agent-continuation payoffs are formulated as a state variable in an HJB equation. Because both  $A$  and  $R$  are not perfectly informed, the HJB approach would require carrying both a state variable of agent continuation and current beliefs, turning the problem into that of solving a partial differential equation. Instead, we decompose the mechanism design problem into several regions over which we use Lagrangian techniques to find the optimal mechanism. The Lagrangian approach allows to more easily derive the qualitative features of the optimal mechanism and allow us to restrict attention to a class of mechanism over which we can derive the quantitative features. In contrast to the model of Sannikov (2007), the moral hazard component is much simpler in our model (in our model the agent can only decide at each point in time whether or not to quit), but the tools available to the mechanism designer are more sparse (since we rule out transfers and the decisions of  $R$  and  $A$  are irreversible). The Lagrangian approach will also be well-suited to the asymmetric information case.

We would like to convert this constrained problem  $SM$  (our primal problem) into an unconstrained form (our dual problem) using Lagrangian techniques. The key technical difficulty lies in the fact that when checking  $DP$ , we must consider all possible quitting rules  $\tau'$  which  $A$  might use. For an arbitrary stopping rule  $(\tau, d_\tau)$ , we might try to solve for the optimal  $\tau'$  which  $A$  would use. However, even for simple time-dependent stopping rules (e.g. using a deadline), solving for  $\tau'$  is difficult and cannot be calculated in closed-form. Given the richness of the set of available  $(\tau, d_\tau)$ , which may be history-dependent, solving for  $\tau'$  is infeasible. Moreover, to use Lagrangian techniques we will need to restrict attention to a finite number of constraints. This means we will need to find a finite number of quitting rules which will approximate the set of binding constraints. Given the dimensionality of the space of quitting rules, it is not immediately clear how to do this.

To make some progress, we will have to restrict attention to a class of quitting rules, which we call threshold quitting rules.

**Definition 7.** A *threshold quitting rule* is a number  $\tau(X_i) \in \mathbb{R}$  such that  $A$  quits experimenting if and only if  $X_t \leq \tau(X_i)$ .

In order to use the Lagrangian approach, we need to restrict attention to a finite number of constraints. Let  $XT_N = \{X_i\}_{i=0}^N$  such that  $X_{i+1} = X_i + \frac{Y}{N}$  for some  $Y \in \mathbb{R}_-$  small enough (we will verify later that the mechanism ends in rejection before  $X_t$  reaches  $Y$ ).

We can rewrite the mechanism design problem as  $RSM_N$

$$\begin{aligned}
& \text{sup}_{(\tau, d_\tau)} \quad \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_0] \\
& \text{subject to} \\
& \text{RDP}_N \quad \forall X_T \in XT_N \quad \mathbb{E}[e^{-r(\tau \wedge \tau')} (d_\tau \mathbb{1}_{\tau' > \tau} - \frac{c}{r}) | Z_0] \leq \mathbb{E}[e^{-r\tau} (d_\tau + \frac{c}{r}) | Z_0]
\end{aligned}$$

It is important to emphasize that it is not obvious that dropping non-threshold constraints is without loss. For many stopping policies  $R$  could use, the best response of  $A$  will not be to use a threshold policy. For example, if  $R$  were to wait until date  $T$  and approve if and only if  $X_T > B$ , then the optimal quitting rule  $A$  would use would in fact not be a threshold policy but would be a time-dependent curve  $\tau' = \inf\{t : X_t = f(t)\}$ . Since we allow for arbitrarily complex history-dependent stopping rules, the quitting rule which is  $A$ 's best response to an arbitrary  $\tau$  will also be a complex history-dependent quitting rule.

Since we have dropped a number of constraints (i.e., all non-threshold quitting rules), the solution to  $RSM_N$  will provide an upper bound on the value to  $R$  of the full problem  $SM$ . Thus if our solution to  $RSM_N$  satisfies the constraints of  $SM$ , then the solution to  $RSM_N$  solves the  $SM$ . Note that we are not restricting the solution of  $RSM$  to be a threshold policy. Instead, we are only checking that  $A$  has no incentive to deviate to a threshold quitting rule rather than following  $R$ 's recommendation.

We can now use Lemma 18 from the Appendix in order to transform our primal problem  $RSM_N$  into the dual problem. We can construct an associated Lagrangian with Lagrange multipliers  $\{\lambda(X_i)\}_i \in \mathbb{R}^n$

$$\begin{aligned}
\mathcal{L} = & \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_0)] \\
& + \sum_i \lambda(X_i) [\mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} (\mathbb{1}(\tau < \tau(X_i)) d_\tau + \frac{c}{r}) | Z_0] - \mathbb{E}[e^{-r\tau} (d_\tau + \frac{c}{r}) | Z_0]]
\end{aligned}$$

For an appropriate choice of  $\{\lambda(X_i)\}_i$ , the solution to the associated Lagrangian will solve the primal problem  $RSM'$  and we will have complementary slackness conditions

$$\forall i, \quad \lambda(X_i) [\mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} (\mathbb{1}(\tau < \tau(X_i)) d_\tau + \frac{c}{r}) | Z_0] - \mathbb{E}[e^{-r\tau} (d_\tau + \frac{c}{r}) | Z_0]] = 0$$

We will decompose the problem into the region in which  $A$  has positive continuation value and the region in which  $A$  has a continuation value of zero. We denote this continuation value for  $R$  of the mechanism which delivers a continuation value of zero to  $A$  by  $H(X_t)$ , which is defined formally as

$$\begin{aligned}
H(X_t) &= \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_t] \\
&\text{subject to} \\
RDP(\tau, d_\tau) &\quad \forall X_i \in XT_N \text{ s.t. } X_i < X_t \quad \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} (d_\tau \mathbb{1}(\tau(X_i) > \tau) - \frac{c}{r}) | Z_t] \leq \mathbb{E}[e^{-r\tau} (d_\tau + \frac{c}{r}) | Z_t] \\
PK(0) &\quad \mathbb{E}[e^{-r\tau} (d_\tau + \frac{c}{r}) | Z_t] = \frac{c}{r}
\end{aligned}$$

where we have added a promise keeping constraint  $PK(0)$  which ensures that the expected utility of  $A$  for continuing until  $R$  rejects is zero. We postpone deriving the mechanism which delivers  $H$  or what values the function  $H$  takes.

With the problem in an unconstrained form, we can use techniques from the single decision-maker stopping problem to find the optimal policy which solves the dual problem. The following Lemma allows us to establish the optimality of a “local” static-threshold rule: the approval threshold stays constant until some lower threshold is reached.

**Lemma 2.** *For all  $N$ , the solution to  $RSM_N$  is a static threshold approval policy until  $X_t$  reaches the first binding constraint  $X^1$  for the first time. The continuation value for  $R$  at  $\tau(X^1)$  is  $H(X^1)$ .*

Lemma 2 establishes the optimality of the initial stationary regime in  $RSM_N$ : This result doesn’t contradict our earlier result that static thresholds are non-optimal: as we show below, the mechanism in the second regime will not turn out to be a static threshold mechanism. A key thing to note is that the value of  $H$  is independent of the history up until time  $\tau(X^1)$ . What happens after  $\tau(X^1)$  is completely bundled into the value  $H$  and therefore doesn’t affect the choice of  $R$  before  $\tau(X^1)$  except through the value of  $H$ .

We must now solve for the optimal mechanism in the second regime (i.e. that which solves the problem which deliver  $H(X^1)$ ).

**Lemma 3.** *As  $N \rightarrow \infty$ , the stopping rule in the second regime is given by the dynamic approval threshold  $\tau = \tau(B(M_t)) \wedge \tau(b^*(0))$  and  $d_\tau = \mathbb{1}(\tau = \tau_{B(M_t)})$  where  $B(M_t) = \underline{B}(M_t)$ .*

The proof of the Lemma mirrors that of Lemma 2. The difference is that when we are trying to solve for the optimal mechanism which deliver  $R$  payoff  $H(X^1)$  (and  $A$  a payoff of zero), we have added a promise keeping constraint to dynamic participation constraints. The difficulty lies in that we are evaluating promise keeping at 0, which is a boundary condition (since  $A$  can never be given a payoff less than 0). To apply the techniques we used to prove the “local” optimality of static-threshold rules in Lemma 19, we can setup and solve an appropriate approximating problem in which we replace  $PK(0)$  with a promise

keeping constraint to deliver  $A$  a utility of  $\epsilon$ . We define a relaxed problem dropping all non-threshold quitting constraints and again use a Lagrangian approach to find that threshold approval rules are “locally” optimal-i.e., the optimal policy will again be a static threshold policy  $B_2^\epsilon$  until  $Z_t$  reaches  $B_2^\epsilon$  or some  $b_2^\epsilon$ , where  $A$  receives zero expected utility conditional on reaching  $b_2^\epsilon$ . Repeating the same approximation at  $b_2$ , we can derive a series of threshold  $\{b_i^\epsilon\}$  such that the approval threshold is static and a function of the lowest  $b_i$  which has been reached. By taking  $\epsilon \rightarrow 0$ , we get the optimal mechanism which solves for  $H(X^1)$  when using the constraint set  $XT_N$ . We can then take  $N \rightarrow \infty$  and show that the  $b_i$ s get arbitrarily close. In the limit  $M_t^X$  is equal to the highest  $b_i$  which has not been reached.

Having derived the solution to  $RSM_N$ , we need to check that the solution to the relaxed problems solves the full problem  $SM$  as  $N \rightarrow \infty$ .

**Lemma 4.** *Let  $(\tau^N, d_\tau^N)$  be the solution to  $RSM_N$  and  $(\tau, d_\tau) = \lim_{N \rightarrow \infty} (\tau^N, d_\tau^N)$ . Then  $(\tau, d_\tau)$  is a solution to  $SM$ .*

The Lagrangian approach helps us see the qualitative features of the optimal mechanism, that the current approval threshold is a function only of  $M_t^X$ . In order to derive the exact form of  $B(M_t^X)$ , we have to do a bit more work outside of the Lagrangian formulation.

To understand why  $B(M_t) = \underline{B}(M_t^X)$ , we first argue that  $\underline{B}(M_t)$  is an upper bound on  $B(M_t)$ . Imagine that  $B(M_t) > \underline{B}(M_t) + \alpha$  (for some  $\alpha > 0$ ) and assume that  $B(M_t)$  is continuous. Suppose that  $Z_t = M_t$ . We know that when  $Z_t = M_t$ , it must be that  $A$  is indifferent between continuing and quitting. Therefore the expected payoff to  $A$  is the same as if he decided to quit whenever his beliefs reach  $Z_t - \delta$ . Since  $A$ 's utility is strictly decreasing in the approval threshold, the payoff that  $A$  expects from waiting to quit until  $M_t - \delta$  is bounded above by a static threshold mechanism with approval threshold  $B_m = \min\{B(M) : M \in [M_t, M_t - \delta]\}$ . This utility given by the upperbound is  $V(B_m, M_t - \delta, M_t)$ . However, we can note that

$$V(B_m, M_t - \delta, M_t) + k < V(\underline{B}(M_t), M_t - \delta, M_t) \approx 0$$

Since the current approval threshold is drifting monotonically downward, it is clear that there should be some point at which it will stop. Since  $R$  will get a negative payoff for approval at any  $Z < 0$ , it will be optimal to stop the downward drift only when  $B(M_t) = 0$ , which is equivalent to rejection at  $Z = b^*(0)$ . Clearly the approval threshold should never drift below zero (since  $R$  would be guaranteeing herself a negative payoff) and should never stop strictly above zero (since the option value of continuing experimentation by decreasing the approval threshold is still strictly positive).

The optimal mechanism leads to an interesting result that the approval threshold drifts downward with  $M_t^X$ . This means that for lower  $M_t^X$ , the probability of Type I error is increasing.

**Proposition 2.** *There exists  $\underline{T}, \bar{T}$  such that for all  $t_1 < \underline{T}$  and  $t_2 > \bar{T}$ , the probability of type I error conditional on approval at time  $t_1$  is less than the probability of type I error conditional on approval at time  $t_2$ .*

The result stands in contrast to the single-decision maker problem in which the probability of type I error is constant across approval times. It also gives us some observable predictions. It tells us that projects that are approved rapidly should be expected to be good more often than projects that take a long time to be approved. In many contexts this fits a natural intuition. If an assistant professor receives tenure very quickly, he is more likely to be judged to be of high quality than if he took a long time to receive tenure.

#### 4.4 No Commitment

We also consider the case in which  $R$  cannot commit to the approval rule. First, we need specify what quitting means in this environment. We consider several set-ups for what happens when the agent stops experimenting:

- 1.  $A$  can irrevocably quit experimenting at any time  $t$  and  $R$  can approve at any time after the agent quits.
- 2.  $A$  can irrevocably quit experimenting at any time  $t$  and  $R$  cannot approve after  $A$  has quit.
- 3.  $A$  can temporarily stop experimenting at any time. While  $A$  is not experimenting,  $A$  pays no flow cost and  $R$  can approve at any time.

We might naturally wonder how much is gained by allowing the principal to commit to the approval rule. If we restrict attention to Markov Perfect Equilibrium using the belief  $Z_t$  as the state variable (as is standard in the literature), then the answer is quite a lot.

**Proposition 3.** *There exists a pair  $(B, b)$  such that  $R$  only at time  $\tau(B)$  and  $A$  quits at time  $\tau(b)$ . In set-up 1.,  $B > 0$  while in set up 2,  $B = 0$  and  $b = b^*(0)$  and  $A$  quits experimenting when  $Z_t \notin (b, B)$ . The value of experimentation to  $R$  is strictly less than under one- or two-sided commitment.*

Set-up 1 corresponds to the model of Kolb (2016). Note that in set-up 2,  $R$  doesn't benefit from experimentation at all: if she approves, she is either approving at  $Z_t = 0$  (which is her myopic threshold) or is approving immediately at  $Z_0$ . The agent is able to benefit from quitting as soon as he knows that  $R$  will approve in the subsequent subgame.

However, the restriction to Markov Perfect Equilibrium has a lot of bite. As the following proposition shows, if we allow for general Nash equilibrium, we find that the optimal mechanism under one-sided commitment can be implemented without commitment. Commitment, in the case of symmetric information and one-sided commitment, is not needed in order to implement  $R$ 's second-best mechanism.

**Proposition 4.** *Under set-up 3, the optimal mechanism under one-sided commitment can be implemented as an equilibrium.*

Note that the optimal mechanism under one-sided commitment is “renegotiation-proof.” Since the threshold is only decreasing downward and  $A$  strictly prefers a downward movement of the threshold,  $A$  would never agree to another mechanism which raised the threshold.

#### 4.5 Comparison of Commitment Structures

The reduction of commitment from two- to one-sided commitment leads to very different mechanisms. Because  $R$  doesn't have to consider dynamic participation constraints under two-sided commitment, she is able to “smooth” out distortions. The Lagrangian approach makes it clear how the static nature of the constraints makes the solution a threshold based policy.

Borrowing financial options terminology, we can see that under two-sided commitment, the optimal mechanism satisfies the record-setting news principle-i.e., a decision (either approval or rejection) happens only at the infimum or supremum of the realized path of  $X_t$ . However, dynamic participation constraints in the one-sided commitment case lead to a loss of the record-setting news principle. Since the approval threshold is decreasing over time, it will happen with positive probability that approval happens at an  $X_t < \sup_s X_s$ .

### 5 Asymmetric Information

We now allow for  $\pi_A$  to take on a binary realization  $\pi_A \in \{\pi_\ell, \pi_h\}$  where  $\pi_\ell < \pi_h$ . Translating into log-likelihood space, we will call the case when  $A$  begins with prior  $Z_0 = Z_\ell = \log(\frac{\pi_\ell}{1-\pi_\ell})$  the low type of  $A$  (who we refer to as  $\ell$ ) and when  $A$  begins with the prior  $Z_0 = Z_h = \log(\frac{\pi_h}{1-\pi_h})$  as the high type of  $A$  (who we refer to as  $h$ ). We let  $p(Z_i)$  be the ex-ante probability of type  $Z_i$ .

We split time  $t = 0$  into three “instances”:  $\{0_-, 0, 0_+\}$ . At time  $t = 0_-$ ,  $A$  is given a signal which conveys some information about the state. Without loss of generality, we assume that the signal is equal to his posterior about  $\omega$ , i.e., his the signal  $s = \pi_A$ . We will focus



on the case when  $\pi_A \in (0, 1)$ . In many applications this is the more realistic assumption than  $\pi_A \in \{0, 1\}$ .  $R$  knows that  $A$  receives a signal, but the realization of the signal is private information to  $A$ . Then, at time  $t = 0$ ,  $A$  can send a message  $m \in \mathcal{M} = \{h, \ell\}$  to  $R$ , after which  $R$  can publicly commit to a mechanism. Finally, at time  $t = 0_+$  a public news process begins. We can think of this public news process as the outcome of some noisy experiments that  $R$  requires  $A$  to perform (e.g., conducting clinical trials).

We now redefine a stopping mechanism to account for the private information of the agent.

**Definition 8.** A *stopping mechanism* is a menu  $\{(\tau^i, d_\tau^i)\}_{i=h,\ell}$  such that  $\tau^i$  is an  $\mathcal{F}_t^X$ -measurable stopping rule and  $d_\tau^i$  is an  $\mathcal{F}_t^X$  measurable decision rule which takes value 0 or 1 and  $R$  implements  $(\tau^i, d_\tau^i)$  when she receives message  $m = i$ .

**Definition 9.** A stopping mechanism is *incentive compatible* if for all  $i$ ,

$$\mathbb{E}[e^{-r\tau^i} (d_\tau^i + \frac{c}{r}) | Z_i] \geq \mathbb{E}[e^{-r\tau^j} (d_\tau^j + \frac{c}{r}) | Z_i]$$

With asymmetric information, there are two competing forces at play,  $R$  must balance two contrasting goals. She must design a menu which is discerning and not too likely to approve low types. On the other hand, she would like to approve high types more quickly. However, these two forces will typically be in competition: offering higher approval standards for low types than high types makes it more attractive for low types to claim to be high types. How can  $R$  design a mechanism that allows him to approve high types quicker while still disincentivizing low types from claiming to be high types? As we will see, the degree of commitment (one- or two-sided) is crucial for determining the best way to do this. Under two-sided commitment,  $R$  can threaten low types with prolonged experimentation as the evidence becomes negative which will be enough to dissuade deviation. However, with one-sided commitment such a threat is no longer credible: the low type always has his outside option available. Instead,  $R$  will sometimes offer a fast-track mechanism, in which  $h$  is initially given a lower approval threshold, but also a high “failure” threshold. If the evidence reaches the “failure” threshold first, then the approval threshold jumps up, from which it slowly drifts back down as a function of  $M_t^X$ . This jump in the approval threshold acts a punishment to dissuade  $\ell$  from claiming to be  $h$ . That the optimal mechanism here depends only on  $M_t$  here is interesting: a similar mechanism to that which works to incentivize  $A$  in the symmetric case also serves to disincentivize  $A$  from misreporting his type. The mechanism also shows how  $R$  “backloads distortions.” The lower approval threshold makes imitating  $h$  more attractive for  $\ell$ . In order to dissuade  $\ell$ ,  $R$  creates the probability of a one-time increase in the approval threshold. Because  $\ell$  and  $h$  have different

beliefs about the probable path of  $X_t$ ,  $\ell$  views the increase as more likely than  $h$  does. Therefore,  $R$  can introduce the distortions (increasing the approval threshold) in such a way as to decrease  $\ell$ 's payoffs differentially than  $h$ 's payoffs.

When we are considering the effects of  $A$  misreporting his type, the beliefs, the beliefs of  $A$  and  $R$  will be different. Note that because initial beliefs enter linearly into  $Z_t$ , after any realization of  $X_t$ , the beliefs of  $A$  and  $R$  (when  $A$  misreports his type) will be different by  $\Delta_z := Z_h - Z_\ell$ .

## 5.1 Two-Sided Commitment

Formally,  $R$ 's problem under two-sided commitment is given by

$$\begin{aligned} \sup_{(\tau^i, d_\tau^i)_{i=\ell, h}} & \sum_{i=\ell, h} \mathbb{E}[e^{-r\tau^i} d_\tau^i \frac{e^{Z_{\tau^i}} - 1}{1 + e^{Z_{\tau^i}}} | Z_i] \cdot p(Z_i) \\ \text{subject to} & \\ P(Z_i) & \mathbb{E}[e^{-r\tau^i} (d_\tau^i + \frac{c}{r}) | Z_i] \geq \frac{c}{r} \\ IC(Z_i, Z_j) & \mathbb{E}[e^{-r\tau^i} (d_\tau^i + \frac{c}{r}) | Z_i] \geq \mathbb{E}[e^{-r\tau^j} (d_\tau^j + \frac{c}{r}) | Z_i] \end{aligned}$$

**Proposition 1.** *The optimal mechanism under two-sided commitment is a menu of static-threshold stopping rules.*

The assumption of two-sided commitment is the same as in Guo (2016). Even though the problems being considered are somewhat different, we find that threshold rules are optimal (similar to her results).

With two-sided commitment, we actually find that the  $h$  type's incentive constraint is the binding one. Because  $R$  can make  $\ell$  commit to experiment past the threshold at which  $\ell$ 's expected continuation value becomes negative,  $R$  can punish  $\ell$  by increasing experimentation on for low beliefs (which also increases  $R$ 's utility). This force is strong enough so that  $R$  can always find a way to dissuade  $\ell$  from claiming to be  $h$ .

**Proposition 2.** *The high type is always made to experiment longer-i.e.,  $b_h > b_\ell$ . If  $b_h \neq -\infty$ , then  $IC(Z_h, Z_\ell)$  and  $P(Z_\ell)$  are binding.*

## 5.2 One-Sided Commitment

We must reformulate the standard participation and incentive constraints to dynamic nature of incentives under one-sided commitment. To do this we will need to define a

dynamic version of incentive compatibility in the same nature as we defined dynamic participation constraints. Incentive constraints for type  $i$ 's value of reporting to be type  $j$  must take into account that  $i$  also considers the value of a deviation where he may choose to quit early.

With this in mind, we write the mechanism design problem with asymmetric information as

$$\begin{aligned}
(AM) \quad & \sup_{(\tau^i, d_\tau^i)_{i=\ell, h}} \sum_{i=\ell, h} \mathbb{E}[e^{-r\tau^i} d_\tau^i \frac{e^{Z_\tau^i} - 1}{1 + e^{Z_\tau^i}} | Z_i] \cdot p(Z_i) \\
& \text{subject to} \\
& DP(Z_i) \quad \sup_{\tau^{i'}} \mathbb{E}[e^{-r(\tau^i \wedge \tau^{i'})} (d_\tau^i \mathbb{1}(\tau^i \leq \tau^{i'}) + \frac{c}{r}) | Z_i] \leq \mathbb{E}[e^{-r\tau^i} (d_\tau^i + \frac{c}{r}) | Z_i] \\
& DIC(Z_i, Z_j) \quad \sup_{\tau^{i'}} \mathbb{E}[e^{-r(\tau^j \wedge \tau^{i'})} (d_\tau^j \mathbb{1}(\tau^j \leq \tau^{i'}) + \frac{c}{r}) | Z_i] \leq \mathbb{E}[e^{-r\tau^i} (d_\tau^i + \frac{c}{r}) | Z_i]
\end{aligned}$$

We include  $\sup_{\tau^{i'}}$  in  $DIC(Z_i, Z_j)$  to convey the fact that  $i$  is comparing correctly declaring his type to be  $i$  to the maximum payoff he could get from reporting to be type  $j$ .

We begin by looking at what the optimal mechanism is when  $DIC(Z_\ell, Z_h)$  is binding and  $DIC(Z_h, Z_\ell)$  is slack. That this would arise seems intuitive: the incentives of  $R$  and  $A$  are more closely aligned when  $A$ 's beliefs are higher. As we will show later, this intuition is correct if  $Z_h$  is high enough.

Let  $V_\ell$  be the utility that  $\ell$  gets from truthfully declaring his type. Then our problem of determining the optimal high type mechanism is given by

$$\begin{aligned}
(AM^h) \quad & \sup_{(\tau^i, d_\tau^i)_{i=\ell, h}} \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_h] \\
& \text{subject to} \\
& DP(Z_h) \quad \sup_{\tau'} \mathbb{E}[e^{-r(\tau \wedge \tau')} (d_\tau \mathbb{1}(\tau \leq \tau') + \frac{c}{r}) | Z_h] \leq \mathbb{E}[e^{-r\tau} (d_\tau + \frac{c}{r}) | Z_h] \\
& DIC(Z_\ell, Z_h) \quad \sup_{\tau^{i'}} \mathbb{E}[e^{-r(\tau \wedge \tau')} (d_\tau^j \mathbb{1}(\tau \leq \tau') + \frac{c}{r}) | Z_\ell] \leq V_\ell + \frac{c}{r}
\end{aligned}$$

We will perform a similar proof technique as in the symmetric information case. First, we will drop all but a finite number of threshold quitting rules ( $XT_N$ ) from  $DP$  and  $DIC$ . This allows us to define a relaxed version of  $AM^h$  which we call  $RAM_N^h$ .

As in the symmetric mechanism, we can decompose the mechanism into two parts: the optimal mechanism up until the time at which, if  $\ell$  had deviated and declared himself to be  $h$  to  $R$ ,  $\ell$  would decide to quit ( $\ell$ 's time of first indifference) and the optimal mechanism

which induces  $\ell$  to quit at that point. The optimal mechanism which induces  $\ell$  to quit will not necessarily be reject and thus the expected payoff of the optimal mechanism which induces  $\ell$  to quit when the evidence reaches  $X_t$  will have a positive value  $H^h(X_t)$  given by

$$\begin{aligned}
H^h(X_t) = \text{sup}_{(\tau, d_\tau)} & \quad \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_t] \\
\text{subject to} & \quad \forall X_i \in \{X_j \in XT_N : X_j < X_t\} \\
RDP(\tau, d_\tau) & \quad \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} (d_\tau \mathbb{1}(\tau(X_i) > \tau) + \frac{c}{r}) | Z_t] \leq \mathbb{E}[e^{-r\tau} (d_\tau + \frac{c}{r}) | Z_t] \\
PK(0) & \quad \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} (d_\tau \mathbb{1}(\tau(X_i) > \tau) + \frac{c}{r}) | Z_t - \Delta_Z] \leq \frac{c}{r}
\end{aligned}$$

By using a similar approach as in Lemma 19, we arrive at the conclusion that static threshold strategies solve  $AM^h$ .

**Lemma 5.** *The solution to  $RAM_N^h$  is given by a stationary threshold  $B_1^h$  until  $X_t$  reaches either  $B_1^h$  or  $b_1^h$ . If  $X_t = b_1^h$ , then the optimal stopping rule after  $\tau(b_1^h)$  is the solution to  $H^h(X^1)$ .*

Once  $b_1^h$  is reached,  $R$  must induce  $\ell$  to quit. There are many ways in which the optimal mechanism could induce  $\ell$  to quit while still providing incentives for  $h$  to experiment. For example, the mechanism could set a high static threshold above  $\underline{B}(X_t, Z_t - \Delta_Z)$ . Since  $h$  is more optimistic about the state (and thus believes that approval is more likely), this could be done in such a way as to strictly induce  $\ell$  to quit and  $h$  to experiment.

It is important to note that while the mechanism must induce  $\ell$  to quit, the payoff relevant beliefs are those of  $h$ . It may be that by inducing  $\ell$  to quit, the mechanism will be setting a stricter approval policy that  $R$  would like to given that the true beliefs are  $h$ . If this is the case, then we would like to “loosen” the approval policy over time while making sure to do in such a way as to not violate the earlier incentives for  $\ell$  to quit. We verify that this intuition is correct in the following Lemma. First though, we define some notation:

$$\begin{aligned}
b_i^*(B, X) & := \text{argmax}_b V(B, b, X, Z_i) \\
\underline{B}_i(X) & := \min\{B : b_i^*(B, X) = X\}
\end{aligned}$$

$b_h^*(B, X_t)$  ( $b_\ell^*(B, X_t)$ ) is the threshold at which a high (low) type would quit when facing a static approval threshold is  $B$ , the current evidence level is  $X_t$  and  $A$ 's beliefs are  $Z_h$  ( $Z_\ell$ ).  $\underline{B}_i(X)$  is the lowest approval threshold which would induce type  $i$  to quit at evidence level  $X$ .

**Lemma 6.** *The optimal mechanism which solves  $H^h(X_t)$  when the current evidence is  $X_t$  is given by a dynamic threshold policy  $\tau = \tau(B^h(M_t^X)) \wedge \tau(\underline{b}(Z_h))$  and  $d_\tau = \mathbf{1}(\tau = \tau(B^h(M_t^X)))$  where*

$$B^h(M_t^X) = \begin{cases} \underline{B}_\ell(M_t^X) & M_t^X \in [b_\ell^*(B_h^2, M_t^X), X_t] \\ B_h^2 & M_t^X \in [b_h^*(B_h^2, M_t^X), b_\ell^*(B_h^2, M_t^X)) \\ \underline{B}_h(M_t^X) & M_t^X \in [b_h^*(\frac{-Z_0\sigma}{\phi}, M_t^X), b_h^*(B_h^2, M_t^X)] \end{cases}$$

When  $DIC(Z_h, Z_\ell)$  is slack,  $B_h^2$  is the same as in the symmetric case with belief  $Z_h$

This second stage of the optimal mechanism consists of several regimes:

- **Punishment Regime:** After exiting the initial stationary regime, the mechanism enters a “punishment” regime which punishes any  $\ell$  type which claims to be  $h$ . The fact that  $B^h(M_t^X) = \underline{B}_\ell(M_t^X)$  ensures that  $\ell$  does indeed want to quit.  $R$  would like to decrease the approval threshold as quickly as possible and we show that  $\underline{B}_\ell(M_t^X)$  provides the best way to do this while maintaining prior incentives for  $\ell$  to quit.
- **Second Stationary Regime:** When  $M_t^X = b_\ell^*(B^*, M_t^X)$  the approval threshold is at a point at which  $R$  would optimally choose when she knows  $A$  to be the high type. At this point,  $R$  can keep the approval threshold constant, which still provides satisfies the prior incentives for  $\ell$  to quit.
- **Incentivization Regime:** If  $M_t^X = b_h^*(B^*, M_t^X)$ , then  $h$  must be incentivized to continue experimenting. As in the symmetric case, the optimal way to do this is to set the approval threshold to be  $\underline{B}_h(M_t^X)$ .

Because  $R$  rejects only when the evidence reaches reach  $b_h^*(\frac{-Z_0\sigma}{\phi}, M_t^X)$  (i.e., when the approval threshold is at the point such that  $Z_t = 0$ ), this implies that there is no distortion at the end of experimentation.  $DICs$  for  $h$  never cause  $R$  to reject earlier that would be optimal in the absence of incentive constraints. This comes about because the  $\ell$  always has more pessimistic beliefs that  $h$  and thus there is always a way to deliver incentives for  $\ell$  to quit that don’t involve rejection.

Having found what the optimal mechanism is for  $h$ , we must also explore what the optimal mechanism for  $\ell$  is. We will consider when  $DIC(Z_h, Z_\ell)$  is not binding. Suppose that the mechanism must deliver utility  $V_\ell$  to the low type correctly declaring his type. Then the mechanism design problem is given by

$$\begin{aligned}
& \text{sup}_{(\tau, d_\tau)} \quad \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_\ell] \\
& \text{subject to} \\
& DP(Z_\ell) \quad \text{sup}_{\tau' \in \mathbb{T}} \mathbb{E}[e^{-r(\tau \wedge \tau')} (d_\tau \mathbf{1}_{\tau' > \tau} - \frac{c}{r}) | Z_\ell] \leq \mathbb{E}[e^{-r\tau} (d_\tau + \frac{c}{r}) | Z_\ell] \\
& PK(V_\ell) \quad \mathbb{E}[e^{-r\tau} (d_\tau + \frac{c}{r}) | Z_\ell] = V_\ell
\end{aligned}$$

Except for the additional promise keeping constraint  $PK(V_\ell)$ , this is identical to the symmetric information mechanism. We should expect that the optimal mechanism for  $\ell$  is qualitatively the same as the symmetric information mechanism, which turns out to be correct.

**Lemma 7.** *The optimal mechanism for  $\ell$  satisfies  $PK(V_\ell, Z_\ell)$  when  $DIC(Z_h, Z_\ell)$  is slack is given by a dynamic approval threshold  $B_\ell(M_t)$ , which is defined as*

$$B^h(M_t^X) = \begin{cases} B_1^\ell & M_t^X \in [b_\ell^*(B_1^\ell, M_t^X), 0) \\ \underline{B}_\ell(M_t^X) & M_t^X \in [b_\ell^*(-\frac{Z_0\sigma}{\phi}, M_t^X), b_\ell^*(B_1^\ell, M_t^X)) \end{cases}$$

for some  $B_1^\ell \in \mathbb{R}$  and  $B_1^\ell$  is less than it would be in the symmetric information case.

Both  $\ell$  and  $h$  receive an initial stationary regime. Qualitatively, the features of the second stage are mostly determined by the initial static phase. We are interested in how these static phases compare for  $h, \ell$ . Is it that  $h$  is offered a lower approval threshold than  $\ell$ ? While it seems intuitive, lowering the approval threshold also introduces other distortions into the mechanism through the  $DIC$  constraints. Additionally, the earlier  $\ell$  is rejected, the lower the optimal approval threshold is for  $\ell$ .

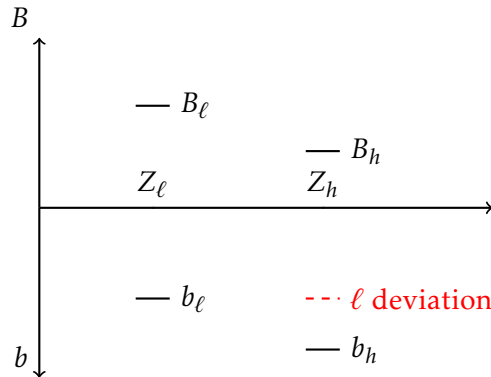
**Theorem 2.** *When  $DIC(Z_\ell, Z_h)$  is binding and  $DIC(Z_h, Z_\ell)$  is slack, the optimal mechanism is given by a stopping rules  $\tau_i = \tau(B^i(M_t^X)) \wedge \tau(\underline{b}(Z_i))$  and  $d_\tau^i = \mathbf{1}(\tau = \tau(B^i(M_t^X)))$  where  $B^i(M_t^X)$  are as in Theorem 1. Let  $(B_1^i, b_1^i)$  be the thresholds of the stationary regimes. Then  $B_1^h \leq B_1^\ell$  and  $b_1^h > b_1^\ell$  if  $B_1^h < B_1^\ell$ .*

- **Low Type Monotonicity:** The mechanism for  $\ell$  closely resembles that of the symmetric mechanism in that the approval threshold will only drift downwards.
- **High Type Jump:** If  $b_1^h > b_1^\ell$ , then the approval threshold for  $h$  takes a jump upwards when  $X_t$  reaches  $b_1^h$  for the first time, after which it is monotonically decreasing in  $M_t^X$ . This jump upward occurs in order to incentivize a deviating  $\ell$  type to quit.

- **No Distortion At The End:**  $R$  never rejects (for either type) at an evidence level higher than she would in the symmetric mechanism.
- **Some Distortion At The Start:** The stationary approval threshold is at least as high as it would be in the symmetric mechanism. If  $R$  were to decrease  $B_1^h$  to  $B_h^*$ , then the punishment regime would have to be more likely to satisfy  $DIC(Z_\ell, Z_h)$ .

We refer to the mechanism given to  $h$  when  $b_1^h > b_1^\ell$  as a **fast-track mechanism**. We can think of  $h$  as being offered a two stage trial: the first trial (a fast-track) is given a low approval threshold, but also a “failure” threshold. If the failure threshold is reached first, then the trial is declared a failure. However, instead of rejecting,  $R$  allows  $h$  to immediately begin a new trial, only now  $h$  is given a higher approval threshold.

This fast-track mechanism illustrates the tradeoffs that must be made under one-sided commitment: in order to grant  $h$  a lower approval threshold,  $R$  must deter deviations by  $\ell$  by increasing the failure threshold. This lower approval threshold is more likely to be reached by  $h$  than  $\ell$ , which allows  $R$  to profitably backload distortions in the “failure” threshold.



We now illustrate some of the ideas behind why the optimal mechanism must take this nested form. For  $\ell$  when declaring himself to be  $\ell$  or  $h$ , his utility is completely determined by the initial static thresholds. We know that  $b_1^\ell = b_\ell^*(B_1^\ell)$ -i.e.,  $R$  keeps the stationary regime until the point at which  $\ell$  reaches his point of first indifference. We begin by noting that if  $B_1^h > B_1^\ell$ , then we cannot have  $DIC(Z_\ell, Z_h)$  binding. The reason for this is clear: since the static approval threshold is higher (which strictly reduces utility to  $\ell$ ),  $\ell$  must gain from experimenting longer on the low end of beliefs when claiming to be  $h$ . But since  $\ell$  when truthfully declaring his type is allowed to experiment up until the point at which he would choose to quit, there is nothing to be gained (relative to truthfully declaring his type) for  $\ell$

from claiming to be  $h$ . If  $B_1^h < B_1^\ell$ , then it must be that  $b_h < b_\ell$ . Otherwise  $\ell$  could profitably deviate by claiming to be  $h$  and quitting when beliefs drift down from initial beliefs by  $b_\ell$ . In this way,  $\ell$  is able to maintain the same lower width as truthfully declaring himself to be  $\ell$  while also achieving a lower approval threshold  $B_h$ .

Theorem 2 assumes that  $DIC(Z_h, Z_\ell)$  is slack and  $DIC(Z_\ell, Z_h)$  is binding. This will not always be the case: there are examples in which  $DIC(Z_h, Z_\ell)$  must bind. This comes about due to the incentivization regime for  $\ell$ . This incentivization regime decreases the approval threshold enough to keep  $\ell$  indifferent. Since  $h$  has a higher belief than  $\ell$  after observing  $X_t$ ,  $h$  (when reporting to be  $\ell$ ) will still have positive continuation value when in  $\ell$ 's incentivization regime, creating incentives for  $h$  to imitate  $\ell$ . However, we can show that if  $Z_h$  is high enough, then the incentives of  $R$  and  $h$  are sufficiently aligned and  $DIC(Z_\ell, Z_h)$  binding is sufficient for  $DIC(Z_h, Z_\ell)$  to be slack.

**Proposition 5.** *For each  $Z_\ell$ ,  $\exists \bar{Z}$  such that  $\forall Z_h > \bar{Z}$ ,  $DIC(Z_h, Z_\ell)$  is slack and  $DIC(Z_\ell, Z_h)$  is binding in the optimal mechanism.*

Although numerical examples show that  $DIC(Z_h, Z_\ell)$  will be slack for  $Z_h$  which are not limiting cases, it will still be the case that for some  $Z_h$ , we will have  $DIC(Z_h, Z_\ell)$  binding in the optimal solution. This comes about because of the incentivization regime: by lowering the approval threshold for  $\ell$ ,  $R$  also makes  $\ell$ 's mechanism more attractive to  $h$ . As we will show, under the assumption that  $Z_\ell < 0$ <sup>6</sup>, we can verify that the optimal mechanism will look very similar to that of Theorem 2. When both  $DIC$ s bind, then the optimal mechanism will introduce distortion into  $\ell$ 's mechanism by inducing early rejection.

**Lemma 8.** *The optimal mechanism for  $\ell$  when  $Z_\ell < 0$  and  $DIC(Z_h, Z_\ell)$  is binding is given by a dynamic approval threshold  $B_\ell(M_t)$ , which is defined as*

$$B^h(M_t^X) = \begin{cases} B_1^\ell & M_t^X \in [b_r \vee b^*(B_1^\ell, Z_\ell), 0) \\ \underline{B}(M_t^X) & M_t^X \in [b_r, b_r \vee b^*(B_1^\ell; Z_\ell)) \end{cases}$$

for some  $(B_1^\ell, b_r) \in \mathbb{R}_2$ .

When we consider the  $h$  mechanism when both constraints are binding, we are simply adding a  $PK$  constraint to deliver some value  $V_h$  to  $h$  when he correctly declares his type. This changes very little about the arguments of Lemma 7. In all, we summarize the optimal mechanism in this case below:

---

<sup>6</sup>We assume that  $Z_\ell < 0$  in order to verify that the optimal mechanism doesn't take an unreasonable form. The assumption that  $Z_\ell < 0$  is reasonable given our application. Many drugs that make it to clinical trials are abandoned or not approved (see ??? (????))



**Theorem 3.** *The optimal mechanism for  $\ell$  when  $Z_\ell < 0$  and both DIC constraints are binding is given by Lemma 8 and the optimal mechanism for  $h$  is given by Lemma 6.*

There are two main differences between the mechanism when  $DIC(Z_h, Z_\ell)$  is binding and when it is slack. When it binds,  $R$  may reject  $\ell$  early (i.e.-  $b_r > \underline{b}(0)$ ) in order to lower incentives for  $h$  to imitate  $\ell$ . Additionally, it may be that the second stationary regime for  $h$  starts below  $R$ 's symmetric information solution (so that  $R$  can provide additional incentives for  $h$  while maintaining  $\ell$ 's incentive to quit).

Qualitatively, this is very similar to the mechanism when  $DIC(Z_h, Z_\ell)$  is slack, except now we have distortion at the bottom by introducing early rejection when the evidence reaches  $b_r$ . Note that we may have a completely static threshold mechanism if  $b_r \geq b^*(B_1^\ell, Z_\ell)$ . In this case, the solution with  $DIC(Z_h, Z_\ell)$  binding and no  $DP(Z_\ell)$  constraints results in a solution in which the rejection threshold is high enough that  $\ell$  never finds it optimal to quit early.

### 5.3 Quantitative Derivation

Our qualitative analysis of the optimal mechanism leaves us with very few parameters over which we must optimize. For high enough  $Z_h$ , the optimal mechanism is completely pinned down by the choice of the thresholds of the stationary regime  $(B_1^i, b_1^i)_{i=h,\ell}$ . When  $DIC(Z_h, Z_\ell)$  is binding, we have consider three parameters each for  $h, \ell$ :  $(B_1^h, b_1^h, B_2^h)$  and  $(B_1^\ell, b_1^\ell, b_r)$ . Given the richness of the available stopping rules, this reduction is somewhat remarkable and makes the problem computationally tractable. The choice of these thresholds will pin down the rest of the mechanism. To find the optimal stationary regime thresholds, we must find what the continuation value to  $R$  is of reaching  $b_i$ .

Define the function  $j(Z, M, b_r)$  (we will drop  $b_r$  for notational convenience) to be the expected value of the principal when the current minimum of evidence is  $M$ , current beliefs are  $Z$  and the project is rejected when beliefs reach  $b_r$ . Using our previous formulas for discounted threshold crossing probabilities, it is easy to see that

$$j(Z, M) = \underbrace{\Psi(B(M), M, Z) \frac{e^Z - e^{Z-B(M)}}{1 + e^Z}}_{\text{Expected utility from approval before } Z_t = M} + \underbrace{\psi(B(M), M, Z) \frac{e^Z + e^{Z-M}}{1 + e^Z}}_{\text{Discounted prob. } M \text{ hit first}} \cdot \underbrace{j(M, M)}_{\text{Continuation value at } M}$$

Thus if we can calculate  $j(M) := j(M, M)$ , the value of  $j(Z, M)$  follows immediately. In order to calculate  $j(M)$ , we use the principle of normal reflection<sup>7</sup>:  $\frac{\partial j(Z, M)}{\partial M} \Big|_{Z=M} = 0$

<sup>7</sup>See Peskir and Sharyayev (2006) for a derivation.

We can then take the derivative with respect to  $M$  to get

$$\begin{aligned} \frac{\partial j(Z, M)}{\partial M} = & B'(M) \left[ \Psi_B \frac{e^Z - e^{Z-B(M)}}{1 + e^Z} + \Psi \frac{e^{Z-B(M)}}{1 + e^Z} + \psi_B \frac{e^Z + e^{Z+M}}{1 + e^Z} j(M) \right] \\ & + \Psi_b \frac{e^Z - e^{Z-B(M)}}{1 + e^Z} + \psi_b \frac{e^Z + e^{Z-b}}{1 + e^Z} j(M) \\ & - \psi \frac{e^{Z-M}}{1 + e^Z} j(M) + j'(M) \psi \frac{e^Z + e^{Z-M}}{1 + e^Z} \end{aligned}$$

Evaluating the above equation at  $Z = M$  and using that  $\frac{\partial j(Z, M)}{\partial M} \Big|_{Z=M+} = 0$ , we get

$$j'(M) = j(M) \left[ \frac{1}{1 + e^M} - \psi_b \right] - \frac{e^M - e^{M-B(M)}}{1 + e^M} \Psi_b$$

where we note that  $\Psi(B(M), M, M) = 0$  and  $\psi(B(M), M, M) = 1$ . This, coupled with the boundary condition  $j(b_r) = 0$  gives the ODE which describes  $j(M)$ .

**Lemma 9.** *The value of experimentation when the current evidence level is  $M_t^X$  and the minimum is  $M_t^x$  is given the unique solution to  $j'(M_t^X)$ .*

For  $\ell$ , we know that the approval mechanism is strictly decreasing in  $M$  and so this equation gives the value of the incentivization for  $\ell$ .

For  $h$ , we know that there may be a second stable regime. In this case, we must solve two differential equations: one for the value in the punishment regime and one for the value of incentivization regime. The ODE for the incentivization regime for  $h$  is identical to the ODE for  $\ell$  with  $b_r = \underline{b}(Z_h)$ . For the ODE describing the punishment regime, the differential equation is identical but the boundary condition is given by

$$j_1(M) = \Psi(B_h^2, b_h^2, M) \frac{e^M - e^{M-B_h^2}}{1 + e^M} + \psi(B_h^2, b_h^2, M) \frac{e^M - e^{M-b_h^2}}{1 + e^M} j_2(M)$$

Since the incentivization regime for  $\ell$  gives a positive value for  $h$ , we must find an equation for the value of  $\ell$ 's incentivization regime for  $h$  (which we will call  $v_\ell$ ). By a similar argument as for  $j(Z, M)$ , we get a differential equation for  $v_\ell(M)$  as

$$v'(M) = [1 - \psi_B B'(M^X, M^Z) - \psi_b] \left( \frac{c}{r} + v(M^Z) \right) - [\Psi_B B'(M) + \Psi_b] \frac{e^{M^Z + \Delta_z} + e^{M-B(M)}}{1 + e^{M^Z + \Delta_z}} \left( 1 + \frac{c}{r} \right)$$

with boundary condition  $v(b_r) = 0$ .

Similarly, we must find the value to  $h$  from the beginning of punishment regime in the mechanism for  $h$ . A similar argument establishes the differential equation to be as the one above. We evaluate the boundary condition as the beginning of the second stationary regime. Since the expected continuation payoff to  $h$  upon reaching the beginning of incentivization regime is zero, we can evaluate the utility to  $h$  of the secondary stationary regime using only the static thresholds. This gives a boundary condition of  $v(M) = \Psi(B_h^2, b_h^2, M)(1 + \frac{c}{r})\frac{e^M + e^{M-B_h^2}}{1+e^M} + \psi(B_h^2, b_h^2, M)\frac{c}{r}\frac{e^M + e^{M-b_h^2}}{1+e^M}$ .

This allow us to write out the mechanism design problem as

$$\begin{aligned} \max_{(B_i, b_i, \underline{b}_i)_{i=h, \ell}} & \sum_{i=h, \ell} \Psi(B_i, b_i) \frac{e^{Z_i} - e^{-B_i}}{1 + e^{Z_i}} + \psi(B_i, b_i) \frac{e^{Z_i} + e^{-b_i}}{1 + e^{Z_i}} j_i(b_i, \underline{b}_i) \\ & \text{subject to} \\ DIC(Z_i, Z_j) : & \Psi(B_i, b_i) \frac{e^{Z_i} - e^{-B_i}}{1 + e^{Z_i}} + \psi(B_i, b_i) \frac{e^{Z_i} + e^{-b_i}}{1 + e^{Z_i}} v_i(b_i, \underline{b}_i) \\ & \geq \Psi(B_j, b_j) \frac{e^{Z_i} - e^{-B_j}}{1 + e^{Z_i}} + \psi(B_j, b_j) \frac{e^{Z_i} + e^{-b_j}}{1 + e^{Z_i}} v_i(b_j, \underline{b}_j) \\ DP(Z_i, Z_j) : & b_i \geq b^*(B_i) \end{aligned}$$

#### 5.4 Comparative Statics

In the symmetric information case, increasing the cost of  $A$  unambiguously hurts  $R$ , since it induces  $R$  to provide more incentivization and reject at a higher beliefs ( $\frac{\partial b^*(0)}{\partial c} > 0$ ). However, with asymmetric information this is no longer the case. Additional costs may be of use as a screening device.

**Proposition 6.** *Under both one- and two-sided commitment, as  $c \rightarrow 0$  the optimal mechanisms for  $h, \ell$  converge to value of the single decision maker problem for  $R$  with prior  $p(Z_h)\pi_h + (1 - p(Z_H))\pi_\ell$ .*

The cost of experimentation to  $A$  is necessary for the creation of separating stopping rules. When  $c$  becomes small, it becomes increasingly harder for  $R$  to induce  $\ell$  to quit while still inducing  $h$  to keep experimenting. With the absence of monetary transfers, costly experimentation provides a tool for screening of types. When information is asymmetric, increasing the cost for  $A$  is always harmful for  $R$  (since it induces  $A$  to quit sooner). However, with asymmetric information, the effect of cost increases is no longer monotonic. In fact, some costs can be useful in that it allows  $R$  to screen types better.

**Proposition 7.** *The value of the optimal mechanism is non-monotonic in  $c$  when  $A$  has private information. When  $A$  has no information, the value of the optimal mechanism is strictly decreasing in  $c$ .*

Consider the limiting cases of  $c$  and suppose that  $Z_h \gg 0 \gg Z_\ell$  and  $p_\ell$  is large. As  $c \rightarrow 0$ , the value of the optimal mechanism converges to that of the first-best symmetric information problem with prior  $p_h\pi_h + (1 - p_h)\pi_\ell$ : all information that  $A$  possess is wasted. If we look at low values of  $\frac{\mu}{\sigma}$ , then the value of the optimal mechanism goes to zero as immediate approval is not optimal and it takes a long time for beliefs to change up to a level at which  $R$  would approve. On the other side, as  $c$  becomes large,  $R$  can always find a mechanism which separates  $h$  and  $\ell$ .  $R$  could find a mechanism which rejects  $\ell$  immediately but for which  $h$  still participates and is approved with positive probability. This will bound the value of the optimal mechanism above zero.

Additionally, we might wonder whether or not it is beneficial to  $R$  for  $A$  to have private information about  $\omega$ . On one hand, if  $R$  can make use of  $A$ 's information, then it is beneficial to  $R$ . On the other hand, private information introduces information rents and can add distortions into  $R$ 's optimal mechanism. Which effect is greater is not ex-ante obvious. To answer this question, we compare the case of symmetric information to the case in which  $A$  has perfect information about  $\omega$ . The following proposition shows that asymmetric information is in fact better for  $R$ .

**Proposition 8.** *Suppose that the prior on  $\omega$  is  $\pi_0$ . Then the value to  $R$  of optimal mechanism under asymmetric information in which  $A$  learns  $\omega$  is higher than the value to  $R$  of the optimal mechanism under symmetric informaton.*

## 6 Conclusion

In this paper, we present a model of a hypothesis testing problem with Brownian learning and agency concerns. We examine how different commitment structures lead to different approval policies. The mechanism we find under one-sided commitment features a history dependent approval threshold, yet can still be solved for in a tractable way and can be written as a function of the minimum of the Brownian motion. We find that the optimal mechanism when the agent posses no private information takes the form of a monotonically decreasing approval threshold. This solution to an optimal stopping problem is novel in the literature and illustrates the use of Lagrangian techniques in stopping problems with agency concerns. We are able to fully characterize the solution in the problem with no adverse selection and are able to pin down the solution to the adverse selection problem up to the choice of a small number of constants.

We also apply the model to the case when the agent has private information. The optimal solution takes the form of a fast-track mechanism: high types are offered a low starting approval threshold, but if the evidence gets too bad, the approval threshold jumps up, entering a punishment phase in which it drifts back down slowly.

Our findings has implications for the design of clinical drug trials. Companies can be offered a menu of trials to choose from: a standard option and a fast-track option. The fast-track option can be implemented using a trial with a low approval threshold and high “failure” threshold. Conditional on reaching the “failure” threshold, the company must conduct a new trial in which the regulator requires more evidence prior to approval. By using a menu of options, the regulator is able to elicit and use the private information of the firms to approve good drugs more quickly. Given the long delays from drug development to market place entry, strategies to speed good drug approval can have large welfare implications.

## References

- [1] Balzer, Thomas, and Klaus Janßen. “A duality approach to problems of combined stopping and deciding under constraints.” *Mathematical methods of operations research* 55, no. 3 (2002): 431-446.
- [2] Bergemann, Dirk, and Ulrich Hege. “Venture capital financing, moral hazard, and learning.” *Journal of Banking & Finance* 22, no. 6 (1998): 703-735.
- [3] Bergemann, Dirk, and Ulrich Hege. “The financing of innovation: Learning and stopping.” *RAND Journal of Economics* (2005): 719-752.
- [4] Bolton, Patrick, and Christopher Harris. “Strategic experimentation.” *Econometrica* 67, no. 2 (1999): 349-374.
- [5] Carpenter, Daniel, and Michael M. Ting. “Regulatory errors with endogenous agendas.” *American Journal of Political Science* 51, no. 4 (2007): 835-852.
- [6] Carpenter, Daniel P. “Protection without capture: Product approval by a politically responsive, learning regulator.” *American Political Science Review* 98, no. 04 (2004): 613-631.
- [7] Chan, Hing Chi Jimmy, Alessandro Lizzeri, Wing Suen, and Leeat Yariv. “Deliberating Collective Decisions.” (2015).

- [8] DiMasi, Joseph. "Innovation in the Pharmaceutical Industry: New Estimates of R&D Costs." (2014)
- [9] Dokuchaev, N.G. "Optimal stopping of stochastic processes in a problem with constraints." *Th. Probab. Appl.* 41 (1997), N 4. 761-768.
- [10] Fong, Kyna. "Evaluating skilled experts: Optimal scoring rules for surgeons." Working Paper, 2009.
- [11] Garfagnini, Umberto. Delegated experimentation. Working paper, 2011.
- [12] Gomes, Renato, Daniel Gottlieb, and Lucas Maestri. "Experimentation and project selection: Screening and learning." *Games and Economic Behavior* 96 (2016): 145-169.
- [13] Grenadier, Steven R., Andrey Malenko, and Nad Malenko. "Timing Decisions in Organizations: Communication and Authority in a Dynamic Environment." *American Economic Review* (Forthcoming).
- [14] Guo, Yingni. "Dynamic delegation of experimentation." *American Economic Review* (2016).
- [15] Halac, Marina, Navin Kartik, and Qingmin Liu. "Optimal contracts for experimentation." *The Review of Economic Studies* (2016): rdw013.
- [16] Halac, Marina, Navin Kartik, and Qingmin Liu. "Contests for experimentation." *Journal of Political Economy*, Forthcoming (2014).
- [17] Henry, Emeric, and Marco Ottaviani. "Research and the approval process." Working Paper, 2014.
- [18] Henry, Emeric, and Marco Ottaviani. "The Organization of Persuasion: Wald Deconstructed." Working Paper, (2015).
- [19] Hörner, Johannes, and Larry Samuelson. "Incentives for experimenting agents." *The RAND Journal of Economics* 44, no. 4 (2013): 632-663.
- [20] Keller, Godfrey, Sven Rady, and Martin Cripps. "Strategic experimentation with exponential bandits." *Econometrica* 73, no. 1 (2005): 39-68.
- [21] Keller, Godfrey, and Sven Rady. "Strategic experimentation with Poisson bandits." *Theoretical Economics* 5, no. 2 (2010): 275-311.

- [22] Kolb, Aaron. "Strategic Real Options." Working Paper (2016).
- [23] Kruse, Thomas, and Philipp Strack. "Optimal stopping with private information." *Journal of Economic Theory* 159 (2015): 702-727.
- [24] Madsen, Erik. "Optimal project termination with an informed agent." Working Paper (2016)
- [25] Moscarini, Giuseppe, and Lones Smith. "The optimal level of experimentation." *Econometrica* 69, no. 6 (2001): 1629-1644.
- [26] Orlov, Dmitry, Andrzej Skrzypacz, and Pavel Zryumov. "Persuading the Regulator To Wait." Working Paper.
- [27] Peskir, Goran, and Albert Shiryaev. *Optimal stopping and free-boundary problems*. Birkhäuser Basel, 2006.
- [28] Sannikov, Yuliy. "A continuous-time version of the principal-agent problem." *The Review of Economic Studies* 75, no. 3 (2008): 957-984.
- [29] Strulovici, Bruno. "Learning while voting: Determinants of collective experimentation." *Econometrica* 78, no. 3 (2010): 933-971.

## 7 Appendix

### 7.1 Properties of $V, J$

**Lemma 10.**  $b^*(B, Z)$  is increasing in  $Z$  and decreasing in  $B$ .

*Proof.*  $b^*(B, Z)$  solves the following first order condition

$$\begin{aligned} 0 &= \Psi_b \frac{e^Z + e^{-B}}{1 + e^Z} \left(1 + \frac{c}{r}\right) + \left[\psi_b \frac{e^Z + e^b}{1 + e^Z} + \psi \frac{e^b}{1 + e^Z}\right] \frac{c}{r} \\ \Rightarrow 0 &= \frac{\partial \tilde{V}}{\partial b} = (R_2 - R_1) \left(1 + \frac{c}{r}\right) (e^Z + e^{-B}) e^{B+b} \\ &\quad + [(R_1 e^{R_1(B+b)} - R_2 e^{R_2(B+b)}) (e^Z + e^b) + (e^{R_1(B+b)} - e^{R_2(B+b)}) e^b] \frac{c}{r} \end{aligned}$$

First we show that  $\frac{\partial b^*(B, Z)}{\partial B} < 0$ . Taking the derivative of the first order condition with respect to  $B$ , we have

$$\begin{aligned} \frac{\partial^2 \tilde{V}}{\partial b \partial B} &= (R_2 - R_1) \left(1 + \frac{c}{r}\right) (e^Z + e^{-B}) e^{B+b} - (R_2 - R_1) \left(1 + \frac{c}{r}\right) e^{-B} e^{B+b} \\ &\quad + [(R_1^2 e^{R_1(B+b)} - R_2^2 e^{R_2(B+b)}) (e^Z + e^b) + (R_1 e^{R_1(B+b)} - R_2 e^{R_2(B+b)}) e^b] \frac{c}{r} \\ &= \frac{c}{r} (R_1^2 e^{R_1(B+b)} - R_2^2 e^{R_2(B+b)}) (e^Z + e^b) - \frac{c}{r} e^b (R_1 e^{R_2(B+b)} - R_2 e^{R_2(B+b)}) \\ &\quad - (R_2 - R_1) e^b \left(1 + \frac{c}{r}\right) + \left(\frac{c}{r} (R_2 e^{R_2(B+b)} + R_1 e^{R_1(B+b)}) (e^Z + e^b) + \frac{c}{r} e^b (e^{R_2(B+b)} + e^{R_1(B+b)})\right) \end{aligned}$$

where the last line follows from substituting in the first-order condition.

If we group terms with  $e^b$ , we have (using the fact that  $R_1 + R_2 = 1$  and  $R_1 < 0 < R_2$ )

$$\begin{aligned} &= (R_2 - R_1) \left(1 + \frac{c}{r}\right) + \frac{c}{r} (e^{R_2(B+b)} (R_2 - 1)^2 - e^{R_1(B+b)} (R_1 - 1)^2) \\ &\geq \frac{c}{r} [e^{R_2(B+b)} (R_2 - 1)^2 - (e^{-R_1(B+b)} (R_1^2 - 2R_1 - 1) + 2R_1 - 1)] \\ &\geq C [e^{R_2(B+b)} (R_2 - 1)^2 - (e^{R_1(B+b)} R_1^2 + 2R_1 (1 - e^{R_1(B+b)}) - 1 - e^{R_1(B+b)})] \\ &\geq \frac{c}{r} [e^{R_2(B+b)} (R_2 - 1)^2 - e^{R_1(B+b)} R_1^2] \\ &= \frac{c}{r} (e^{R_2(B+b)} - e^{R_1(B+b)}) (R_2 - 1)^2 > 0 \end{aligned}$$

If we group terms with  $e^Z$ , a similar manipulation gives us



$$(e^{R_2\Delta} - e^{R_1\Delta})(R_2 - 1)^2 > 0$$

Therefore we must have  $\frac{\partial^2 \tilde{V}}{\partial B \partial b} < 0$ .

By the second-order condition, we know that  $\frac{\partial^2 \tilde{V}}{\partial b^2} < 0$  and so we must have  $\frac{\partial b^*(B, Z)}{\partial B} < 0$ .

To show that  $\frac{\partial b^*(B, Z)}{\partial Z} > 0$ , we note that  $\frac{\partial V}{\partial b}$  is proportional to

$$(R_2 - R_1)\left(1 + \frac{c}{r}\right)(e^Z + e^{-B})e^{B+b} + [(R_1 e^{R_1(B+b)} - R_2 e^{R_2(B+b)})(e^Z + e^b) + (e^{R_1(B+b)} - e^{R_2(B+b)})e^b] \frac{c}{r}$$

which is negative as  $b \rightarrow \infty$ . Since  $V$  is single-peaked in  $b$ , it must be that  $b^*(B, Z)$  is finite.  $\square$

**Lemma 11.** For each  $Z$ ,  $\exists B_Z$  such that  $b^*(B, Z) = 0 \forall B \geq B_Z$  and  $b^*(B, Z)$  is finite for all  $B > 0$ .

*Proof.* Suppose not. Then for all  $B$  we must have  $b > 0$  (since  $b^*(B, Z)$  is decreasing in  $B$ ) and hence the first order condition must hold.

$$(R_1 - R_2)\left(1 + \frac{c}{r}\right)(e^Z + e^{-B})e^{B+b} = [(R_1 e^{R_1(B+b)} - R_2 e^{R_2(B+b)})(e^Z + e^b) + (e^{R_1(B+b)} - e^{R_2(B+b)})e^b] \frac{c}{r}$$

As  $B \rightarrow \infty$ , we have

$$\begin{aligned} (R_1 - R_2)\left(1 + \frac{c}{r}\right)e^{Z+B+b} &\approx [-R_2 e^{R_2(B+b)}(e^Z + e^b) - e^{R_2(B+b)}e^b] \frac{c}{r} \\ \Rightarrow R_1\left(\left(1 + \frac{c}{r}\right)e^{Z+B+b} - e^{R_2(B+b)} \frac{c}{r}\right) &\approx R_2\left(\left(1 + \frac{c}{r}\right)e^{Z+B+b} - e^{R_2(B+b)} \frac{c}{r}\right) \end{aligned}$$

which is clearly a contradiction since  $R_1 < 0 < R_2$ .

To show that  $b^*(B, Z)$  is finite, we note that as  $b \rightarrow \infty$ , the first-order condition becomes

$$\begin{aligned} (R_1 - R_2)\left(1 + \frac{c}{r}\right)(e^{Z+B+b} + e^b) &\approx [-R_2 e^{R_2(B+b)}(e^Z + e^b) - e^{R_2(B+b)}e^b] \frac{c}{r} \\ \Rightarrow R_1\left(\left(1 + \frac{c}{r}\right)e^{Z+B+b} - e^{R_2(B+b)} \frac{c}{r}\right) &\approx R_2\left(\left(1 + \frac{c}{r}\right)e^{Z+B+b} - e^{R_2(B+b)} \frac{c}{r}\right) \end{aligned}$$

$\square$

**Lemma 12.**  $J$  is single peaked in  $B$  for a fixed  $b$ .

*Proof.* Fix a rejection threshold  $b$  and let the derivative of  $R$ 's utility with respect to  $B$  be zero. We can then show that the second derivative is negative, since the second derivative is

$$\begin{aligned}
\frac{\partial^2 J}{\partial B^2} &= \frac{(R_2^2 e^{-R_2 \Delta} - R_1^2 e^{-R_1 \Delta})(e^{-R_1 \Delta} - e^{-R_2 \Delta}) + (R_2^2 e^{-R_2 \Delta} - R_1^2 e^{-R_1 \Delta})^2}{(e^{-R_1 \Delta} - e^{-R_2 \Delta})^2} (1 - e^{-B}) \\
&\quad + e^{-B} \left[ \frac{R_1 e^{-R_1 \Delta} - R_2 e^{-R_2 \Delta}}{e^{-R_1 \Delta} - e^{-R_2 \Delta}} - 1 - \frac{(R_2 - R_1)(R_2 e^{R_2 \Delta} - R_1 e^{R_1 \Delta})}{(e^{R_2 \Delta} - e^{R_1 \Delta})^2} H(Z_r)(1 + e^{-b}) \right] \\
&\leq \frac{(R_2^2 e^{-R_2 \Delta} - R_1^2 e^{-R_1 \Delta})}{e^{-R_1 \Delta} - e^{-R_2 \Delta}} (1 - e^{-B}) - \frac{R_1 e^{-R_1 \Delta} - R_2 e^{-R_2 \Delta}}{e^{-R_1 \Delta} - e^{-R_2 \Delta}} e^{-B} + e^{-B} \left[ \frac{R_1 e^{-R_1 \Delta} - R_2 e^{-R_2 \Delta}}{e^{-R_1 \Delta} - e^{-R_2 \Delta}} - 1 \right] \\
&= \frac{(R_2^2 e^{-R_2 \Delta} - R_1^2 e^{-R_1 \Delta})}{e^{-R_1 \Delta} - e^{-R_2 \Delta}} (1 - e^{-B}) - e^{-B} \\
&= \frac{(R_2^2 e^{-R_2 \Delta} - R_1^2 e^{-R_1 \Delta})}{e^{-R_1 \Delta} - e^{-R_2 \Delta}} (e^Z - e^{-B}) + \frac{R_1 e^{-R_1 \Delta} - R_2 e^{-R_2 \Delta}}{e^{-R_1 \Delta} - e^{-R_2 \Delta}} (1 - e^{-B}) \\
&= \frac{e^{-R_2 \Delta} (R_2 - 1) R_2 - e^{-R_1 \Delta} (R_1 - 1) R_1}{e^{-R_1 \Delta} - e^{-R_2 \Delta}} (1 - e^{-B}) \\
&\leq 0
\end{aligned}$$

where  $\Delta = B - b$  and we use the assumption that  $\frac{\partial J}{\partial B} = 0$ .

□

**Lemma 13.**  $V$  is single peaked in  $b$ .

*Proof.* We will show that if  $\frac{\partial V}{\partial b} = 0$ , then  $\frac{\partial^2 V}{\partial b^2} < 0$ .

We know that

$$\frac{\partial V}{\partial b} = [(R_1 - R_2) e^{B+b} (e^Z + e^{-B}) + C(R_2 e^{R_2(B+b)} - R_1 e^{R_1(B+b)}) (e^Z + e^b) - C e^b (e^{R_2(B+b)} - e^{R_1(B+b)})] \frac{\psi}{e^{R_2(B+b)} - e^{R_1(B+b)}}$$

so

$$\begin{aligned}
\frac{\partial^2 V}{\partial b^2} &= [-(R_1 - R_2)e^{B+b}(e^Z + e^{-B}) - C(R_2e^{R_2(B+b)} - R_1e^{R_2(B+b)})e^b - C(R_2^2e^{R_2(B+b)} - R_1^2e^{R_1\Delta})(e^Z + e^{Z'-b}) \\
&\quad + Ce^{Z'-b}(e^{R_2(B+b)} - e^{R_1(B+b)}) + Ce^b(R_2e^{R_2(B+b)} - R_1e^{R_1(B+b)})] \frac{\psi}{e^{R_2(B+b)} - e^{R_1(B+b)}} \\
&= [-(R_1 - R_2)e^{B+b}(e^Z + e^{-B}) - C(R_2^2e^{R_2(B+b)} - R_1^2e^{R_1(B+b)})(e^Z + e^b) + Ce^b(e^{R_2(B+b)} - e^{R_1(B+b)})] \frac{\psi}{e^{R_2(B+b)} - e^{R_1(B+b)}} \\
&= [C(R_2e^{R_2(B+b)} - R_1e^{R_1(B+b)})(e^Z + e^b) - Ce^b(e^{R_2(B+b)} - e^{R_1(B+b)}) \\
&\quad - C(R_2^2e^{R_2(B+b)} - R_1^2e^{R_1(B+b)})(e^Z + e^b) + Ce^b(e^{R_2(B+b)} - e^{R_1(B+b)})] \frac{\psi}{e^{R_2(B+b)} - e^{R_1\Delta}} \\
&= [C(R_2e^{R_2(B+b)} - R_1e^{R_1(B+b)})(e^Z + e^b) - C(R_2^2e^{R_2(B+b)} - R_1^2e^{R_1(B+b)})(e^Z + e^b)] \frac{\psi}{e^{R_2(B+b)} - e^{R_1(B+b)}}
\end{aligned}$$

which is less than zero because  $\frac{\psi}{e^{R_2(B+b)} - e^{R_1(B+b)}} > 0$  and

$$R_2e^{R_2(B+b)} - R_1e^{R_1(B+b)} < R_2^2e^{R_2(B+b)} - R_1^2e^{R_1(B+b)}$$

□

**Lemma 14.** *V has a single inflection point in X.*

*Proof.* Suppose that  $\frac{\partial V}{\partial X} = 0$  which implies that

**Need to Finish**

□

**Lemma 15.** *V is strictly decreasing in B.*

*Proof.* if we take the derivative of V with respect to B, we get

$$\begin{aligned}
\frac{\partial V}{\partial B} &= \Psi \left[ \left( \frac{R_1e^{-R_1(B+b)} - R_2e^{-R_2(B+b)}}{e^{-R_1(B+b)} - e^{-R_2(B+b)}} \cdot \frac{e^Z + e^{-B}}{1 + e^Z} - \frac{e^{-B}}{1 + e^Z} \right) \left( 1 + \frac{c}{r} \right) + \frac{c}{r} \frac{(R_2 - R_1)e^{-(B+b)}}{e^{-R_1(b+B)} - e^{R_2(b+B)}} \cdot \frac{e^Z + e^b}{1 + e^Z} \right] \\
\Rightarrow \frac{\partial V}{\partial B} &\leq 0 \\
\iff 0 &\geq [(R_1e^{-R_1(B+b)} - R_2e^{-R_2(B+b)})(e^Z + e^{-B}) - (e^{-R_1(B+b)} - e^{-R_2(B+b)})e^{-B}] \left( 1 + \frac{c}{r} \right) \\
&\quad + \frac{c}{r} (R_2 - R_1)e^{-(B+b)}(e^Z + e^b)
\end{aligned}$$

Focusing on terms with  $e^Z$ , we have

$$\begin{aligned}
& [R_1 e^{-R_1(B+b)} - R_2 e^{-R_2(B+b)}] \left(1 + \frac{c}{r}\right) + \frac{c}{r} (R_2 e^{-(B+b)} - R_1 e^{-(B+b)}) \\
& \leq R_1 (e^{-R_1(B+b)} - e^{-(B+b)}) - R_2 (e^{-R_2(B+b)} - e^{-(B+b)}) \\
& \leq 0
\end{aligned}$$

Focusing on the remaining terms, we have

$$\begin{aligned}
& [(R_1 e^{-R_1(B+b)} - R_2 e^{-R_2(B+b)})e^{-B} - (e^{-R_1(B+b)} - e^{-R_2(B+b)})e^{-B}] \left(1 + \frac{c}{r}\right) + \frac{c}{r} (R_2 e^{-(B+b)} - R_1 e^{-(B+b)})e^b \\
= & [(R_1 - 1)e^{-R_1(B+b)} - (R_2 - 1)e^{-R_2(B+b)}] \left(1 + \frac{c}{r}\right) e^{-B} + \frac{c}{r} (R_2 e^{-B} - R_1 e^{-B}) \\
= & e^{-B} \left( (e^{-R_1(B+b)}(R_1 - 1) - R_1 - (e^{-R_2(B+b)}(R_2 - 1) + R_2)) \frac{c}{r} + [(R_1 - 1)e^{-R_1(B+b)} - (R_2 - 1)e^{-R_2(B+b)}] \right) \\
\leq & 0
\end{aligned}$$

Hence we can conclude that  $\frac{\partial V}{\partial B} < 0$

□

We first present a technical lemma which will be of use during the following proofs.

**Lemma 16.** *Let  $Z_t$  be a solution to  $dZ_t = \mu(Z_t)dt + \sigma(Z_t)dW_t$ , where  $W_t$  is a standard Brownian motion. Then for the problem*

$$\sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} (d_\tau g_1(Z_\tau) + (1 - d_\tau)g_2(Z_\tau)) | Z_0]$$

*There exists a solution of the form  $\tau = \inf\{t : Z_t \notin (Z_r, Z_a)\}$  with  $d_\tau = \mathbb{1}_{Z_\tau = Z_i}$  for  $Z_i = Z_a$  or  $Z_i = Z_r$ .*

*Proof.* We can note that conditional on stopping, it will be optimal to choose  $d_\tau = 1 \iff g_1(Z_\tau) \geq g_2(Z_\tau)$ . We can define  $g(Z_\tau) = \max\{g_1(Z_\tau), g_2(Z_\tau)\}$  and rewrite the optimal problem as

$$\sup_\tau \mathbb{E}[e^{-r\tau} g(Z_\tau) | Z_0]$$

Because the process  $Z_t$  is Markov and we have exponential discounting (and hence time consistency), the principle of optimality tells us that  $Z_t$  is a sufficient state variable for the optimal policy from time  $t$  onward.

Let us define the value function when current beliefs are  $Z$  as

$$U(Z) := \sup_\tau \mathbb{E}[e^{-r\tau} g(Z_\tau) | Z]$$

As is standard, we can describe  $\tau$  be a continuation region  $\mathcal{C} = \{Z : U(Z) > g(Z)\}$  and a stopping region  $\mathcal{D} = \{Z : U(Z) = g(Z)\}$ . Although the continuation region could take a non-interval form (e.g.,  $\mathcal{C} = [Z_1, Z_2] \cup [Z_3, Z_4]$  where  $Z_1 \leq Z_2 \leq Z_3 \leq Z_4$ ), we are only concerned with the continuation region around  $Z_0$ . Since the diffusion process is continuous, for any  $\mathcal{C}$  which depends only on  $Z$ , there is another continuation region  $\mathcal{C}' = (Z'_1, Z'_2)$  which delivers the same expected value when starting at  $Z_0$  (where  $Z'_1 = \sup_Z \{Z \in \partial\mathcal{C} : Z \leq Z_0\}$  is the highest boundary of  $\mathcal{C}$  which is below  $Z_0$  and  $Z'_2 = \inf_Z \{Z \in \partial\mathcal{C} : Z \geq Z_0\}$  is the lowest boundary point of  $\mathcal{C}$  above  $Z_0$ ). Therefore, there is an optimal stopping policy in the form of a threshold strategy around  $Z_0$ . □

**Lemma 17** (Duality). *Let  $\{\phi_i\}_{i=1}^n$  and  $\Phi$  be bounded  $\mathcal{F}_t^X$  measurable functions and define  $\mathcal{C} := \{(\tau, d_\tau) : \mathbb{E}[\phi_i(\tau, \omega, d_\tau)|Z_0] \leq 0 \forall i = 1, \dots, n\}$ . Suppose that  $\exists(\tau, d_\tau)$  such that  $\mathbb{E}[\phi_i(\tau, \omega, d_\tau)|Z_0] < 0 \forall i = 1, \dots, n$ . Then there is no duality gap, i.e.*

$$\sup_{(\tau, d_\tau) \in \mathcal{C}} \mathbb{E}[\Phi(\tau, \omega, d_\tau)|Z_0] = \inf_{\lambda \in \mathbb{R}_+^n} \sup_{(\tau, d_\tau)} \mathbb{E}[\Phi(\tau, \omega, d_\tau)|Z_0] + \sum_{i=1}^n \lambda_i \mathbb{E}[\phi_i(\tau, \omega, d_\tau)|Z_0]$$

and the infimum is obtained by some finite  $\lambda \in \mathbb{R}_+^n$ . Additionally, we have complementary slackness conditions:

$$\forall i, \lambda_i \cdot \mathbb{E}[\phi_i(\tau, \omega, d_\tau)|Z_0] = 0$$

*Proof.* See Balzer and Janben (2002) or Dokuchaev (1996). □

## 7.2 Symmetric Information

### 7.2.1 Two-Sided Commitment

We begin with a proof of proposition 1.

**Proposition 1.** *The solution to the symmetric information problem with two-sided commitment takes the form of a static-threshold policy. If  $b \neq -\infty$ , then the optimal approval and rejection thresholds  $(B, b)$  are the solution to the following equations:*

$$\frac{\Psi_B(e^{Z_0} - e^{-B}) + e^{-B}\Psi}{\Psi_B(e^{Z_0} + e^{-B}) - e^{-B}\Psi + \frac{c}{r+c}\psi_B(e^{Z_0} + e^{-b})} = \frac{\Psi_b(e^{Z_0} - e^{-B})}{\Psi_b(e^{Z_0} + e^{-B}) - \frac{c}{r+c}\psi e^{-b} + \frac{c}{r+c}\psi_b(e^{Z_0} + e^{-b})}$$

$$\Psi(e^{Z_0} + e^{-B}) + \frac{c}{r+c}\psi(e^{Z_0} + e^{-b}) = \frac{c}{r+c}(1 + e^{Z_0})$$

if  $b = -\infty$ , then  $B = \log(\frac{R_2}{R_1})$ .

*Proof.* We start by proving the conditions of Lemma 18 are met. To see this, note that the stopping policy  $\tau = \epsilon$  and  $d_\tau = 1$  will keep the participation constraint slack for  $\epsilon$  small enough. The other conditions of Lemma 18 are easily checked.

By applying Lemma 18, we can use a Lagrangian in order to turn the primal problem:

$$\begin{aligned} & \sup_{(\tau, d_\tau)} \quad \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_0] \\ & \text{subject to} \\ & P: \quad \mathbb{E}[e^{-r\tau} (d_\tau + \frac{c}{r}) | Z_0] - \frac{c}{r} \geq 0 \end{aligned}$$

into the dual problem

$$\begin{aligned} \mathcal{L} &= \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_0] + \lambda [\frac{c}{r} - \mathbb{E}[e^{-r\tau} (d_\tau + \frac{c}{r}) | Z_0]] \\ &= \mathbb{E}[e^{-r\tau} (d_\tau (\frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} - \lambda) - \lambda \frac{c}{r}) | Z_0] - \lambda \frac{c}{r} \end{aligned}$$

By Lemma 17, we can verify that the solution is of a threshold form. Let  $(B, b)$  be the approval and rejection threshold respectively. Then we know that the primal problem must solve

$$\mathcal{L} = \Psi \frac{e^{Z_0} - e^{Z_0 - B}}{1 + e^{Z_0}} + \lambda [\frac{c}{r} - \Psi \frac{e^{Z_0} - e^{Z_0 - B}}{1 + e^{Z_0}} (1 + \frac{c}{r}) - \psi \frac{e^{Z_0} - e^{Z_0 - b}}{1 + e^{Z_0}} \frac{c}{r}]$$

Taking first-order conditions are rearranging yields the equality in the proposition.  $\square$

Now we move to the proof of Corollary ??

*Proof.* TBF

$\square$

### 7.2.2 One-Sided Commitment

We begin the proof of Theorem 1 by first solving the relaxed problem  $RSM'$ .

**Lemma 18.** *For all  $N$ , the solution to  $RSM_N$  is a static threshold approval policy until  $X_t$  reaches the first binding constraint  $X^1$  for the first time. The continuation value for  $R$  at  $\tau(X^1)$  is  $H(X^1)$ .*

*Proof.* We can now convert the constrained problem  $RSM'$  into the dual form

$$\begin{aligned}\mathcal{L} &= \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_0] \\ &\quad + \sum_i \lambda(X_i) [\mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} (\mathbb{1}_{\tau < \tau(X_i)} d_\tau + \frac{c}{r}) | Z_0] - \mathbb{E}[e^{-r\tau} (d_\tau + \frac{c}{r}) | Z_0]]\end{aligned}$$

with appropriate complementary slackness

$$\forall i, \lambda(X_i) [\mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} (\mathbb{1}_{\tau < \tau(X_i)} d_\tau + \frac{c}{r}) | Z_0] - \mathbb{E}[e^{-r\tau} (d_\tau + \frac{c}{r}) | Z_0]] = 0$$

Let  $X^1 = \max\{X_i : \lambda(X_i) > 0\}$  be the first binding constraint. We will argue that as long as  $X_t$  has not crossed  $X^1$ , the optimal policy by  $R$  must be a threshold policy. To see this, define the value of the optimal stopping rule after crossing  $X^1$  as

$$\begin{aligned}& \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z^1] \\ & + \sum_{Z_i \neq Z^1} \lambda(Z_i) [\mathbb{E}[e^{-r(\tau \wedge \tau(Z_i))} (\mathbb{1}_{\tau < \tau(Z_i)} d_\tau + \frac{c}{r}) | Z^1] - \mathbb{E}[e^{-r\tau} (d_\tau + \frac{c}{r}) | Z^1]] \\ & - \lambda(X^1) \mathbb{E}[e^{-r\tau} (d_\tau + \frac{c}{r}) | Z^1]\end{aligned}$$

Clearly the solution to this problem is independent of the previous history  $h_t$  and hence the value of this objective function above can be written as a function  $K^R(X^1)$  which depends only on  $X^1$ . By the principle of optimality, we know that our solution to the original problem will solve  $K^R(X^1)$  after  $X^1$  has been crossed. We can focus on the region where  $X^1$  has not yet been hit and treat  $K^R(X^1)$  as the continuation value for reaching  $X^1$ . The solution to our Lagrangian must solve

$$\begin{aligned}& \sup_{(\tau, d_\tau)} \mathbb{E}[\mathbb{1}_{\tau < \tau(X^1)} [e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} + \sum_i \lambda(X_i) [e^{-r\tau} (d_\tau + \frac{c}{r}) - e^{-r\tau} (d_\tau + \frac{c}{r})]] \\ & + \mathbb{1}_{\tau > \tau(X^1)} e^{-r\tau(X^1)} [K^R(X^1) + \lambda(X^1) \frac{c}{r}] | Z_0] \\ & = \sup_{(\tau, d_\tau)} \mathbb{E}[\mathbb{1}_{\tau < \tau(X^1)} e^{-r\tau} (d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}}) + \mathbb{1}_{\tau > \tau(X^1)} e^{-r\tau(X^1)} [K^R(X^1) + \lambda(X^1) \frac{c}{r}] | Z_0]\end{aligned}$$

Because we assume that the filtration is right-continuous, we can assume that the value of reaching  $X^1$  is  $G(X^1) := \max\{\frac{e^{Z^1} - 1}{1 + e^{Z^1}}, 0, K^R(X^1) + \lambda(X^1) \frac{c}{r}\}$ . The optimization problem is then

$$\sup_{(\tau, d_\tau)} \mathbb{E}[\mathbb{1}_{\tau < \tau(X^1)} e^{-r\tau} (d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}}) + \mathbb{1}_{\tau > \tau(X^1)} e^{-r\tau(X^1)} G(Z^1) | Z_0]$$

Consider a diffusion process with  $dX_t = \mu(X_t)dt + \sigma(X_t)dW_t$  where

$$\mu(X_t) = \begin{cases} \frac{e^{Z_t} - 1}{1 + e^{Z_t}} \cdot \frac{2\mu^2}{\sigma^2} & \text{if } X_t > X^1 \\ 0 & \text{if } X_t \leq X^1 \end{cases} \quad \text{and} \quad \sigma(X_t) = \begin{cases} \frac{4\mu^2}{\sigma^2} & \text{if } X_t > X^1 \\ 0 & \text{if } X_t \leq X^1 \end{cases}$$

and gain functions

$$g_1(X_t) = \begin{cases} \frac{e^{Z_t} - 1}{1 + e^{Z_t}} & \text{if } X_t > X^1 \\ G(X^1) & \text{if } X_t \leq X^1 \end{cases} \quad \text{and} \quad g_2(X_t) = \begin{cases} 0 & \text{if } X_t > X^1 \\ G(X^1) & \text{if } X_t \leq X^1 \end{cases}$$

and consider the following optimization problem:

$$\sup_{(\bar{\tau}, \bar{d}_\tau)} \mathbb{E}[e^{-r\bar{\tau}} (\bar{d}_\tau g_1(X_{\bar{\tau}}) + (1 - \bar{d}_\tau) g_2(X_{\bar{\tau}})) | Z_0]$$

By applying Lemma 17,, we can conclude that the optimal policy is a threshold policy with thresholds  $(B^1, b^1)$ : approve if  $X_t \geq B^1$ , move to  $H(X_r)$  if  $X_t \leq b^1$ , or move to  $G(X^1)$  if  $X^1 > b^1$  and  $X^1$  is hit before  $B^1$ .

It is easy to see that upon reaching  $X^1$ , the optimal stopping rule is to stop immediately since  $G(X^1) \geq 0$  so  $\tau \geq \tau(X^1) \Rightarrow \tau = \tau(X^1)$ . This allows us to write out (putting in function forms for  $g_1, g_2$ ) the above optimization problem as

$$\sup_{(\bar{\tau}, \bar{d}_\tau)} \mathbb{E}[\mathbb{1}_{\tau < \tau(X^1)} e^{-r\bar{\tau}} \bar{d}_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} + \mathbb{1}_{\bar{\tau} \geq \tau(X^1)} e^{-r\tau(X^1)} G(Z^1) | Z_0]$$

which is identical to the relaxed problem. Therefore the solution to relaxed problem must also be a static threshold policy until  $X^1$  is reached.

Suppose that in the optimal stopping rule  $X^1$  has been reached and the process has not yet been stopped. The optimal stopping rule from this point on is that which deliver  $K^R(X^1)$ . Looking at the definition of  $K^R(X^1)$ , we can see that the optimal stopping rule will be the same regardless of the time  $t$  at which  $X^1$  is reached. Since the stopping rule above  $X^1$  is a threshold rule and the stopping rule when  $X^1$  has been reached is a doesn't depend on when  $X^1$  is reached, we can conclude that the continuation payoff when beliefs reach  $X^1$  (for the first time) to  $A$  from following  $(\tau, d_\tau)$  does not depend on the time  $t$  at which  $X^1$  was first reached. Let  $K^A(X^1)$  be the continuation payoff to  $A$  from stopping rule  $(\tau, d_\tau)$  when  $Z^1$  has just been reached for the first time. Because it doesn't depend on the time at



which  $X^1$  was reached,  $K^A(X^1)$  is a constant. We can write  $A$ 's expected utility from  $(\tau, d_\tau)$  as

$$\begin{aligned}\mathbb{E}[e^{-r\tau}(d_\tau + \frac{c}{r})|Z_0] &= \mathbb{E}[e^{-r(\tau \wedge \tau(X^1))}(\mathbb{1}_{\tau < \tau(X^1)}d_\tau + \frac{c}{r} + \mathbb{1}_{\tau \geq \tau(X^1)}K^A(X^1))|Z_0] \\ &= \mathbb{E}[e^{-r(\tau \wedge \tau(X^1))}(\mathbb{1}_{\tau < \tau(X^1)}d_\tau + \frac{c}{r})|Z_0] + \mathbb{E}[e^{-r(\tau \wedge \tau(X^1))}\mathbb{1}_{\tau \geq \tau(X^1)}K^A(X^1)|Z_0]\end{aligned}$$

By complementary slackness, we know that

$$\begin{aligned}\mathbb{E}[e^{-r(\tau \wedge \tau(X^1))}(\mathbb{1}_{\tau < \tau(X^1)}d_\tau + \frac{c}{r})|Z_0] &= \mathbb{E}[e^{-r\tau}(d_\tau + \frac{c}{r})|Z_0] \\ &= \mathbb{E}[e^{-r(\tau \wedge \tau(X^1))}(\mathbb{1}_{\tau < \tau(X^1)}d_\tau + \frac{c}{r} + \mathbb{1}_{\tau \geq \tau(X^1)}K^A(X^1))|Z_0] \\ &= \mathbb{E}[e^{-r(\tau \wedge \tau(X^1))}(\mathbb{1}_{\tau < \tau(X^1)}d_\tau + \frac{c}{r})|Z_0] + \mathbb{E}[e^{-r(\tau \wedge \tau(X^1))}\mathbb{1}_{\tau \geq \tau(X^1)}K^A(X^1)|Z_0] \\ \Rightarrow 0 &= \mathbb{E}[e^{-r(\tau \wedge \tau(X^1))}\mathbb{1}_{\tau \geq \tau(X^1)}K^A(X^1)|Z_0] \\ &= K^A(X^1)\mathbb{E}[e^{-r(\tau \wedge \tau(X^1))}\mathbb{1}_{\tau \geq \tau(X^1)}|Z_0] \\ \Rightarrow K^A(X^1) &= 0\end{aligned}$$

Therefore upon reaching the first binding constraint, the continuation value of  $A$  must be zero. This means that the optimal stopping rule once  $X^1$  is reached will be the optimal stopping rule which gives  $A$  zero expected utility. This delivers a value of  $H(X^1)$  and hence the optimal stopping rule  $(\tau, d_\tau)$  will take the form of an approval threshold  $B$  and a lower threshold  $b$ , where reaching the lower threshold shifts the optimal stopping rule that which delivers  $H(b)$ .

We can now show that  $K^R(X^1) = H(X^1)$ . Since any  $(\tau, d_\tau)$  which satisfies the constraints of  $H$  is admissible in  $K^R$ , we must have  $K^R(X^1) \geq H^R(X^1)$ . Let  $(\tau^1, d_\tau^1)$  be the optimal stopping rule for  $K^R(X^1)$ . We want to show that  $(\tau^1, d_\tau^1)$  is admissible in the maximization problem of  $H(X^1)$ . From the fact that  $K^A(X^1) = 0$ , we know that  $(\tau^1, d_\tau^1)$  satisfies  $PK(0)$ . To check that the *RDP* conditions are met, we note that for  $X_i < X^1$  we have

$$\begin{aligned}\mathbb{E}[e^{-r(\tau \wedge \tau(X_i))}(d_\tau \mathbb{1}(\tau \leq \tau(X_i)) + \frac{c}{r})|Z_0] &= \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))}(d_\tau \mathbb{1}(\tau \leq \tau(X_i)) + \frac{c}{r})\mathbb{1}(\tau > \tau(X^1))|Z_0] \\ &\quad + \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))}(d_\tau \mathbb{1}(\tau \leq \tau(X_i)) + \frac{c}{r})\mathbb{1}(\tau \leq \tau(X^1))|Z_0] \\ &= \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))}(d_\tau \mathbb{1}(\tau \leq \tau(X_i)) + \frac{c}{r})\mathbb{1}(\tau > \tau(X^1))|Z_0] \\ &\quad + \mathbb{E}[e^{-r\tau}(d_\tau + \frac{c}{r})\mathbb{1}(\tau \leq \tau(X^1))|Z_0]\end{aligned}$$

and similarly

$$\begin{aligned}\mathbb{E}[e^{-r\tau}(d_\tau + \frac{c}{r})|Z_0] &= \mathbb{E}[e^{-r\tau}(d_\tau + \frac{c}{r})\mathbb{1}(\tau > \tau(X^1))|Z_0] \\ &\quad + \mathbb{E}[e^{-r\tau}(d_\tau + \frac{c}{r})\mathbb{1}(\tau \leq \tau(X^1))|Z_0]\end{aligned}$$

Therefore, we can rewrite the initial *RDP* constraint at  $Z_0$  as

$$\begin{aligned}\mathbb{E}[e^{-r(\tau \wedge \tau(X_i))}(d_\tau \mathbb{1}(\tau \leq \tau(X_i)) + \frac{c}{r})\mathbb{1}(\tau > \tau(X^1))|Z_0] &+ \mathbb{E}[e^{-r\tau}(d_\tau + \frac{c}{r})\mathbb{1}(\tau \leq \tau(X^1))|Z_0] \\ &\leq \mathbb{E}[e^{-r\tau}(d_\tau + \frac{c}{r})\mathbb{1}(\tau > \tau(X^1))|Z_0] + \mathbb{E}[e^{-r\tau}(d_\tau + \frac{c}{r})\mathbb{1}(\tau \leq \tau(X^1))|Z_0] \\ &\Rightarrow \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))}(d_\tau \mathbb{1}(\tau \leq \tau(X_i)) + \frac{c}{r})\mathbb{1}(\tau > \tau(X^1))|Z_0] \leq \mathbb{E}[e^{-r\tau}(d_\tau + \frac{c}{r})\mathbb{1}(\tau > \tau(X^1))|Z_0]\end{aligned}$$

Using the Markov property of  $X_t$  and the fact that  $(\tau^1, d_\tau^1)$  doesn't rely on the history up until  $\tau(X^1)$ , we have that

$$\begin{aligned}\mathbb{E}[e^{-r(\tau \wedge \tau(X_i))}(d_\tau \mathbb{1}(\tau \leq \tau(X_i)) + \frac{c}{r})\mathbb{1}(\tau > \tau(X^1))|Z_0] &= \mathbb{E}[e^{-r\tau(X^1)}\mathbb{E}[e^{-r(\tau^1 \wedge \tau(X_i))}(d_\tau^1 \mathbb{1}(\tau^1 \leq \tau(X_i)) + \frac{c}{r})|Z^1]|Z_0] \\ &= \mathbb{E}[e^{-r\tau(X^1)}|Z_0]\mathbb{E}[e^{-r(\tau^1 \wedge \tau(X_i))}(d_\tau^1 \mathbb{1}(\tau^1 \leq \tau(X_i)) + \frac{c}{r})|Z^1]\end{aligned}$$

and similarly for  $\mathbb{E}[e^{-r\tau}(d_\tau + \frac{c}{r})\mathbb{1}(\tau > \tau(X^1))|Z_0]$ . Together, we get that

$$\begin{aligned}\mathbb{E}[e^{-r(\tau \wedge \tau(X_i))}(d_\tau \mathbb{1}(\tau \leq \tau(X_i)) + \frac{c}{r})\mathbb{1}(\tau > \tau(X^1))|Z_0] &\leq \mathbb{E}[e^{-r\tau}(d_\tau + \frac{c}{r})\mathbb{1}(\tau > \tau(X^1))|Z_0] \\ &\Rightarrow \mathbb{E}[e^{-r\tau(X^1)}]\mathbb{E}[e^{-r(\tau^1 \wedge \tau(X_i))}(d_\tau^1 \mathbb{1}(\tau^1 \leq \tau(X_i)) + \frac{c}{r})|Z^1] \leq \mathbb{E}[e^{-r\tau(X^1)}]\mathbb{E}[e^{-r\tau^1}(d_\tau^1 + \frac{c}{r})|Z^1] \\ &\Rightarrow \mathbb{E}[e^{-r(\tau^1 \wedge \tau(X_i))}(d_\tau^1 \mathbb{1}(\tau^1 \leq \tau(X_i)) + \frac{c}{r})|Z^1] \leq \mathbb{E}[e^{-r\tau^1}(d_\tau^1 + \frac{c}{r})|Z^1]\end{aligned}$$

Therefore the solution  $(\tau^1, d_\tau^1)$  satisfies all *RDP* constraints and hence is in the admissible stopping rules for  $H$ . Therefore, we conclude that  $H(X^1) \geq K^R(X^1)$ , which together with our previous observation implies that  $K^R(X^1) = H(X^1)$ .

We have assumed that  $R$  approves at  $B^1$  and moves to  $G(X^1)$  at  $b^1 = Z^1$ . In order to justify this, we note that rejection at both  $B, b$  implies immediate rejection (since otherwise the *DP* constraint for  $\tau(X_0)$  would be binding). We also note that it will never be optimal to approve at  $B$  and  $b$ . We also note that it will never be optimal to approve at both  $B$  and  $b$ ; if  $\frac{e^{Z^1}-1}{1+e^{Z^1}} > 0$ , then immediate approval at  $t = 0$  dominates waiting. To see this, we

translate  $B, b$  into the  $\pi$  belief space as cutoffs  $(\pi_B, \pi_b)$ . It is standard to show that the value function of these two cutoffs satisfies  $J(\pi_0) = (\pi_0(1 - \pi_0))^2 J''(\pi_0)$ . Since  $J(\pi_0) > 0$  by  $\frac{e^{Z^1} - 1}{1 + e^{Z^1}} > 0$ , we must have  $J''(\pi_0) > 0$ . Let  $\alpha = \frac{\pi_B - \pi_0}{\pi_B - \pi_b}$  so that  $\pi_0 = \alpha\pi_B + (1 - \alpha)\pi_b$ . Using the fact that  $J(\pi_B) = 2\pi_B - 1$  and  $J(\pi_b) = 2\pi_b - 1$ , we have

$$\begin{aligned} J(\pi_0) &= J(\alpha\pi_B + (1 - \alpha)\pi_b) \\ &< \alpha J(\pi_B) + (1 - \alpha)J(\pi_b) \\ &= \alpha(2\pi_B - 1) + (1 - \alpha)(2\pi_b - 1) = 2\pi_0 - 1 \end{aligned}$$

Therefore immediate approval is better than waiting.

If  $\frac{e^{Z^1} - 1}{1 + e^{Z^1}} < 0$ , then it is better to move to  $H(Z^1)$  at  $Z^1$  since  $H(Z^1) \geq 0$ . □

Next, we show that the solution to this relaxed problem is indeed as a solution to the full problem  $SM'$ .

The result is stated in the next Lemma.

**Lemma 3.** *As  $N \rightarrow \infty$ , the stopping rule in the second regime is given by the dynamic approval threshold  $\tau = \tau(B(M_t)) \wedge \tau(b^*(0))$  and  $d_\tau = \mathbb{1}(\tau = \tau_{B(M_t)})$  where  $B(M_t) = \underline{B}(M_t)$ .*

*Proof.* We want to now find what the optimal policy is once we have reached the point at which  $A$  is being promised 0 utility. For many values of  $Z$  it will not be optimal to let  $A$  quit (see the example, where reducing the approval threshold can lead to higher utility for  $R$ ).

Suppose that we have just reached the binding constraint  $X^i$ . Then  $R$ 's problem can be written as

$$\begin{aligned} \sup_{(\tau, d_\tau)} & \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z^i] \\ \text{subject to} & \quad RDP \\ & \quad PKI(0) \end{aligned}$$

where, in addition to the dynamic participation constraint we add a promise keeping interval constraint<sup>8</sup>.

---

<sup>8</sup>The use of a promise keeping interval rather than a simple promise keeping constraint (e.g., that expected utility of  $A$  must be equal to  $x$ ) is done only for technical reasons to allow for the application of Lemma 18. One can think of  $PKI$  in terms of a simple promise keeping constraint and the logic of the proof will go through.

$$PKI(x): \quad \mathbb{E}[e^{-r\tau}(d_\tau + \frac{c}{r})|Z^i] - \frac{c}{r} \in [\frac{x}{2}, \frac{3x}{2}]$$

Define the feasible set of  $(\tau, d_\tau)$  which delivers utility in  $[\frac{x}{2}, \frac{3x}{2}]$  when beliefs start at  $Z_0$  to be

$$C(x, Z^i) := \{(\tau, d_\tau) : RDP, PKI(x) \text{ both hold}\}$$

We can then write  $R$ 's problem as

$$\sup_{(\tau, d_\tau) \in C(0, Z^i)} \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z^i]$$

Let  $(\tau^\epsilon, d_\tau^\epsilon) = \operatorname{argsup}_{(\tau, d_\tau) \in C(\epsilon, Z^i)} \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z^i]$  be the solution to these “nearby” problems and define  $(\tau^0, d_\tau^0) := \lim_{\epsilon \rightarrow 0} (\tau^\epsilon, d_\tau^\epsilon)$ .

We first argue that the limiting stopping rule is indeed optimal. For the sake of contradiction, suppose that  $(\tau^0, d_\tau^0)$  was not optimal. Let  $(\tau^*, d_\tau^*)$  be the optimal stopping rule in  $C(0, Z^i)$ . Then we know that

$$\mathbb{E}[e^{-r\tau^*} d_\tau^* \frac{e^{Z_{\tau^*}} - 1}{1 + e^{Z_{\tau^*}}} | Z^i] > \mathbb{E}[e^{-r\tau^0} d_\tau^0 \frac{e^{Z_{\tau^0}} - 1}{1 + e^{Z_{\tau^0}}} | Z^i]$$

Consider a stopping rule which, at time 0, approves with probability  $\epsilon$  and moves to  $(\tau^*, d_\tau^*)$  with probability  $1 - \epsilon$ . Since  $(\tau^*, d_\tau^*) \in C(0, Z^i)$ , this new stopping rule will deliver utility  $\epsilon$  and will be in  $C(\epsilon, Z^i)$ . Since  $(\tau^\epsilon, d_\tau^\epsilon)$  was optimal in the problem with constraint set  $C(\epsilon, Z^i)$ , we know that

$$\mathbb{E}[e^{-r\tau^\epsilon} d_\tau^\epsilon \frac{e^{Z_{\tau^\epsilon}} - 1}{1 + e^{Z_{\tau^\epsilon}}} | Z^i] \geq \mathbb{E}[e^{-r\tau^*} d_\tau^* \frac{e^{Z_{\tau^*}} - 1}{1 + e^{Z_{\tau^*}}} | Z^i](1 - \epsilon) + \epsilon \frac{e^{Z^i} - 1}{1 + e^{Z^i}}$$

Taking the limits of both sides as  $\epsilon \rightarrow 0$ , we have that (using the Dominated Convergence Theorem)

$$\mathbb{E}[e^{-r\tau^0} d_\tau^0 \frac{e^{Z_{\tau^0}} - 1}{1 + e^{Z_{\tau^0}}} | Z^i] \geq \mathbb{E}[e^{-r\tau^*} d_\tau^* \frac{e^{Z_{\tau^*}} - 1}{1 + e^{Z_{\tau^*}}} | Z^i]$$

which is a contradiction. Therefore, it must be that  $(\tau^0, d_\tau^0)$  is optimal.

What does  $(\tau^0, d_\tau^0)$  look like? We first have to know what  $(\tau^\epsilon, d_\tau^\epsilon)$  look like. By using a Lagrangian approach and then applying Lemma 19 (adding the  $PKI$  constraints doesn't qualitatively change the analysis), we get that  $(\tau^\epsilon, d_\tau^\epsilon)$  consists of a static approval threshold  $B$  until beliefs hit some lower bound  $b$ , at which point the stopping rule delivers  $H(b)$ . As

$\epsilon \rightarrow 0$ , we will have  $b = X_1 := \max\{X_j \in XT : X_j < X^i\}$ . By Lemma 15,  $V(B, b, X)$  has at most one inflection point in  $X$  and  $\frac{\partial^2 V}{\partial X^2}$  is bounded away from zero. Therefore, if  $b$  was further below  $X_1$ , then  $V(B, b, X_1) < 0$  for small  $\epsilon$ , which would violate the dynamic participation constraint for  $\tau(X_1)$ .

We can repeat this argument for each time the mechanism delivers  $H(b)$ . This tells us that the optimal mechanism which delivers 0 utility to  $A$  is a series of approval thresholds  $\{B_i\}_{i=1}^N$  and lower thresholds  $\{b_i\}_{i=1}^N$  such that

- $R$  approves if and only if beliefs reach  $B_i$  where  $B_i$  depends only on the highest  $b_i$  which has not yet been reached. We can write the current approval belief as  $B^N(M_t^X)$ .
- $b_i = \max\{X_j \in XT : X_j < b_{i-1}\}$ .
- The expected continuation payoff for  $A$  is zero when  $X_t$  reaches  $b_i$  for the first time.

Now we take the limit of the mechanism which solves our relaxed problem as the grid of thresholds in our constraint set  $XT$  approaches the continuum. To show such a limit exists, we need to show that  $B^N(M_t^X)$  has uniformly bounded variation for each  $N$ .

Define a function  $\bar{B}(b, X)$  as the solution to  $V(\bar{B}(b, X), b, X) = 0$ . At  $X = b_i$ , we must have  $B^N(X_i) = \bar{B}(b_{i+1}, b_i)$ . By Lemma ??,  $\bar{B}(b, X)$  is a well-defined and using the implicit function theorem,  $\bar{B}(b, X)$  is continuously differentiable. Therefore, we can find a  $K \in \mathbb{R}_+$  such that for all  $X_i \in [X_0, Y]$  we have  $|B^N(X_i) - B^N(X_{i+1})| < K|X_i - X_{i+1}|$ . For each  $N$ , the total variation in  $B^N(M_t^X)$  is

$$\sum_{i=1}^N |B^N(X_i) - B^N(X_{i+1})| \leq \sum_{i=1}^N K|X_i - X_{i+1}| = K(X_0 - Y)$$

Since  $\{B^N(M_t^X)\}_{N=1}^\infty$  has uniformly bounded total variation and is uniformly bounded, we can apply Helly's Selection Theorem to conclude that  $B(M_t^X) := \lim_{N \rightarrow \infty} B^N(M_t^X)$  exists.

We now want to argue that  $B(M_t) = \underline{B}(M_t)$ . Suppose this were not the case. We can apply the theorem of the maximum to conclude that the optimal stopping rule which delivers utility  $\epsilon$  is continuous in  $Z_0$  and, taking the limit as  $\epsilon \rightarrow 0$ , is continuous in  $Z_0$  when it must deliver utility 0 to  $A$ . Thus we can conclude that if  $B(M_t) < \underline{B}(M_t) - 2\delta$  for a positive measure of  $M_t$  and some small  $\delta > 0$ , then  $B(M'_t) < \underline{B}(M'_t) - \delta$  for  $M'_t$  close to  $M_t$ . Therefore we can assume that there exists an interval  $[Z_0, Z_0 - \alpha]$  over which  $B(M_t) < \underline{B}(M_t) - \delta$  for  $M_t \in [Z_0, Z_0 - \alpha]$ .

Let  $\bar{B} = \max\{B(M_t) : M_t \in [Z_0, Z_0 - \alpha]\}$ . Then the utility for  $A$  is bounded below by a stopping rule which uses  $\bar{B}$  as long as  $M_t \in [Z_0, Z_0 - \alpha]$ . The utility of this lower bound when starting at  $Z_0$  can be written as

$$\Psi(\bar{B}, -\alpha, 0) \left( \frac{e^{Z_0} + e^{-\bar{B}}}{1 + e^{Z_0}} \right) \left( 1 + \frac{c}{r} \right) + \psi(\bar{B}, -\alpha, 0) \frac{e^{Z_0} + e^\alpha}{1 + e^{Z_0}} \frac{c}{r} - \frac{c}{r}$$

where we use that if the minimum reaches  $-\alpha$ , the continuation payoff must be zero. Since  $A$ 's utility is single peaked in the lower threshold, then for a low enough  $\alpha$ , we will have  $\bar{B} < \underline{B}(Z_0) - \delta$  and the above expression will be strictly positive. This is a contradiction, so it cannot be that  $b > b^*(B) - \delta$ .

Finally, we should check that  $(\tau^0, d_\tau^0)$  is a valid stopping rule and is in  $C(0, Z_0)$ . To verify that this stopping rule yields zero utility to  $\ell$ , we will show that beliefs will cross the lower threshold immediately with probability one. To see this, suppose the contrary. Then by Blumenthal's 0-1 Law, we know that for  $A = \{\exists \epsilon > 0, Z_t \geq Z_0 \forall t \in [0, \epsilon]\}$ , we have  $\mathbb{P}(A) = 1$ . We can rewrite  $A$  as  $A = \{\exists \epsilon > 0, W_t \geq -\frac{e^{Z_0 - \Delta z} - 1}{1 + e^{Z_0 - \Delta z}} \frac{\mu}{\sigma} t \forall t \in [0, \epsilon]\}$ . For small  $t$ , we must have  $W_t \geq -\frac{e^{Z_0 - \Delta z} - 1}{1 + e^{Z_0 - \Delta z}} \frac{\mu}{\sigma} t$ , almost surely. But we know that  $W_t$  is distributed  $N(0, t)$ , therefore with positive probability we will have  $W_t < -\frac{e^{Z_0 - \Delta z} - 1}{1 + e^{Z_0 - \Delta z}} \frac{\mu}{\sigma} t$ . Therefore, it must be that  $\mathbb{P}(A) = 0$ , which implies that for all  $\epsilon > 0$ , we have that there is a  $t \in [0, \epsilon)$  such that  $W_t < -\frac{e^{Z_0 - \Delta z} - 1}{1 + e^{Z_0 - \Delta z}} \frac{\mu}{\sigma} t$ . Therefore, the process crosses the lower boundary almost instantly and the expected payoff to  $\ell$  is zero.

To show that  $(\tau^0, d_\tau^0)$  is a valid stopping rule, we note that  $M_t \in \mathcal{F}_t^Z = \sigma(Z_t)$  and that

$$\{\tau^0 \leq t\} = \cup_{q \in \mathbb{Q} \cap [0, t]} \{Z_t \geq a^S(M_t)\} \in \mathcal{F}_t^Z$$

□

**Lemma 4.** *Let  $(\tau^N, d_\tau^N)$  be the solution to  $RSM_N$  and  $(\tau, d_\tau) = \lim_{N \rightarrow \infty} (\tau^N, d_\tau^N)$ . Then  $(\tau, d_\tau)$  is a solution to  $SM$ .*

*Proof.* We need to verify that after any history  $h_t$ , the continuation value for  $A$  is weakly positive. Since the solution to the relaxed problem depends only on  $X_t, M_t^X$ , we need only check that  $\mathbb{E}[e^{-r\tau}(d_\tau + \frac{c}{r}) | Z_t] - \frac{c}{r} \geq 0$ .

First consider  $(X_t, M_t)$  in the incentivization regime. By the fact that the continuation value at  $X_t = M_t$  is zero, we know that the utility for  $A$  is given by  $V(\underline{B}(M_t), M_t, X_t)$ . Using the definition of  $\underline{B}(M_t)$ , this is weakly positive for all  $X_t$ .

Now consider  $(X_t, M_t^X)$  in the stationary regime. We know that the value for  $A$  in this region is given by  $V(B_h^1, b_h^1, X_t)$ . Since  $b_h^1 = b^*(B_h^1, 0)$ , we know that  $V(B_h^1, b_h^1, X_t)$  will always be weakly positive for  $X_t \in B_h^1$ .

Since the dynamic participation constraints are satisfied after all histories, the relaxed problem *RSM* solves the full problem *SM*. □

**Proposition 2.** *There exists  $\underline{T}, \bar{T}$  such that for all  $t_1 < \underline{T}$  and  $t_2 > \bar{T}$ , the probability of type I error conditional on approval at time  $t_1$  is less than the probability of type I error conditional on approval at time  $t_2$ .*

*Proof.* For small  $\underline{T}$ , the probability that  $X_t$  dips below  $b^1$  and returns to  $B(M_t^X)$  is negligible. Therefore the probability of type I error conditional on approval at  $t < \underline{T}$  will be approximately  $\frac{1}{1+e^{Z^1}}$  where  $Z^1 = Z_0 + \frac{\psi}{\sigma} B^1$ .

For large enough  $\bar{T}$  and  $t > \bar{T}$ , the probability that  $X_t$  has dipped below  $b^1$  is strictly positive. Therefore there is a strictly positive probability that approval is happening when  $X_t$  reaches  $\underline{B}(M_t^X)$ . Therefore the probability of type I error will be strictly below  $\frac{1}{1+e^{Z^1}}$ . □

### 7.2.3 No Commitment

**Proposition 3.** *There exists a pair  $(B, b)$  such that  $R$  only at time  $\tau(B)$  and  $A$  quits at time  $\tau(b)$ . In set-up 1.,  $B > 0$  while in set up 2,  $B = 0$  and  $b = b^*(0)$  and  $A$  quits experimenting when  $Z_t \notin (b, B)$ . The value of experimentation to  $R$  is strictly less than under one- or two-sided commitment.*

*Proof.* Set-up 1 follows directly from Kolb (2016), so let us focus on the case of set-up 2. Let us check whether  $R$  or  $A$  has an incentive to deviate.  $A$  has no incentive to deviate. He is quitting at his optimal level  $b^*(0)$  given  $R$ 's approval threshold and has no incentive to quit early since  $R$  will not approve at any  $Z_t < 0$ . Moreover,  $R$  will always approve at any  $Z_t \geq 0$  whenever  $A$  has quit and so  $A$  will always quit experimenting immediately whenever  $Z_t \geq 0$ .  $R$  also has no incentive to deviate; if he approves at  $Z_t < 0$  he earns a strictly negative payoff while if he rejects early, he gets a payoff of zero (which is equal to his equilibrium payoffs). Since  $A$  and  $R$  have no incentive to deviate, this is an equilibrium. □

**Proposition 4.** *Under set-up 3, the optimal mechanism under one-sided commitment can be implemented as an equilibrium.*

*Proof.* Suppose that  $R$  uses the mechanism from the case of one-sided commitment  $(\tau^*, d_\tau^*)$  and  $A$  uses the following strategy:

- Experiment until  $\tau^*$ .

- If  $d_\tau^* = 0$ , then stop experimenting and do not restart.
- If  $d_\tau^* = 1$ , then stop experimenting and do not restart.

We claim that this is an equilibrium. To see this, let's first consider the incentives of  $R$  to deviate. Suppose that the equilibrium calls for  $R$  to approve at time  $\tau^*$ . If she doesn't approve, then the agent quits experiment at time  $\tau^*$  forever. Since no new learning occurs,  $R$  has a strict incentive to approve immediately at  $\tau^*$  since  $Z_{\tau^*} > 0$ . Suppose  $R$  had a profitable deviation  $\tau'$  such that  $\tau' \leq \tau^*$  and there is some history such that  $\tau'$  approves strictly sooner than  $\tau^*$ .<sup>9</sup> Then  $\tau'$  will not violate any  $DP$  constraints (approving sooner would only slacken the  $DP$  constraints), contradicting the optimality of  $\tau^*$ . Therefore no such deviation can exist.

Next, we consider the incentives of  $A$  to deviate from the proposed equilibrium. Note that under the proposed approval rule, since all the  $DP$  constraints hold,  $A$  has no incentive to quit early. If he were to quit early,  $R$  would believe that  $A$  will restart experimenting immediately and therefore not find it optimal to approve. Moreover,  $A$  has an incentive to stop experimenting at  $\tau^*$  since he believes that  $R$  will approve immediately. In the off-path event that  $R$  doesn't approve,  $A$  believes that  $R$  will approve in the next instant and has no incentive to restart experimentation since it is costly and will not increase the probability of approval.

Since neither  $A$  nor  $R$  have an incentive to deviate,  $(\tau^*, d_\tau^*)$  is indeed an equilibrium. □

## 7.3 Asymmetric Information

### 7.3.1 Two-Sided Commitment

TBF

### 7.3.2 One-Sided Commitment

**Lemma 5.** *The solution to  $RAM_N^h$  is given by a stationary threshold  $B_1^h$  until  $X_t$  reaches either  $B_1^h$  or  $b_1^h$ . If  $X_t = b_1^h$ , then the optimal stopping rule after  $\tau(b_1^h)$  is the solution to  $H^h(X^1)$ .*

*Proof.* As in the symmetric mechanism case, we define a relaxed problem in which we drop all but a finite number of threshold quitting rule constraints (given by  $\{X_i\}_{i=1}^N$ ). We can hypothesize that as long as  $\ell$  continues to experiment,  $h$  will also continue to experiment. Therefore, we relax the problem by dropping all  $h$  related participation constraints (we

---

<sup>9</sup> $R$  will never find it profitable to reject earlier



leave a promise keeping constraint in for in which  $h$  receives  $V_h^2$  if  $\tau = \tau(X_1)$ ). Since  $X_1$  is the first binding constraint, our relaxed problem is equivalent to

$$\begin{aligned}
\text{sup}(\tau, d_\tau) \quad & \mathbb{E}[e^{-r\tau}(d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}})|Z_h] \\
\text{subject to} \quad & \\
\text{RDIC}(Z_\ell, Z_h): \quad & \forall X_i \in \{X_i\}_{i=1}^N \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))}(\mathbf{1}(\tau \leq \tau(X_1))d_\tau + \frac{c}{r})|Z_h - \Delta_Z] \leq \frac{c}{r} + W_\ell \\
\text{RDP}(Z_h): \quad & \forall X_i \in \{X_i\}_{i=1}^N \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))}(\mathbf{1}(\tau \leq \tau(X_1))d_\tau + \frac{c}{r})|Z_h] \leq \frac{c}{r} + W_h \\
\text{PK}_h(V_h): \quad & \mathbb{E}[e^{-r\tau}(d_\tau \frac{c}{r})|Z_h] \geq W_h + \frac{c}{r}
\end{aligned}$$

with associated Lagrangian

$$\begin{aligned}
\mathcal{L} = \quad & \mathbb{E}[e^{-r\tau}(d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}}|Z_h] \\
& + \sum_{i=1}^N \lambda_\ell(X_i)[\mathbb{E}[e^{-r(\tau \wedge \tau(X_i))}(\mathbf{1}_{\tau \leq \tau(X_i)}d_\tau + \frac{c}{r})|Z_0 - \Delta_Z] - \frac{c}{r} - V_\ell] \\
& + \sum_{i=1}^n \lambda_h(X_i)[\mathbb{E}[e^{-r(\tau(X_i) \wedge \tau)}(\mathbf{1}_{\tau \leq \tau(X_i)}d_\tau + \frac{c}{r})|Z_h] - V_h] \\
& + \lambda_h^P[W_h + \frac{c}{r} - \mathbb{E}[e^{-r\tau}(d_\tau + \frac{c}{r})|Z_h]]
\end{aligned}$$

One complication we face is that the expectation terms in the Lagrangian are taken conditioning on different starting beliefs. It is not clear that the previous argument for a threshold approval rule will apply. However, we can note that

$$\begin{aligned}
\lambda \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))}(\mathbf{1}_{\tau \leq \tau(X_i)}d_\tau + \frac{c}{r})|Z_h - \Delta_z] &= \lambda \frac{e^{Z_h - \Delta_z}}{1 + e^{Z_h - \Delta_z}} \mathbb{E}[e^{-r(\tau \wedge \tau(Z_i))}(\mathbf{1}_{\tau \leq \tau(X_i)}d_\tau + \frac{c}{r})|\omega = 1] \\
&+ \lambda \frac{1}{1 + e^{Z_h - \Delta_z}} \mathbb{E}[e^{-r(\tau \wedge \tau(Z_i))}(\mathbf{1}(\tau \leq \tau(X_i))d_\tau + \frac{c}{r})|\omega = 0] \\
&= \lambda \frac{1 + e^{Z_h}}{1 + e^{Z_h - \Delta_z}} [\frac{e^{Z_h}}{1 + e^{Z_h}} \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))}e^{-\Delta_z}(\mathbf{1}(\tau \leq \tau(X_i))d_\tau + \frac{c}{r})|\omega = 1] \\
&+ \frac{1}{1 + e^{Z_h}} \mathbb{E}[e^{-r(\tau \wedge \tau(Z_i))}(\mathbf{1}(\tau \leq \tau(X_i))d_\tau + \frac{c}{r})|\omega = 0]] \\
&= \tilde{\lambda} \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))}(\mathbf{1}(\tau \leq \tau(X_i))d_\tau + \frac{c}{r}) \frac{e^{-\Delta_z} e^{Z_{\tau \wedge \tau(X_i)}} + 1}{1 + e^{Z_{\tau \wedge \tau(X_i)}}}|Z_h]
\end{aligned}$$

where  $\tilde{\lambda} := \lambda \frac{1+e^{Z_h}}{1+e^{Z_h-\Delta_z}}$ . Thus by using a modified Lagrange multiplier, we can convert the Lagrangian into a standard stopping problem, where there is a single expectation. This allows us to write the Lagrangian as

$$\begin{aligned}
\mathcal{L} &= \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_h)] \\
&+ \sum_{i=1}^N \tilde{\lambda}_\ell(X_i) [\mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} (\mathbb{1}_{\tau \leq \tau(X_i)} d_\tau + \frac{c}{r}) \frac{e^{-\Delta_z} e^{Z_{\tau \wedge \tau(X_i)}} + 1}{1 + e^{Z_{\tau \wedge \tau(X_i)}}} | Z_0] - \frac{c}{r} - V_\ell] \\
&+ \sum_{i=1}^n \lambda_h(X_i) [\mathbb{E}[e^{-r(\tau(X_i) \wedge \tau)} (\mathbb{1}_{\tau \leq \tau(X_i)} d_\tau + \frac{c}{r}) | Z_h] - V_h] \\
&+ \lambda_h^P [W_h + \frac{c}{r} - \mathbb{E}[e^{-r\tau} (d_\tau + \frac{c}{r}) | Z_h]]
\end{aligned}$$

Let  $X^1$  be the first binding constraint for  $\ell$ . Then by the same arguments as in Lemma 19, the value of continuing to  $R$  time  $\tau(X^1)$  is given by

$$\begin{aligned}
K_h^R(X^1) &= \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z^1] \\
&+ \sum_{X_i < X^1} \tilde{\lambda}_\ell(X_i) [\mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} (\mathbb{1}_{\tau \leq \tau(X_i)} d_\tau + \frac{c}{r}) \frac{e^{-\Delta_z} e^{Z_{\tau \wedge \tau(X_i)}} + 1}{1 + e^{Z_{\tau \wedge \tau(X_i)}}} | Z^1] - \frac{c}{r} - V_\ell] \\
&+ \sum_{X_i < X^1} \lambda_h(X_i) [\mathbb{E}[e^{-r(\tau(X_i) \wedge \tau)} (\mathbb{1}_{\tau \leq \tau(X_i)} d_\tau + \frac{c}{r}) | Z^1] - V_h] \\
&+ \lambda_h^P [W_h + \frac{c}{r} - \mathbb{E}[e^{-r\tau} (d_\tau + \frac{c}{r}) | Z^1]]
\end{aligned}$$

It is natural to conjecture that  $\ell$ 's constraints will bind before  $h$ 's will. Let us drop all constraints  $DP$  constraints for  $h$  above  $X^1$  (we will check that this valid below). This allows us to write the Lagrangian prior to  $X^1$  as

$$\begin{aligned}
\mathcal{L} = & \mathbb{E}[e^{-r(\tau \wedge \tau(X^1))}(\mathbb{1}(\tau < \tau(X^1))d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} + \mathbb{1}(\tau \geq \tau(X^1))K^R(X^1)) \\
& + \mathbb{1}(\tau < \tau(X^1))[\sum_{i=1}^N \tilde{\lambda}_\ell(X_i)(e^{-r(\tau \wedge \tau(X_i))}(d_\tau + \frac{c}{r}) \frac{e^{-\Delta_z} e^{Z_{\tau \wedge \tau(X_i)}} + 1}{1 + e^{Z_{\tau \wedge \tau(X_i)}}} - \frac{c}{r} - V_\ell) \\
& + \sum_{i=1}^n \lambda_h(X_i)(e^{-r(\tau(X_i) \wedge \tau)}(d_\tau + \frac{c}{r}) - V_h) + \lambda_h^p(W_h + \frac{c}{r} - e^{-r\tau}(d_\tau + \frac{c}{r}))] \\
& + \mathbb{1}(\tau(X_i) \leq \tau) \tilde{\lambda}(X^1) e^{-r\tau(X^1)} \frac{c}{r} \frac{e^{-\Delta_z} e^{Z^1} + 1}{1 + e^{Z^1}} |Z^1]
\end{aligned}$$

Let  $K_h^A := \mathbb{E}[e^{-r\tau}(d_\tau + \frac{c}{r})|Z^1]$ . We can use Lemma 17 and a similar argument to the symmetric information case to conclude  $R$  uses a threshold strategy in which he approves if  $X_t \geq B$  and switches to the optimal stopping rule which deliver  $H^h(X^1)$  if  $X_t \leq X^1$ .

From here, it is easy to verify that dropping the  $h$  DP constraints above  $X^1$  is without loss. Since  $h$  is more optimistic that the state is good, his expected utility when an upper approval threshold is used is higher than that of  $\ell$ , which is positive by assumption that  $X^1$  is the first binding constraint. □

**Lemma 6.** *The optimal mechanism which solves  $H^h(X_t)$  when the current evidence is  $X_t$  is given by a dynamic threshold policy  $\tau = \tau(B^h(M_t^X)) \wedge \tau(\underline{b}(Z_h))$  and  $d_\tau = \mathbb{1}(\tau = \tau(B^h(M_t^X)))$  where*

$$B^h(M_t^X) = \begin{cases} \underline{B}_\ell(M_t^X) & M_t^X \in [b_\ell^*(B_h^2, M_t^X), X_t] \\ B_h^2 & M_t^X \in [b_h^*(B_h^2, M_t^X), b_\ell^*(B_h^2, M_t^X)) \\ \underline{B}_h(M_t^X) & M_t^X \in [b_h^*(\frac{-Z_0\sigma}{\phi}, M_t^X), b_h^*(B_h^2, M_t^X)] \end{cases}$$

When  $DIC(Z_h, Z_\ell)$  is slack,  $B_h^2$  is the same as in the symmetric case with belief  $Z_h$

*Proof.* Lets consider a slightly relaxed problem in which  $\ell$  must receive less than  $\epsilon$  utility when initial beliefs are  $Z_0$ . Our constraints are a bounded promise keeping constraint ( $BPK_\ell$ ) which ensures that  $\ell$ , for any potential quitting rule, cannot get more than  $\epsilon$  utility,  $DP_h$  constraints and a promise keeping constraint to verify that we deliver at least  $V_h^2$  utility to  $h$ . These can be written as

$$\begin{aligned}
BPK_\ell(\epsilon, Z_0): \quad & \forall X_k \in \{X_j\}_{j=i}^N. \mathbb{E}[e^{-r(\tau(X_i) \wedge \tau)}(\mathbb{1}_{\tau \leq \tau(X_i)} d_\tau + \frac{c}{r}) | Z^i - \Delta_Z] - \frac{c}{r} \leq \epsilon \\
DP_h(Z^i): \quad & \forall X_k \in \{X_j\}_{j=i}^N. \mathbb{E}[e^{-r(\tau(X_i) \wedge \tau)}(\mathbb{1}_{\tau \leq \tau(X_i)}(d_\tau + \frac{c}{r}) | Z^i)] \leq K_h^A + \frac{c}{r} \\
PK_h(Z^i): \quad & \mathbb{E}[e^{-r\tau}(d_\tau + (1 - d_\tau)V_h^3 + \frac{c}{r}) | Z_h] \geq K_h^A + \frac{c}{r}
\end{aligned}$$

Suppose that  $BPK_\ell(\epsilon, Z_0)$  is a binding constraint for some  $\epsilon > 0$  (note that this implies that  $BPK_\ell(\epsilon', Z_0)$  will be binding for all  $\epsilon' \in (0, \epsilon)$ ). By applying the same argument as in the previous Lemma, we have that the optimal mechanism will be a static approval threshold until the first time that the next binding constraint  $X^2$  is reached. When  $X^2$  is reached, the mechanism will solve  $H_h(X^2, K_h^{A,2})$  for some  $K_h^{A,2}$ .

By employing a similar argument to the symmetric information case, we can conclude that as  $\epsilon \rightarrow 0$ , the optimal stopping rules converge to one where the approval threshold is given by  $\underline{B}_\ell(M_t^X)$ .

Suppose that  $BPK_\ell(\epsilon, Z^i)$  is not binding for all  $\epsilon > 0$  and let  $(\tau^{DIC, \epsilon}, d_\tau^{DIC, \epsilon})$  be the solution to the problem including  $DIC$  constraints when  $BPK_\ell(\epsilon, Z_0)$  is not binding. Define  $(\tau^{DIC}, d_\tau^{DIC}) = \lim_{\epsilon \rightarrow 0} (\tau^{DIC, \epsilon}, d_\tau^{DIC, \epsilon})$ . Let  $(\tau', d'_\tau)$  be the solution to

$$(\tau', d'_\tau) = \operatorname{argsup}_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z^i]$$

such that

$$\begin{aligned}
DP_h(Z^i): \quad & \forall X_k \in \{X_j\}_{j=i}^N. \mathbb{E}[e^{-r(\tau(X_i) \wedge \tau)}(\mathbb{1}_{\tau \leq \tau(X_i)}(d_\tau + \frac{c}{r}) | Z^i)] \leq K_h^A + \frac{c}{r} \\
PK_h(Z^i): \quad & \mathbb{E}[e^{-r\tau}(d_\tau h + \frac{c}{r}) | Z_h] \geq K_h^A + \frac{c}{r}
\end{aligned}$$

Note that this problem is identical to that of the symmetric mechanism except for the addition of the promise keeping constraint. If  $DIC(Z_h, Z_\ell)$  is not binding, then the promise keeping constraint will be slack and the optimal mechanism will be equal to the solution to the problem with symmetric information. When the promise keeping constraint is not slack, then it is straightforward from our previous work to see that the solution will consist of a stationary regime (with approval threshold  $B_2$ ) followed by an incentivization regime with approval threshold  $\underline{B}_h(M_t^X)$ . Once the incentivization regime has begun for  $h$ , the value to  $R$  is identical to that of the symmetric information case. Hence the project will continue until  $M_t^X = b_h^*(0)$ . □

**Lemma 7.** *The optimal mechanism for  $\ell$  satisfies  $PK(V_\ell, Z_\ell)$  when  $DIC(Z_h, Z_\ell)$  is slack is given by a dynamic approval threshold  $B_\ell(M_t)$ , which is defined as*

$$B^h(M_t^X) = \begin{cases} B_1^\ell & M_t^X \in [b_\ell^*(B_1^\ell, M_t^X), 0) \\ \underline{B}_\ell(M_t^X) & M_t^X \in [b_\ell^*(-\frac{Z_0\sigma}{\phi}, M_t^X), b_\ell^*(B_1^\ell, M_t^X)) \end{cases}$$

for some  $B_1^\ell \in \mathbb{R}$  and  $B_1^\ell$  is less than it would be in the symmetric information case.

*Proof.* When  $DIC(Z_h, Z_\ell)$  is slack, the problem of determining  $\ell$ 's mechanism is identical to that of the symmetric mechanism except for the inclusion of  $PK_\ell(V_\ell)$ . Following the same steps as in the symmetric mechanism, we get that the optimal mechanism is a stationary regime with static thresholds  $(B_\ell^1, b_\ell^1)$  followed by an incentivization regime with a dynamic approval threshold given by  $B(M_t^X, M_t^Z) = \underline{B}(M_t^X, M_t^Z)$ .

The only step we need to verify is that  $b_\ell^1 = b^*(M_t^X, M_t^Z)$ . Suppose for the sake of contradiction that  $b_\ell^1 > b^*(M_t^X, M_t^Z)$ . Intuively the  $DP$  constraints are slack since  $\ell$  would like to keep experimenting at  $b_\ell^1$ . However, given our formulation, the  $DP$  constraints are binding since threshold quitting rules  $\tau(X)$  for  $X < b_\ell^1$  are never reached. Hence, to get at the idea that the  $DP$  constraints are not binding, we write down a relaxed version of the relaxed problem

$$\begin{aligned} & \sup_{(\tau, d_\tau)} \quad \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_\ell] \\ & \text{subject to} \\ RDP(Z_\ell, \epsilon): \quad & \sup_{\tau'} \mathbb{E}[e^{-r(\tau' \wedge \tau)} (\mathbf{1}(\tau \leq \tau') d_\tau + \frac{c}{r}) | Z_\ell] \leq \mathbb{E}[e^{-r\tau} (d_\tau + \frac{c}{r}) | Z_\ell] + \epsilon \end{aligned}$$

By applying the technique of Theorem 1, we can see that the solution will be a stationary regime with static-thresholds  $(B_\ell^{1,\epsilon}, b_\ell^{1,\epsilon})$  followed by an incentivization regime for any  $\epsilon > 0$ . It is easy to show that the Theorem of the Maximum applies here (as it did in the symmetric mechanism) and we must have  $\lim_{\epsilon \rightarrow 0} b_\ell^{1,\epsilon} = b_\ell^1$ . However, for  $\epsilon > 0$  sufficiently small, we will have all quitting rules slack. Hence, when we turn the problem into a Lagrangian, by complementary slackness only the objective function will be left and the first-best solution for  $R$  will solve the Lagrangian (and have  $b_\ell^{1,\epsilon} = \infty$ ). This is a contradiction as we must have  $b_\ell^{1,\epsilon} \rightarrow b_\ell^1$ . □

**Theorem 2.** *When  $DIC(Z_\ell, Z_h)$  is binding and  $DIC(Z_h, Z_\ell)$  is slack, the optimal mechanism is given by a stopping rules  $\tau_i = \tau(B^i(M_t^X)) \wedge \tau(\underline{b}(Z_i))$  and  $d_\tau^i = \mathbf{1}(\tau = \tau(B^i(M_t^X)))$  where  $B^i(M_t^X)$*

are as in Theorem 1. Let  $(B_1^i, b_1^i)$  be the thresholds of the stationary regimes. Then  $B_1^h \leq B_1^\ell$  and  $b_1^h > b_1^\ell$  if  $B_1^h < B_1^\ell$ .

*Proof.* We know that  $b_\ell^1 = b^*(B_\ell^1, Z_\ell)$ . Because  $\ell$  receives zero expected utility conditional on reaching  $b_h^1$ ,  $\ell$ 's expected utility is given by  $V(B_\ell^1, b_\ell^1, Z_\ell)$ . First we want to show that  $B_h \leq B_\ell$ . Suppose that  $DIC(Z_\ell, Z_h)$  is binding. For the sake of contradiction, suppose that  $B_h > B_\ell$ . The utility that  $\ell$  gets from claiming to be  $h$  is bounded above by  $\max_b V(B_h^1, b, Z_\ell)$ . Because  $b_\ell^1 = b^*(B_\ell^1, Z_\ell)$ , then the utility  $\ell$  gets from truthfully reporting his type is given by  $\max_b V(B_\ell^1, b, Z_\ell)$ . By Lemma 16, we know that  $V$  is strictly decreasing in  $B$ . Therefore, we have  $\max_b V(B_h^1, b, Z_\ell) < \max_b V(B_\ell^1, b, Z_\ell)$ , which contradicts  $DIC(Z_\ell, Z_h)$  binding.

Now suppose that  $b_r > b_\ell^1$ . Again let  $B_h^1 > B_\ell^1$ . Now consider the alternative mechanism in which  $\ell$  is given a stationary regime with  $\tilde{B}_\ell^1 = B_h^1$  and  $\tilde{b}_\ell^1 = b_h^1$  (with rejection at  $\tilde{b}_\ell^1$ ). Because this is not optimal, we must have

$$\Psi(B_\ell^1, b_r, 0) \frac{e^{Z_\ell} - e^{-B_\ell^1}}{1 + e^{Z_\ell}} \geq \Psi(\tilde{B}_\ell^1, \tilde{b}_\ell^1, 0) \frac{e^{Z_\ell} - e^{-\tilde{B}_\ell^1}}{1 + e^{Z_\ell}} \Rightarrow \Psi(B_\ell^1, b_r, 0) > \Psi(\tilde{B}_\ell^1, \tilde{b}_\ell^1, 0)$$

Thus the probability of approval when  $\omega = H$  in the stationary regime is higher for  $\ell$  when reporting  $\ell$  rather than  $h$  in the stationary regime. But, because the  $B_\ell^1 < \tilde{B}_\ell^1$ , we also have that the probability of approval when  $\omega = L$  in the stationary regime is higher for  $\ell$  when reporting  $\ell$  rather than  $h$ . But since the expected costs are lower in the stationary regime for  $\ell$  than  $h$ , we cannot have  $DIC(Z_\ell, Z_h)$  binding since the probability of approval is higher and costs are lower. Therefore, we must have  $B_h^1 \leq B_\ell^1$ .

Suppose that  $B_h^1 < B_\ell^1$ . Then we must have  $b_\ell^1 < b_h^1$ ; if we had  $b_h^1 \leq b_\ell^1$ , then  $\ell$  could choose to report  $h$  and quit if the evidence reaches  $b_\ell^1$ . This deviation is identical to lowering the approval standard for  $\ell$ , which strictly increases utility for  $\ell$ . Therefore, in order to not violate  $DIC(Z_\ell, Z_h)$ , we must have  $b_h^1 < b_\ell^1$ . □

**Proposition 5.** For each  $Z_\ell$ ,  $\exists \bar{Z}$  such that  $\forall Z_h > \bar{Z}$ ,  $DIC(Z_h, Z_\ell)$  is slack and  $DIC(Z_\ell, Z_h)$  is binding in the optimal mechanism.

*Proof.* First we establish that  $\exists Z_h$  such that  $DIC(Z_\ell, Z_h)$  binding implies that  $DIC(Z_h, Z_\ell)$  is slack. As  $Z_h \rightarrow \infty$ , we have that the probability of approval conditional on the state being  $H$  ( $p_h(d_\tau = 1|H)$ ) approaches 1 in the optimal mechanism for  $h$ . This is due to the fact that as  $Z_h \rightarrow \infty$ , we have that  $R$  will never reject. For a fixed  $Z_\ell$ , we will have  $p_h(d_\tau = 1|H) > p_\ell(d_\tau = 1|H)$ . Therefore, the optimal mechanism must give  $h$  a lower expected

Let  $\alpha_i = \mathbb{E}[e^{-r\tau} | \omega = H]$  be the expected discounted time till  $\tau$  when  $\omega = H$ . As  $Z_h \rightarrow \infty$ , we have that  $\mathbb{E}[e^{-r\tau}(d_\tau + \frac{c}{r}) | Z_h] \approx \alpha_i(1 + \frac{c}{r})$  since  $p(d_\tau = 1 | H) \rightarrow 1$ . Similarly  $\alpha_i$  approximates the utility to  $R$ .

Suppose that we solve the optimal mechanism dropping  $DIC(Z_h, Z_\ell)$ . Since  $R$  could always offer the  $\ell$ s mechanism to  $h$ , we must have  $\alpha_h \geq \alpha_\ell$ . This will imply that  $DIC(Z_h, Z_\ell)$  is slack. Since  $d_\tau = 0$  with positive probability under  $\ell$ , we have that the utility to  $h$  for claiming to be  $\ell$  is strictly below  $\alpha_\ell(1 + \frac{c}{r}) - \epsilon$  for some  $\epsilon > 0$ .

By applying the Theorem of the Maximum to  $\pi_h$  and taking  $\pi_h \rightarrow 1$ , we get that as long the limiting mechanism is bounded away from immediate approval, there is a  $\bar{Z}$  such that for all  $Z_h > \bar{Z}$ , we have  $\alpha(Z_h) \geq \alpha(Z_\ell) - \delta$ . □

**Lemma 8.** *The optimal mechanism for  $\ell$  when  $Z_\ell < 0$  and  $DIC(Z_h, Z_\ell)$  is binding is given by a dynamic approval threshold  $B_\ell(M_t)$ , which is defined as*

$$B^h(M_t^X) = \begin{cases} B_1^\ell & M_t^X \in [b_r \vee b^*(B_1^\ell, Z_\ell), 0) \\ \underline{B}(M_t^X) & M_t^X \in [b_r, b_r \vee b^*(B_1^\ell; Z_\ell)) \end{cases}$$

for some  $(B_1^\ell, b_r) \in \mathbb{R}_2$ .

*Proof.* Since  $h$  will always have a higher belief than  $\ell$ , we can conjecture that  $h$  will never quit as long  $\ell$  still finds it optimal to experiment. This leads us to define a relaxed problem in which we replace  $DIC(Z_h, Z_\ell)$  with

$$RDIC(Z_h, Z_\ell) : \quad \mathbb{E}[e^{-r\tau}(d_\tau + \frac{c}{r}) | Z_\ell + \Delta_z] \leq V_h$$

It is easily verified using the technique of Theorem 1 that the optimal mechanism will take the form of a stationary threshold rule followed by an incentivization regime. We can then verify that  $RDIC(Z_h, Z_\ell)$  is sufficient for  $DIC(Z_h, Z_\ell)$ . Since  $DIC(Z_h, Z_\ell)$  binds, rejection will take place before  $\underline{b}(Z_\ell)$ .

Suppose that  $b_r < b^*(B_1^\ell, Z_\ell)$ . We want to show that the stationary regime lasts until  $b_\ell^1 = b^*(B_\ell^1, Z_\ell)$ . Suppose not: then by following a similar relaxation of  $DP$  as in Lemma 3, we can conclude that the optimal mechanism will be a stationary regime until  $b_r$ , a contradiction. Therefore we must have  $b_r \vee b_\ell^1 \leq b^*(B_\ell^1, Z_\ell)$ .

The only remaining step is verifying that  $R$  rejects when beliefs reach  $b_r$  and approves when beliefs reach  $B_\ell^1(M_t^X)$  (if we prove this, then it will imply that  $h$  will never quit before  $\ell$  would quit because  $h$  is more optimistic about the probability of approval). This is not

immediate: for example, if  $R$  must deliver zero utility to both  $h, \ell$  when beliefs are  $Z_t > 0$ , then approving after waiting until time  $T = \frac{1}{-r} \log(\frac{c}{r+c})$  would deliver zero utility to both  $h, \ell$  and would be strictly better than rejection.

However, we can rule such a case out because  $Z_\ell < 0$  which implies that  $b_r < 0$ . Suppose that we have reached the last binding quitting constraint for  $\ell$ . Then the problem  $R$  faces is

$$\sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} (d_\tau (\frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} - (\sum_i \lambda_\ell(X_i) + \lambda_\ell - \lambda_h \frac{e^\Delta e^{Z_\tau} + 1}{1 + e^{Z_\tau}})) - \frac{c}{r} (\sum_i \lambda_\ell(X_i) + \lambda_\ell - \lambda_h \frac{e^\Delta e^{Z_\tau} + 1}{1 + e^{Z_\tau}}))] | Z_t]$$

where  $\lambda_\ell, \lambda_h$  are the promise keeping and  $RDIC$  constraints and  $\lambda_\ell(X_i)$  are the  $DP$  constraints.

We begin by noting that it cannot be that  $R$  rejects at both  $B, b$ , since this would violate the  $DIC$  constraints of  $\ell$ .

Suppose that  $R$  chooses to approve at  $Z_\tau = b_r$ . Then because  $\frac{e^{b_r} - 1}{1 + e^{b_r}} < 0$ , we must have  $\sum_i \lambda_\ell(X_i) + \lambda_\ell - \lambda_h \frac{e^\Delta e^{Z_\tau} + 1}{1 + e^{Z_\tau}} < 0$ . If  $R$  also approves at the upper threshold  $B$ , then we can apply a similar argument to that of Lemma 19 to conclude that immediate approval would be better than waiting. Intuitively, if  $\sum_i \lambda_\ell(X_i) + \lambda_\ell - \lambda_h \frac{e^\Delta e^{Z_\tau} - 1}{1 + e^{Z_\tau}} < 0$ , then  $R$  views experimentation as costly. Since approval will happen with probability one, there is no benefit to experimentation costs since the same decision is made regardless of the outcome of experimentation.

Next, we rule out the situation in which  $R$  rejects at  $B$  and approves at  $b$ . Suppose that at least one  $DP$  constraint for  $\ell$  has been reached and we have reached the final binding constraint  $X_N$  for  $\ell$ . Then, right before this  $DP$  threshold has been reached,  $R$  must have been using an upper approval threshold of some  $B$ . Suppose that after  $b$  has been reached,  $R$  begins to use upper threshold  $B'$  but now rejects at  $B'$ . Note that after  $X_N$  has been reached, the continuation value upon reaching  $X_N$  is always the same as the continuation value the first time it has been reached (since  $R$  uses locally stationary threshold rules). Let  $\tau$  be the stopping mechanism approving at threshold  $B$  and let  $\tau'$  be the stopping mechanism rejecting at threshold  $B'$ . By optimality of  $B$  before  $X_N$  has been reached, we have that

$$\begin{aligned} & \mathbb{E}[e^{-r\tau} (\mathbb{1}(\tau = \tau(X_N))(V(X_N) + \lambda(X_N) \frac{c}{r}) \\ & + \mathbb{1}(\tau = \tau(B)) (\frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} - (\sum_i \lambda_\ell(X_i) + \lambda_\ell - \lambda_h \frac{e^\Delta e^{Z_\tau} + 1}{1 + e^{Z_\tau}})) - \frac{c}{r} (\sum_i \lambda_\ell(X_i) - \lambda_h \frac{e^\Delta e^{Z_\tau} + 1}{1 + e^{Z_\tau}}))] | Z_t] \\ & \geq \mathbb{E}[e^{-r\tau'} (\mathbb{1}(\tau' = \tau(X_N))(V(X_N) + \lambda(X_N) \frac{c}{r}) - \frac{c}{r} (\sum_i \lambda_\ell(X_i) + \lambda_\ell - \lambda_h \frac{e^\Delta e^{Z_\tau} + 1}{1 + e^{Z_\tau}}))] | Z_t] \end{aligned}$$



and when evaluated after  $X_N$  has been reached, we have

$$\begin{aligned} & \mathbb{E}[e^{-r\tau}(\mathbb{1}(\tau = \tau(X_N))(V(X_N))) \\ & + \mathbb{1}(\tau = \tau(B))\left(\frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} - \left(\sum_i \lambda_\ell(X_i + \lambda_\ell) - \lambda_h \frac{e^\Delta e^{Z_\tau} + 1}{1 + e^{Z_\tau}}\right) - \frac{c}{r} \left(\sum_i \lambda_\ell(X_i) + \lambda_\ell - \lambda_h \frac{e^\Delta e^{Z_\tau} + 1}{1 + e^{Z_\tau}}\right)\right) | Z_t] \\ & \leq \mathbb{E}[e^{-r\tau'}(\mathbb{1}(\tau' = \tau(X_N))(V(X_N))) - \frac{c}{r} \left(\sum_i \lambda_\ell(X_i) + \lambda_\ell - \lambda_h \frac{e^\Delta e^{Z_\tau} + 1}{1 + e^{Z_\tau}}\right) | Z_t] \end{aligned}$$

Adding these two together, we have that  $\mathbb{E}[e^{-r\tau} \mathbb{1}(\tau = \tau(X_N)) \lambda(X_N) \frac{c}{r}] \geq \mathbb{E}[e^{-r\tau'} \mathbb{1}(\tau' = \tau(X_N)) \lambda(X_N) \frac{c}{r}]$ . Since  $B$  and  $B'$  are thresholds, we must have that  $B' < B$  (since  $\lambda(X_N) < 0$ ). But since it was optimal to approve at  $B$ , we must have

$$\frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} - \left(\sum_{i < N} \lambda_\ell(X_i) + \lambda_\ell - \lambda_h \frac{e^\Delta e^{Z_\tau} + 1}{1 + e^{Z_\tau}}\right) > 0$$

But then adding  $-\lambda(X_N)$  to gain function when we reach  $X_N$ , we increase the gain function for approving. Therefore, for it to be optimal to approve at  $B$  but reject at  $B'$ , we must have that

$$\frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} + \lambda_h \frac{e^\Delta e^{Z_\tau} + 1}{1 + e^{Z_\tau}}$$

is increasing in  $Z$ . But then the fact that approval happens at  $b_r$  implies that the gain function is positive at  $b_r$  and it would be optimal to approve at  $b_r$ , a contradiction. Therefore, we cannot have rejection at  $B'$ .

Assume that no  $DP(Z_\ell)$  constraints are binding. We can solve a relaxed problem in which we only put in  $IC(Z_\ell, Z_h)$  and  $IC(Z_h, Z_\ell)$ . As argued in Proposition ??, the solution to this problem will take the form of an upper threshold  $B_\ell$  and a lower threshold  $b_\ell$ . We need to show that the optimal policy doesn't involve approval at  $b_\ell$ .

Suppose that the optimal policy involved rejection at  $B_\ell$  and approval at  $b_\ell$ . Then the utility of  $h$  will be lower than the utility of  $\ell$ . However, when declaring to be type  $h$ , we have that the utility of  $h$  is higher than the utility of  $\ell$ . If  $V_i^j$  is the utility of type  $i$  declaring to be type  $j$ , the above statement implies that  $V_h^\ell < V_\ell^\ell = V_\ell^h < V_h^h$ , a contradiction of  $DIC(Z_h, Z_\ell)$  binding.

Now suppose that the optimal policy involves approval at both  $B_\ell$  and  $b_\ell$ . Let  $Z(b_\ell)$  ( $Z(B_\ell)$ ) be the belief of  $\ell$  at  $b_\ell$  ( $B_\ell$ ). Approval at  $b_\ell$  and  $B_\ell$  implies that

$$\begin{aligned}\frac{e^{Z(b_\ell)} - 1}{1 + e^{Z(b_\ell)}} - \lambda_\ell + \lambda_h \frac{e^\Delta e^{Z(b_\ell)} + 1}{1 + e^{Z(b_\ell)}} &> 0 \\ \frac{e^{Z(B_\ell)} - 1}{1 + e^{Z(B_\ell)}} - \lambda_\ell + \lambda_h \frac{e^\Delta e^{Z(B_\ell)} + 1}{1 + e^{Z(B_\ell)}} &> 0\end{aligned}$$

Note that if we take the derivative of  $\frac{e^{Z-1}}{1+e^Z} - \lambda_\ell + \lambda_h \frac{e^\Delta e^Z + 1}{1+e^Z}$ , we get

$$\frac{e^Z}{(1 + e^Z)^2} [2 + \lambda_h (e^\Delta - 1)]$$

which is always positive or always negative. Therefore, we must have that  $\frac{e^{Z-1}}{1+e^Z} - \lambda_\ell + \lambda_h \frac{e^\Delta e^Z + 1}{1+e^Z} > 0$  for all  $Z \in [b_\ell, B_\ell]$ .

Now let us formulate the the problem in terms of  $\pi$ .  $V(\pi)$  be the solution to

$$V(\pi) = \mathbb{E}[e^{-r\tau}(2\pi_\tau - 1 - \lambda_\ell + \lambda_h((e^\Delta - 1)\pi_\tau - 1))] + \frac{c}{r}(-\lambda_\ell + \lambda_h \lambda_h((e^\Delta - 1)\pi_\tau - 1))|\pi]$$

As is well known, the infintesimal generator of  $\pi$  is  $\mathbb{L} = \frac{\mu}{2\sigma^2}\pi^2(1 - \pi)^2 \frac{\partial^2}{\partial \pi^2}$  and so we can write  $V(\pi)$  as the solution to

$$rV(\pi) = \frac{\mu}{2\sigma^2}\pi^2(1 - \pi)^2 \frac{\partial^2}{\partial \pi^2} V''(\pi)$$

with boundary conditions given by

$$V(\pi_B) = (2\pi_B - 1 - \lambda_\ell + \lambda_h((e^\Delta - 1)\pi_B - 1)) + \frac{c}{r}(-\lambda_\ell + \lambda_h \lambda_h((e^\Delta - 1)\pi_B - 1))$$

$$V(\pi_b) = (2\pi_b - 1 - \lambda_\ell + \lambda_h((e^\Delta - 1)\pi_b - 1)) + \frac{c}{r}(-\lambda_\ell + \lambda_h \lambda_h((e^\Delta - 1)\pi_b - 1))$$

Since  $V(\pi) > 0$ , it must be that  $V''(\pi) > 0$ . Let  $\alpha = \frac{\pi - \pi_b}{\pi_B - \pi_b}$ . Then we have that

$$\begin{aligned}V(\pi) &= V(\alpha\pi_b + (1 - \alpha)\pi_B) < \alpha V(\pi_b) + (1 - \alpha)V(\pi_B) \\ &= \alpha((2\pi_b - 1 - \lambda_\ell + \lambda_h((e^\Delta - 1)\pi_b - 1)) + \frac{c}{r}(-\lambda_\ell + \lambda_h \lambda_h((e^\Delta - 1)\pi_b - 1))) \\ &\quad + (1 - \alpha)(2\pi_B - 1 - \lambda_\ell + \lambda_h((e^\Delta - 1)\pi_B - 1)) + \frac{c}{r}(-\lambda_\ell + \lambda_h \lambda_h((e^\Delta - 1)\pi_B - 1)) = 2\pi - 1\end{aligned}$$

where the last line is the payoff to immediate approval, a contradiction of the optimality of  $\tau$ . Therefore, we cannot have approval at both  $B, b$ .

This leads to the conclusion that we must have approval at  $B$  and rejection at  $b$ . From this, it is easy to see that  $DP_\ell$  being satisfied implies that  $h$  has no incentive to quit early. Therefore, dropping the quitting constraints for  $h$  was without loss and our solution is optimal. □

**Lemma 9.** *The value of experimentation when the current evidence level is  $M_t^X$  and the minimum is  $M_t^x$  is given the unique solution to  $j'(M_t^X)$ .*

*Proof.* Let  $j(Z, M)$  be given by

$$j(Z, M) = \Psi(B(M), M, Z) \frac{e^Z - e^{Z-B(M)}}{1 + e^Z} + \psi(B(M), M, Z) \frac{e^Z + e^{Z-M}}{1 + e^Z} j(M)$$

which is the solution to the Dirichlet problem (for  $Z \in (M, \underline{B}(M))$ )

$$\begin{aligned} \mathbb{L}_Z j(Z, M) &= rj(Z, M) \\ j(B(M), M) &= B(M) \\ j(M, M) &= j(M) \end{aligned}$$

and  $j(M)$  is the solution to the differential equation (derived using the principle of normal reflection  $\frac{\partial j(Z, M)}{\partial M} \Big|_{Z=M} = 0$ )

$$j'(M) = j(M) \left[ \frac{1}{1 + e^M} - \psi_B B'(M) - \psi_b \right] - \frac{e^M - e^{M-B'(M)}}{1 + e^M} [\Psi_B B'(M) + \Psi_b]$$

with boundary condition  $j(\underline{M}) = 0$ .

We first argue that there is a unique solution to the differential equation for  $j'(M)$ . To do this, we need to establish that  $j$  is Lipschitz continuous in  $j(M)$  and continuous in  $M$ . Once this is established, the Picard-Lindelof Theorem completes the argument.

Lipschitz continuity is clear. Continuity in  $M$  follows from the continuity of  $\underline{B}(M), \underline{B}'(M)$ .

Now we want to argue that  $j(Z_0, M_0) = \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{B(M_\tau)} - 1}{1 + e^{B(M_\tau)}} | Z_0, M_0]$  where  $\tau = \tau(B(M_t)) \wedge \tau(\underline{M})$ . By applying Ito's Lemma to  $j(Z, M)$ , we have

$$\begin{aligned} e^{-rt} j(Z_t, M_t) &= j(Z_0, M_0) + \int_0^t e^{-rs} \left[ \sigma \frac{\partial j(Z_s, M_s)}{\partial Z} dB_s + \frac{\partial j(Z_s, M_s)}{\partial Z_s} \mu(Z_s) ds \right. \\ &\quad \left. + \frac{\partial^2 j(Z_s, M_s)}{\partial Z^2} \frac{\sigma^2}{2} ds - rV(Z_s, M_s) ds + \frac{\partial j(Z_s, M_s)}{\partial M_s} dM_s \right] \\ &= j(Z_0, M_0) + S_t + \int_0^t e^{-rt} [\mathbb{L}_Z V(Z_s, M_s) - rV(Z_s, M_s)] ds \end{aligned}$$

where we use the fact that  $\frac{\partial j(M,M)}{\partial M} = 0$  and  $\Delta M_s = 0$  when  $Z_s > M_s$  and we define  $S_t$  to be

$$S_t = \int_0^t e^{-rt} \frac{\partial j(Z_s, M_s)}{\partial Z_s} dB_s$$

which is a continuous local martingale.

We now note that  $\mathbb{L}_Z V(Z_s, M_s) - rV(Z_s, M_s) = 0$  for all  $Z_s \in (M_s, \underline{B}(M_s))$ . Therefore, we can reduce the above equation for  $e^{-rt} j(Z_t, M_t)$  to

$$e^{-r\tau} j(Z_\tau, M_\tau) = j(Z_0, M_0) + S_\tau$$

where  $\tau = \tau_{B(M_t)} \wedge \tau_{\underline{M}}$ . When the process is stopped, the value  $j(Z_\tau, M_\tau)$  is always equal to  $\mathbb{1}(Z_\tau = B(M_t)) \frac{e^{B(M_t)} - 1}{1 + e^{B(M_t)}}$ . Therefore, we have that

$$e^{-r\tau} \mathbb{1}(Z_\tau = B(M_t)) \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} = e^{-r\tau} j(Z_\tau, M_\tau) = j(Z_0, M_0) + S_\tau$$

Taking expectations of both sides, we have

$$\mathbb{E}[e^{-r\tau} \mathbb{1}(Z_\tau = B(M_t)) \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_0] = j(Z_0, M_0) + \mathbb{E}[S_t | Z_0]$$

It follows from Doob's optimal sampling theorem that  $\mathbb{E}[S_t | Z_0, M_0] = 0$ . Noting that  $\mathbb{E}[e^{-r\tau} d_\tau \frac{e^{B(M_t)} - 1}{1 + e^{B(M_t)}} | Z_0, M_0] = \mathbb{E}[e^{-r\tau} \mathbb{1}(Z_\tau = B(M_t)) \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_0]$ , we can conclude that  $j(Z_0, M_0) = \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{B(M_t)} - 1}{1 + e^{B(M_t)}} | Z_0, M_0]$  □

**Proposition 6.** *Under both one- and two-sided commitment, as  $c \rightarrow 0$  the optimal mechanisms for  $h, \ell$  converge to value of the single decision maker problem for  $R$  with prior  $p(Z_h)\pi_h + (1 - p(Z_h))\pi_\ell$ .*

*Proof.* Consider the case of  $c = 0$ . Let  $\alpha_i = \mathbb{E}[e^{-r\tau_i} \mathbb{1}(d_\tau^i = 1) | \omega = 1]$  and  $\beta = \mathbb{E}[e^{-r\tau_i} \mathbb{1}(d_\tau^i = 1) | \omega = 0]$  be the discounted probability of approval for type  $Z_i$  when  $\omega = 1$  and  $\omega = 0$  (respectively). In order to preserve incentive compatibility, we must have

$$\begin{aligned} \pi_h \alpha_h + (1 - \pi_h) \beta_h &\geq \pi_h \alpha_\ell + (1 - \pi_h) \beta_\ell \\ \pi_\ell \alpha_\ell + (1 - \pi_\ell) \beta_\ell &\geq \pi_\ell \alpha_h + (1 - \pi_\ell) \beta_h \end{aligned}$$

By optimality of  $\tau_h, \tau_\ell$ , we also must have

$$\begin{aligned}\pi_h \alpha_h - (1 - \pi_h) \beta_h &\geq \pi_h \alpha_\ell - (1 - \pi_h) \beta_\ell \\ \pi_\ell \alpha_\ell - (1 - \pi_\ell) \beta_\ell &\geq \pi_\ell \alpha_h - (1 - \pi_\ell) \beta_h\end{aligned}$$

Adding the equations using  $\pi_\ell$ , we get  $\alpha_\ell \geq \alpha_h$ . Doing the same with  $\pi_h$ , we get that  $\alpha_h \geq \alpha_\ell$ . Therefore we must have  $\alpha_h = \alpha_\ell$  and therefore  $\beta_h = \beta_\ell$ . Therefore, it is without loss to offer both types the same mechanism. An application of the Theorem of the Maximum gives us the result as stated in the Proposition.  $\square$

**Proposition 7.** *The value of the optimal mechanism is non-monotonic in  $c$  when  $A$  has private information. When  $A$  has no information, the value of the optimal mechanism is strictly decreasing in  $c$ .*

*Proof.* Suppose that  $\pi_h \approx 1$  and  $\pi_\ell \approx 0$ .

We examine a limiting case where the signal to noise ratio  $\frac{\mu}{\sigma} \rightarrow 0$  and  $c \rightarrow 0$ . We claim that the value of the optimal mechanism is zero. By Lemma ??, we know that the value of the optimal mechanism converges to that of a single decision maker. As  $\frac{\mu}{\sigma} \rightarrow 0$ , learning becomes impossible and the expected time to approval becomes infinitely long.

Next, we want to show that for  $c$  large enough,  $\ell$  will drop out immediately and  $h$  will be approved with strictly positive probability. To do this, we propose a testing rule which approves if and only if  $X_{dt} > X_c$ , where  $X_c$  is set such that

$$-cdt + (1 - rdt) \int_{X_c}^{\infty} \frac{1}{2\sqrt{\sigma^2 dt \pi}} e^{-\frac{(x+\mu dt)^2}{\sigma^2 dt}} dx = 0$$

This will be solved for some  $X_c < 0$  (so that  $h$  is approved more than half the time) and  $h$  will choose to experiment since  $\int_{X_c}^{\infty} \frac{1}{2\sqrt{\sigma^2 dt \pi}} e^{-\frac{(x+\mu dt)^2}{\sigma^2 dt}} dx < \int_{X_c}^{\infty} \frac{1}{2\sqrt{\sigma^2 dt \pi}} e^{-\frac{(x-\mu dt)^2}{\sigma^2 dt}} dx$ . In this case, the value of the project to  $R$  is bounded below by  $\frac{p(Z_h)}{2} > 0$ . Therefore, the value for high  $c$  is higher than the value for low  $c$ .  $\square$