

Experimentation and Approval Mechanisms

Andrew McClellan*

New York University

[Click here for latest version](#)

December 17, 2017

Abstract

We study the design of approval rules when experimentation must be delegated to an agent with misaligned preferences with an application to FDA decision rules. Consider a dynamic learning relationship in which an agent experiments on a project, submitting the outcomes without concealment or distortion to a principal who must decide to approve or reject the project. The agent bears the cost of experimentation, and while she cannot approve, she can unilaterally walk away from the relationship at any time. With such interim participation constraints, the approval threshold is no longer time-stationary. We characterize the history-dependent optimal rule and show that, conditional on continued experimentation, the approval threshold moves downward sporadically. Specifically, the threshold in force at any time depends on the history of play only via the minimum of the belief history and the current belief regarding project success. While we derive this outcome as a full commitment solution, it can be implemented as an equilibrium even in the absence of regulator commitment. When the agent has private information about the state, these results change along one significant dimension: an agent can choose to receive fast-track approval in the form of an initially depressed approval threshold. On expiry of the fast track, however, the threshold jumps up, in contrast to the previous exercise. Thereafter, the monotone dynamics described earlier reappear. Our results help us understand how approval rules optimally change over time and provide a theoretical foundation for both fast-track mechanisms and the possible loosening of later thresholds on longer experimentation paths. They also have empirical implications that run counter to predictions from single-decision maker problems. Using data on FDA approval decisions, we look at the qualitative and quantitative features of fast track programs, and uncover a new relationship between Type I error rates and the length of clinical trials. The agency considerations in our model provide an explanation for these relationships.

* Email: aam618@nyu.edu.

I am deeply indebted to and would especially like to thank Debraj Ray and Ennio Stacchetti for their constant support, guidance and insightful suggestions throughout this project. 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 trade-off between making a decision quickly and accurately. 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. In the approval process the FDA must trade off the need for haste (in order to alleviate the suffering of those currently afflicted) and the need for patience (so as to gather more information in order to prevent the use of harmful drugs by the public). A key element of this environment is that the information is not only generated by nature but is controlled by an agent (the drug company) with incentives that are not aligned with that of the FDA: the drug companies which perform the clinical trials bear the cost of experimentation and may have different preferences on when to approve the drug. Thus the approval rule used by the FDA will determine how much experimentation the company is willing to perform. Additionally, the drug companies may possess private information which the FDA must elicit. 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. The misalignment of incentives will prevent straightforward elicitation: the company, which wants the drug to be approved more quickly, may have an incentive to exaggerate their optimism about the drug's quality. Understanding these agency considerations is important for determining the best approval rule.

In this paper, we revisit the canonical Wald hypothesis-testing problem with the new feature that approval and experimentation are controlled by separate players. We look at how a regulator can design stopping and decision rules (without monetary transfers) which incentivize an agent to perform experimentation and truthfully reveal any private information they have about the state of nature. The players have misaligned incentives in that the regulator prefers more experimentation before making a decision than the agent does. Thus the regulator has additional incentives to consider when designing his optimal stopping rule: in addition to the trade off between haste and discretion, the regulator must also consider how to continually incentivize the agent to continue experimentation. We study what the regulator's optimal mechanism will be in the presence and absence of private information and under different levels of commitment. The misalignment of incentives between the regulator controlling the decision and the agent controlling experimentation will introduce rich and novel dynamics into the optimal stopping problem.

The contribution of this paper is several-fold: first, we look at a novel class of ap-

approval rules and, using new techniques for dynamic contracting, prove their optimality among all approval rules. We find that these rules generate interesting testable predictions (implying that longer experimentation is associated with more erroneous approval) and possess a number of desirable properties. Second, we study the effects of private information and find that it adds new qualitative features to the approval rules. We also empirically study FDA approval decisions; gathering data on the length of clinical trials and Type I error, we find that longer clinical trials are associated with higher error rates, a relationship our theoretical results can justify. Finally, we extend our model to encompass a large class of stopping games and solve for optimal mechanisms.

We start our analysis by investigating the case in which the agent has no private information (*symmetric information*), focusing on how the regulator provides incentives for the agent to experiment. A robust result from the optimal stopping literature with a single decision-maker is the optimality of *stationary threshold* strategies, in which the decision-maker stops whenever the state crosses a time- and history-independent threshold. If the incentives of the regulator and agent were aligned, such a rule would be optimal in our environment. Moreover, we show such that a threshold strategy is again optimal when there are misaligned incentives but the agent can perfectly commit to the regulator’s mechanism (a situation we call *two-sided commitment* in which there is only an *ex-ante participation constraint* for the agent).

However, in most real life situations the agent has an outside option that he could always take; if the clinical trials begin to go poorly and approval appears unlikely, the drug company may pull the plug on the clinical trial. In this setting, the regulator must consider *interim participation constraints* for the agent (which we call *one-sided commitment*). We show that under one-sided commitment, stationary threshold rules are no longer optimal: once the agent is about to quit, the regulator has an incentive to change the approval rule so that the agent chooses to continue experimentation. However, there are many ways the regulator could change the approval rule to give the agent incentives for experimentation. Because the beliefs over the state of nature are changing over time, standard dynamic contracting methods become intractable. Given the richness of the set of approval rules, solving for the optimal rule can appear quite daunting. However, we develop a new method for solving such problems and find that a novel optimal stopping rule which is history-dependent and non-stationary but still remarkably tractable.

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. The approval rule consists of an approval threshold which moves downward sporadically; more specifically, the principal starts with a threshold which initially stays fixed as beliefs drift

downwards. If beliefs descend low enough that the agent would optimally quit against if the current threshold were to remain fixed forever, then the principal begins to lower just enough to incentivize the agent to continue experimentation. When beliefs move higher than the current minimum, the threshold stays fixed (never increasing) and will only decrease when beliefs reach a new low. This drift downwards of the approval threshold is bounded; if beliefs reach a lower fixed threshold, the regulator allows the agent to quit. This gives us an interesting testable prediction because, unlike in the case of a single decision-maker, the probability of Type I error is not constant over time. This mechanism also possess a number of attractive features; for example, we find that the optimal mechanism's thresholds are independent of the initial beliefs, which would not be the case if we were to restrict the principal to only consider stationary threshold rules.

At a first pass, the optimal mechanism may seem to depend strongly on the assumption that the regulator can commit to the approval rule. To understand the role of regulator commitment, we examine what the regulator-optimal equilibrium is and find that our one-sided commitment mechanism is implementable without commitment. Whereas previous literature has restricted attention to simple Markov equilibria which use only current beliefs as a state variable, this finding shows that, if we remove this restriction, we can implement the regulator-optimal using only one additional state-variable (the minimum of beliefs up to the current time). We also show that every Pareto efficient equilibrium can be implemented using a similar mechanism to our own (with a different initial threshold). These equilibria possess the desirable feature that they are Pareto efficient after every history, giving them a strong robustness to renegotiation concerns.

When we give private information to the agent (i.e., a more informed prior at time zero about the state), we find that the optimal mechanism may take the form of a “fast-track” menu option. Low types select into a mechanism which is qualitatively similar to the case with no adverse-selection-i.e., the approval threshold is monotonically decreasing when beliefs reach new lows. However, high types select a qualitatively different mechanism. They are initially given a low approval threshold, but they also face a stationary “failure” threshold. If the failure threshold is reached, the project is not rejected but the approval threshold takes a discrete jump upward (they are thrown out of the fast-track); that is, they are allowed to continue to experiment but face a more stringent standard. This result shows how adverse selection creates a back-loading of costly distortions (raising the approval threshold) for the high type. By introducing a higher approval threshold, the regulator hurts both his and the agent's 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 and increase the probability of quicker approval for the high type.

While our model can be viewed through a normative lens on how decision rules should be designed, we can also look at the model through a positive lens to see if our model can justify phenomenon about decision rules which standard models do not capture. To do this, we gather data on FDA decisions and Type I error rates to explore the relationship between the length of experimentation and the probability of Type I error. In Section 6, we show that our model’s testable predictions (in contrast to standard single-decision maker models) empirically match Type I error rates in FDA approval decisions.

Many of the results will go through for a large class of environments. In Section 7, we show that form of the optimal symmetric information mechanism, a threshold decreasing as beliefs reach new lows, holds in a much richer class of stopping games which allow for more general diffusion processes (not just a learning environment) as well as general payoff functions and outside options. This allows us to generalize our findings to apply to many more principal-agent problems and show that the dynamics we study are a feature of a wide class of such problems. We relate them to a number of other models, such as a promotion model, lobbying game and real-option game.

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. In Section 6 we look at the relationship between the length of clinical trials and Type I error in the context of FDA drug approvals. Section 7 provides extensions and generalize the model to a wider class of diffusion processes and payoff functions.

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 whereas in her model the principal controls experimentation throughout. 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. Madsen (2016) also studies a principal-agent stopping problem with transfers in the case of the quickest detection problem.

Within a model of dynamic information revelation, Orlov et al. (2017) look the interaction between an agent who can supply information and a principal who can exercise an option. 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 analyzed 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 for 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. Frank et. al (2014) and Carpenter et. al (2008) look at the effects of regulatory changes at the FDA on the probability of Type I error.

Henry and Ottaviani (2017) study 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 and study the social costs and benefits of different allocations of authority and commitment.

3 Model

3.1 Environment

Following our motivating example, we study the interaction between a (female) regulator R and a (male) agent A in an infinite-horizon continuous-time model. Both players share a common discount rate $r > 0$. A project, which is up for approval, is of two types: good ($\theta = H$) or bad ($\theta = L$). The regulator wants to approve only good projects. The benefit to approving a good project is a_H^i and the loss to approving a bad project is a_L^i :

	R 's Payoffs			A 's Payoffs	
	H	L		H	L
Approve	a_H^R	a_L^R	Approve	a_H^A	a_L^A
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 this assumption makes the analysis simpler).

In general, the terminal payoffs of R and A might differ. For example, A might only care about the project being approved (if, for example, $a_H^A = a_L^A = 1$). While we allow for general terminal payoffs (and extend the results further in Section 7.2), to simplify notation we take in the rest of the text $a_H^R = a_H^A = 1$ and $a_L^R = -1 \leq a = a_L^A \leq 1$ (we assume $a \leq 1$ so that A weakly prefers approval when $\theta = H$ over approval when $\theta = L$). By changing a , we vary the bias A has over terminal decisions from being aligned with R to always preferring approval. When $a = -1$ the only difference between A and R 's payoffs is that A bears the cost of experimentation. Taking $a = -1$ makes clear the difference in costs is key to the rich dynamics in the optimal mechanism.

Both players begin the game with a common-prior $\pi_0 = \mathbb{P}(\theta = H)$. Over the course of the game, both players learn about the underlying state of nature θ (which we call experimentation). Information about the state is revealed via a Brownian diffusion process with a state-dependent drift. Formally, while experimentation is ongoing, both players publicly observe

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

where $W = \{W_t, \mathcal{F}_t, 0 \leq t < \infty\}$ is a standard one-dimensional Brownian motion¹ on the state space (Ω, \mathcal{F}, P) and $\mu_L = -\mu < 0 < \mu = \mu_H$. By observing X_t , both players update beliefs about the state. After observing X_t , a player's posterior belief is given by Bayes rule as

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

where f_t^θ is the density of a normal distribution with mean $\mu_\theta t$ and variance $\sigma^2 t$. To simplify the belief updating procedure, we note that we can write the beliefs in terms of log-likelihoods²- i.e.,

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

Putting in our terms for π_t , we have (after some algebra)

¹Which implies that $W_0 = 0$ so $X_0 = 0$.

²We subsequently abuse notation by referring to Z_t as beliefs.

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

where $\phi = \frac{2\mu}{\sigma}$, which is called the signal-to-noise ratio, 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

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

This transformation of the belief process is useful because both X_t and the initial Z_0 enter linearly into the current Z_t . As discussed before, in some situation it is reasonable that the agent has more information before the public news process begins. We model such situations by allowing the agent's initial belief Z_0 to be private information.

Definition 1. *The model has **symmetric information** if the initial belief Z_0 of the agent is common-knowledge. The model has **asymmetric information** if the initial belief Z_0 is private information of the agent.*

Note that in a model with asymmetric information, all the private information of A is realized at $t = 0$; that is, all information observed by A after $t = 0$ is also observed by R . For now we postpone further analysis of the asymmetric information model until Section 5 and continue describing the model under symmetric information.

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 and assume it satisfies standard restrictions (see Karatzas and Shreve (1991)). A history $h_t = \omega|_{[0,t]}$ is the realization of a path of X_t (from time 0 to 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., the belief at which she would approve if she were myopic.

3.1.1 Remarks

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

- **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, if X_t corresponds to patient's health during a clinical trial, then the choice of Brownian motion reflects the gradual nature of learning and the noisiness of health outcomes. Even when administered good drugs, a patient's health will still sometimes decline. However, the drift of a patient's 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 us to avoid the situation in which R and A 's beliefs diverge over time, which would make the model intractable.
- **Costs:** 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)). The cost of delaying approval of a good drug is internalized by R in the discounting of future payoffs.

3.2 Mechanism

Our question of interest is to understand how R can optimally design approval standards to elicit the private information of the firm. We will assume that transfers are infeasible (as is in the case in the example of FDA approval decisions). Formally, we allow R to design a stopping mechanism, which consists of a stopping time and a decision rule (to approve or reject conditional upon stopping):

Definition 2. A *stopping mechanism* is a pair $(\tau, d_\tau) \in \mathbb{T} \times \mathbb{D}$, where \mathbb{T} is the set of \mathcal{F}_t^X -measurable stopping rules and \mathbb{D} is the set of \mathcal{F}_t^X -measurable decision rules taking values in $\{0, 1\}$.

When discussing our stopping mechanisms, it will be useful to discuss how the mechanism behaves after a particular history h_t .

Definition 3. For stopping mechanism (τ, d_τ) and history h_t , the *continuation mechanism at h_t* is $(\tau[h_t], d_\tau[h_t])$ and is defined for each ω with history h_t by $\tau[h_t](\zeta_t\omega) = \tau(\omega) - t$ and $d_\tau[h_t](\zeta_t\omega) = d_\tau(\omega)$, where $\zeta_t : \Omega \rightarrow \Omega$ is the shift operator defined such that $X_t(\omega) = X_0(\zeta_t\omega)$.

We can then use our notion of a continuation mechanism to define when the mechanism is Markov with respect to certain variables.

Definition 4. For a set of \mathcal{F} -measurable stochastic processes $\{S^i\}_{i=1}^N$, a stopping mechanism is **(S^1, \dots, S^N)-Markov** if for all histories h_t, h'_t such that S_t^1, \dots, S_t^N are equal, $(\tau[h_t], d_\tau[h_t]) = (\tau[h'_t], d_\tau[h'_t])$.

For example, a stopping rule which approves at the first time $X_t \geq B$ (for some $B \in \mathbb{R}$) will be X_t -Markov, since the continuation stopping time will be the same regardless of the history that led up to X_t .

The agency considerations in the model will impose constraints on the mechanisms which are allowed, which will impose constraints on which problems we consider admissible.

Definition 5. Let $\Delta_C \subseteq \mathbb{T} \times \mathbb{D}$ and define the constrained problem CP to be

$$\sup_{(\tau, d_\tau) \in \Delta_C} \mathbb{E}[e^{-r\tau} g(X_\tau, d_\tau) | X_0].$$

We say that (τ, d_τ) is **admissible with respect to CP** if $(\tau, d_\tau) \in \Delta_C$.

For most of the paper we will endow R with perfect commitment power, allowing us to focus on direct-revelation stopping mechanisms in the asymmetric information model, and we assume that the decision to approve or reject is irrevocable.³ The utility of R for a particular mechanism (τ, d_τ) 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

$$V(\tau, d_\tau, Z_0) = \mathbb{E}[e^{-r\tau} (\pi_0 \tau + a(1 - \pi_\tau)) d_\tau - \int_0^\tau e^{-rt} c dt | Z_0] = \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} - \frac{1 - e^{-r\tau}}{r} c | Z_0]$$

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 mechanism is characterized by a pair of thresholds: the regulator approves if her beliefs ever reach B and rejects in her beliefs ever reach b . We will refer to B as the static approval threshold and b as the static rejection threshold.

³This irrevocability assumption is without loss if we allow experimentation to be stopped and restarted and the agent only pays his flow cost while experimentation is ongoing.

Definition 6. A *static threshold mechanism* is a pair $(b, B) \in \mathbb{R}^2$ such that $b \leq Z_0 \leq B$, $\tau = \inf\{t : Z_t \notin (b, B)\}$ and $d_\tau = \mathbf{1}(Z_\tau \geq B)$. We define $\tau_{\geq}(B) := \inf\{t : Z_t \geq B\}$ and $\tau_{\leq}(b) := \inf\{t : Z_t \leq b\}$.

This focal class of stopping mechanisms are tractable, easily implemented and have the useful property that they allow us to calculate the expected utility for R and A in closed form. To express these utilities, we must know the expected discounted probability that a threshold is reached. The formula⁴ for the discounted probability of reaching B before b when the $\theta = H$ is given by

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

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

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

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

Doing a bit of algebra (see Henry and Ottaviani (2017)) allows us to show that the discounted probability that B is crossed before b if $\theta = L$ is

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

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

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

This allows us to rewrite the utility of R, A when (τ, d_τ) takes a threshold form

$$\begin{aligned} \tilde{J}(B, b, Z_0) &:= J(\tau_{\geq}(B) \wedge \tau_{\leq}(b), \mathbf{1}(\tau_{\geq}(B) < \tau_{\leq}(b)), Z_0) = \frac{e^{Z_0}}{1 + e^{Z_0}} \Psi(B, b, Z_0)(1 - e^{-B}) \\ \tilde{V}(B, b, Z_0) &:= V(\tau_{\geq}(B) \wedge \tau_{\leq}(b), \mathbf{1}(\tau_{\geq}(B) < \tau_{\leq}(b)), Z_0), \\ &= -\frac{c}{r} + \frac{e^{Z_0}}{1 + e^{Z_0}} \left(\Psi(B, b, Z_0) \left(1 + \frac{c}{r} + \left(a + \frac{c}{r}\right)e^{-B}\right) + \frac{c}{r} \psi(B, b, Z_0)(1 + e^{-b}) \right). \end{aligned}$$

We will generally drop the dependence of Ψ, ψ on B, b, Z when the choice of B, b, Z is clear. To simplify notation, we will also use $\Psi_b := \frac{\partial \Psi}{\partial b}$ and $\Psi_B := \frac{\partial \Psi}{\partial B}$ (with similar notation for the derivatives of ψ).

⁴See Stokey (2009).

4 Symmetric Information

We begin our analysis by studying the design of optimal mechanisms under 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 of independent interest.⁵ Our model extends the canonical hypothesis-testing model, which is well-studied in single decision-maker problems, into a mechanism-design framework in which the decision maker faces an additional trade-off in that she incentivize A to continue to experiment. Additionally, we examine how optimal mechanisms change depending on the level of commitment of R , A . Exploring this dimension yields new dynamics in the optimal mechanism.

4.1 Principal Optimal

We begin by solving for the principal optimal mechanism. Note that R has no experimentation costs: Therefore she will never find it optimal to reject: since the news process will never lead R to know for sure that the state 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, and so standard arguments imply that her optimal 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

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

Taking the derivative with respect to B , we get a first-order condition of the optimal approval threshold B^{FB} as

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

which, as we should expect, implies that the optimal threshold choice of R is invariant to the initial belief. This threshold mechanism $(B^{FB}, b^{FB}) = (-\log(\frac{R_1}{-R_2}), -\infty)$ is the optimal mechanism for R and fits the standard result that static-threshold mechanisms are optimal in single decision-maker problems.

⁵Similar settings have been considered in the existing literature: Henry and Ottaviani (2017) study a model with the same payoff structure as ours, but restrict attention to the class of static threshold mechanisms.

4.2 Two-Sided Commitment

We now consider the case in which the control of experimentation and approval is decentralized-i.e., A controls experimentation and R controls the approval decision. This separation of authority introduces a new consideration for R : she must design her approval rule so that it provides incentives for A to perform experimentation. Because A alone bears the cost of experimentation, even when A and R 's terminal payoffs are aligned (i.e., $a = -1$), A will prefer less experimentation than R . This tension in the amount of experimentation both players desire will be the main driving force of our results.

We begin by introducing the agency problem in the most mild way possible and allow R to present a binding contract to A which specifies the amount of experimentation that A must perform. However, unlike the principal optimal solution, A has some say in the design of the mechanism: More specifically, A has authority to accept or reject the mechanism at $t = 0$ (and at that time alone). If A accepts the mechanism, then A commits to continue experimentation until the mechanism specifies that experimentation ends. We define this environment below:

Definition 7. *A mechanism has **two-sided commitment** if once A has accepted (τ, d_τ) , experimentation continues until τ .*

Since A has the option of rejecting R 's proposed mechanism, R 's mechanism must satisfy a participation constraint. Formally, this means that A 's expected utility from R 's proposed mechanism must be at least as high as A 's outside option, which we take to be 0, when rejecting R 's mechanism.

Definition 8. *A mechanism (τ, d_τ) satisfies the **participation constraint** if $V(\tau, d_\tau, Z_0) \geq 0$.*

Since A has no private information about the state and the contract, once agreed to, is binding, the participation constraint will be the sole constraint on R 's choice of a mechanism. Formally, the mechanism design problem faced by R is given by

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

where P is a participation constraint that ensures that the agent finds it optimal to agree to the mechanism.

Our problem takes the form of a constrained optimal stopping problem. A robust finding from the single decision-maker problem (and as seen in our principal-optimal solution) is the optimality of static-threshold rules (e.g., Wald (1947) and Moscaroni and Smith (2001)). However, with agency considerations the participation constraint prevents the use of standard time-consistency arguments that imply the optimality of threshold rules. Despite this difficulty, we are able to show in Proposition 1 that static threshold mechanisms remain optimal.

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:*

$$\frac{\Psi_b(1 - e^{-B})}{\frac{c}{r}[\psi_b(1 + e^{-b}) - \psi e^{-b}] + \Psi_b[(1 + \frac{c}{r} + (a + \frac{c}{r})e^{-B})]} = \frac{\Psi_B(1 - e^{-B}) + e^{-B}\Psi}{a\Psi e^{-B} + \Psi_B(1 + \frac{c}{r} + (a + \frac{c}{r})e^{-B}) + \psi_B(1 + e^{-b})\frac{c}{r}}$$

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

where $B < B^{FB}$ and $P(\tau, d_\tau)$ is binding. If $b = -\infty$, then $B = B^{FB}$.

Proof. All proofs are relegated to the Appendices. □

We see that, as long as the the principal optimal mechanism is not achievable, A must be indifferent between accepting the mechanism and taking his outside option. Any mechanism in which the project is rejected and A strictly prefers to accept the mechanism can be improved upon by lowering the rejection threshold slightly, thereby increasing R utility (since experimentation is always valuable to R) while preserving the participation constraint. This result also establishes that the solution under two-sided commitment is qualitatively the same as in the single decision-maker (thereby justifying some of the focus on threshold mechanisms in the literature), albeit quantitatively different. This quantitative difference is described in the following corollary.

Corollary 1. *If (B^{TS}, b^{TS}) is the optimal two-sided commitment thresholds, then $B^{FB} \geq B^{TS}$ and $b^{TS} \geq b^{FB}$.*

Whenever there there is rejection in the two-sided commitment the option value of experimentation is lower in the two-sided commitment case than in the principal-optimal case. Since with two-sided commitment experimentation will be ended earlier at low

beliefs relative to the principal-optimal, the value of experimentation at intermediate beliefs decreases relative to approval leading to a lower approval threshold.

Interestingly, as in Henry and Ottaviani (2017), the choice of B, b will depend on the initial Z_0 . This differs from the single-decision maker problem. Unlike in the single-decision maker problem, the introduction of the participation constraint at $t = 0$ means that the pure time-consistency of single-decision maker problems is no longer present.

4.3 One-Sided Commitment

In many applications, the assumption of two-sided commitment is unreasonable. If, over the course of the trial, the company becomes pessimistic that the drug will ever be approved, a drug company may decide to cut their losses and end the trial early. While the FDA can commit to approval standards, the ability to compel the company to continue running a clinical trial 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 at any time. Thus, A must be incentivized to continue experimentation even after $t = 0$. We call this environment one of one-sided commitment.

Definition 9. *A mechanism has **one-sided commitment** if after any history h_t A can quit experimenting and take an outside payoff 0.*

We also allow for A , once R has approved to immediately quit and take his outside option of zero rather than the payoff for approval. Since A has the ability at any time to take an outside option, we must reformulate what a participation constraint means in this new 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. The mechanism must then satisfy a dynamic version of the usual participation constraint.

Definition 10. *A mechanism (τ, d_τ) satisfies the **dynamic participation constraint** if after any history h_t the expected continuation to A from (τ, d_τ) is non-negative:*

$$\forall h_t, \quad \mathbb{E}[e^{-r(\tau-t)}(d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_t, h_t] - \frac{c}{r} \geq 0.$$

Because there is a participation constraint for each history and given the richness of potential histories, writing out all the constraints is infeasible. Intuitively, another

way of stating the idea behind the dynamic participation constraint is to say that A never finds it strictly optimal to quit. Suppose that the agent chooses to quit before R ends experimentation-i.e., A chooses a *quitting rule* $\tau' \in \mathbb{T}$ by which he takes his outside option of 0 at time τ' . This strategy will give A an expected utility of

$$\mathbb{E}[e^{-r(\tau \wedge \tau')} (d_\tau \mathbf{1}(\tau \leq \tau') \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})] - \frac{c}{r}.$$

Following this idea and restricting R to choose from mechanisms which incentivize A to not quit prematurely, we define R 's problem as

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

subject to

$$DP : \sup_{\tau' \in \mathbb{T}} \mathbb{E}[e^{-r(\tau \wedge \tau')} (d_\tau \mathbf{1}(\tau \leq \tau') \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_0] \leq \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_0].$$

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. Specifying that the optimal mechanism must not let A quit before R approves or rejects the drug is without loss: If a mechanism allows A to quit after a history h_t , then we could specify another mechanism in which R rejects at the same moment that A quits. This will not change any incentives for A to quit earlier than time t (since quitting and rejection lead to the same payoff for A) and hence the expected payoff to R from the two mechanisms will be the same.

As we formally prove in Lemma 1, the constraint DP is essentially a rewritten version of the definition of dynamic participation constraints from an ex-ante perspective.⁶ To see that the two are equivalent, note that if there was such positive probability placed on a set of histories that A had a strictly negative expected continuation payoff, then the quitting rule

$$\tau' = \inf\{t : \mathbb{E}[e^{-r(\tau-t)} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_t, h_t] = \frac{c}{r}\}$$

(i.e., quitting when the continuation payoff to A is zero) would lead to a strictly higher payoff, violating DP . Similarly, if (τ, d_τ) satisfies dynamic participation constraints, then it also must satisfy DP : If A 's continuation value is always positive, then it cannot be optimal to quit early since A would be giving up positive utility.

⁶With a slight abuse of notation, we will refer to DP as dynamic participation constraints for the rest of the paper.

Lemma 1. *Any mechanism (τ, d_τ) which satisfies all dynamic participation constraints must satisfy DP. For any mechanism $(\tilde{\tau}, \tilde{d}_\tau)$ which satisfies DP, there exists another mechanism which satisfies all dynamic participation constraints and yields the same payoff as $(\tilde{\tau}, \tilde{d}_\tau)$.*

Given the previous results for the principal-optimal and two-sided commitment it seems natural to conjecture the optimality of static threshold mechanisms. Surprisingly, we find that the conjecture fails, as we illustrate with a simple example below to show how threshold rules can be improved upon. Simply put, whenever R rejects in a static threshold mechanism, she would be better off lowering the threshold (“cutting the A some slack”) in order to incentivize A to continue experimenting. By changing the threshold after some histories, R is better able to fine tune the incentives for experimentation to A .

Suppose that R is using a static approval threshold of $B_1 > 0$ (If $B_1 < 0$, then the static threshold mechanism would only approve at beliefs which give R negative utility and, therefore, R would be better off rejecting immediately). Since R would always benefit from continued experimentation at all beliefs below this approval threshold, he will never reject the project before the agent would decide to quit. Let b_Z^* be the value at which A will choose to quit experimenting when R uses a static threshold of B and the current beliefs are Z and let $B_{Z,A}^*$ be the optimal threshold that A would choose if she could choose both the approval and rejection threshold. In general, this choice will depend on the approval threshold B . It is straightforward to show that the *argmax* over b of $\tilde{V}(B, b, Z_0)$ is independent of Z_0 and thus the optimal quitting threshold is only a function of B . More formally, we define $b_Z^*, B_{Z,A}^*$ as

$$\begin{aligned} b_Z^*(B) &:= \operatorname{argmax}_b \tilde{V}(B, b, Z_0), \\ B_{Z,A}^* &:= \operatorname{argmax}_B \tilde{V}(B, b_Z^*(B), Z_0). \end{aligned}$$

Let us assume that $B_1 > B_{Z,A}^*$ (otherwise, R could raise B_1 to $B_{Z,A}^*$ and be strictly better off). In order to satisfy DP it must be that R rejects when $Z_t = b_Z^*(B)$.

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

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

However, upon reaching $b_Z^*(B_1)$, R rejects the project and takes her outside option. At this point, she would be better off if she could convince A to keep experimenting. 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 $b_Z^*(B_1)$ first, then, instead of rejecting, R lowers the approval threshold to αB_1 for some $\alpha < 1$ such that $b_Z^*(B_1) < \alpha B_1$ and $\max\{0, B_{Z,A}^*\} < \alpha B_1$. Note that A will now only quit experimenting if the evidence reaches $b_Z^*(\alpha B_1)$, 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

$$\begin{aligned} & \Psi(B_1, b_Z^*(B_1), Z_0) \frac{e^{Z_0}(1 - e^{-B_1})}{1 + e^{Z_0}} \\ & + \psi(B_1, b_Z^*(B_1), Z_0) \frac{e^{Z_0}(1 + e^{-b_Z^*(B_1)})}{1 + e^{Z_0}} \Psi(\alpha B_1, b_Z^*(\alpha B_1), b_Z^*(B_1)) \frac{e^{b_Z^*(B_1)}(1 - e^{-\alpha B_1})}{1 + e^{b_Z^*(B_1)}}. \end{aligned}$$

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.

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

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

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

This is multiplied by the discounted probability that beliefs hit $b_Z^*(B_1)$ before B_1

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

Note that $\Psi(\alpha B_1, b_Z^*(\alpha B_1), b_Z^*(B_1)) \frac{e^{b_Z^*(B_1)}(1 - e^{-\alpha B_1})}{1 + e^{b_Z^*(B_1)}}$ is strictly positive (since $e^{-\alpha B_1} < 1$). Therefore Mechanism 2 yields a higher payoff for R . Since the choice of B_1 was arbitrary, we can see that any static-threshold mechanisms are not optimal.

Intuitively, R is being too stubborn by sticking to the static threshold B_1 . Once the beliefs have gone low enough, R would be better off by decreasing his approval threshold in order to provide incentives for A to continue experimentation-i.e., 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 ensuring that A doesn't find it optimal to cease experimenting.

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 large, it is not clear if there is salient class of mechanisms that the optimal policy will lie in or know if the optimal policy is feasible to derive. The key difficulty comes from the *sup* over $\tau' \in \mathbb{T}$ in the *DP* constraint. Unlike standard mechanism design, where the agent can deviate by misreporting along some one-dimensional interval, the *DP* constraint allows the agent to deviate across an infinite dimensional class. Overcoming this difficulty is the main challenge of this section.

Remarkably, we are still able to solve for the optimal mechanism and show that it possesses a relatively simple structure. In the interest of keeping notation consistent throughout the following sections, we will describe the mechanism in terms of X_t rather than Z_t .⁷ We define an equivalent version of $b_Z^*, B_{Z,A}^*$ for the process measured in terms of X_t as

$$b^*(B) := b^*(B; Z_0) := [b_Z(Z_0 + \frac{\phi}{\sigma}B) - Z_0] \frac{\sigma}{\phi},$$

$$B_A^* := B_A^*(Z_0) := [B_{Z,A}^* - Z_0] \frac{\sigma}{\phi}.$$

We find that the optimal mechanism turns out to depend on the realized path of X_t only through the current minimum of the evidence path $M_t^X := \min\{X_s : s \in [0, t]\}$. The optimal mechanism approves at the first time that X_t crosses a threshold $B(M_t^X)$. We can think of the optimal mechanism as consisting 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 the first time X_t crosses $B(M_t^X)$ 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. The optimal mechanism features a gradual movement downward of the approval threshold from R 's preferred level to A 's preferred level. We define $\underline{B}(X)$ is the lowest static approval threshold B above A 's preferred threshold B_A^* such that A would choose to

⁷This will be useful when we introduce asymmetric information so that we don't have to describe the mechanism in terms of the beliefs of both A and R .

quit at evidence level X when R is using a static threshold mechanism with approval threshold B -i.e.,

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

Formally, the resulting optimal mechanism is stated in the following theorem.

Theorem 1. *The optimal stopping mechanism under symmetric information is given by the stopping rule $\tau = \inf\{t : X_t \geq B(M_t^X)\} \wedge \tau_{\leq}(b^*(B_A^*) \vee b^*(-\frac{\sigma}{\phi}Z_0))$ and $d_{\tau} = \mathbf{1}(X_{\tau} \geq B(M_{\tau}^X))$ where $B(M_t^X)$ is defined as*

$$B(M_t^X) = \begin{cases} B^1 & M_t^X \in [b^*(B^1), 0] \\ \underline{B}(M_t^X) & M_t^X \in [b^*(B_A^*) \vee b^*(-\frac{\sigma}{\phi}Z_0), b^*(B^1)]. \end{cases}$$

Note that A can never be incentivized to experiment at beliefs below $b^*(B_A^*)$: in his first best, A would be quitting at $Z < b^*(B_A^*)$, a payoff he can replicate even when he doesn't have control of the approval threshold by quitting immediately. Moreover, R will never choose to lower the approval threshold below her myopic threshold $-\frac{\sigma}{\phi}Z_0$ (since doing so would only guarantee her a negative payoff and she would be better off letting the agent quit). Thus experimentation is not extended indefinitely but ends whenever the approval threshold reaches either the agent optimal level or R 's myopic cutoff.

We note several interesting features of the optimal mechanism:

- **Monotonicity:** The approval threshold only drifts downward and only changes in order to provide incentives to keep A from quitting. The times at which the current approval threshold decreases are stochastic (since they are a function of M_t^X).
- **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 $X_t = M_t^X$. This means that even though the threshold is moving towards A 's preferred level, this does not increase A 's time zero utility since whenever the threshold decreases, A would be just as well off quitting immediately.
- **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 the initial beliefs Z_0 are. This property, which is common in single-decision

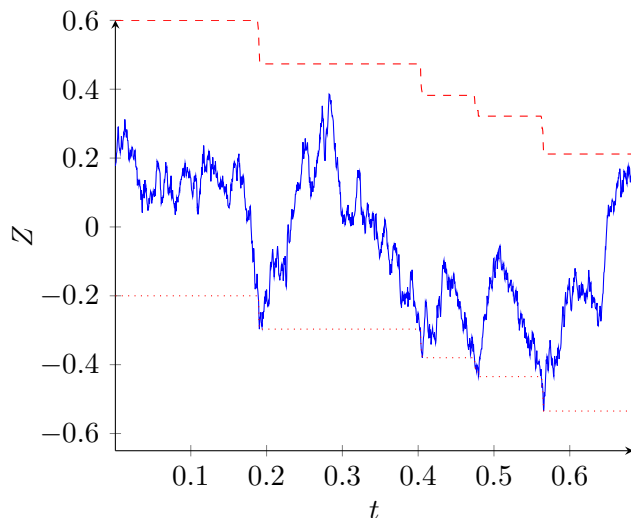


Figure 2: The graph above corresponds to the changing approval threshold for a particular realization of X_t . The upper dashed line corresponds to the current approval threshold. This approval threshold will stay at the same level until X crosses the current minimum of the process, which is given by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

maker problems, is absent if we were to restrict attention to a choice over static threshold rules (see Henry and Ottaviani (2017)).

The rest of this section will be devoted to sketching out the ideas of the proof. The tractability afforded by continuous time has led to a growing literature in mechanism and contract design. Our approach differs from the standard continuous-time approach (e.g., Sannikov (2007), where 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 they are learning about the underlying state, using the HJB approach in our model would require carrying both a state variable of agent continuation and current beliefs about the state; finding the solution to the HJB equation would require solving a difficult partial differential equation and is impractical for analyzing the optimal stopping rule. Instead, we use a different approach by finding a relaxed problem over which Lagrangian techniques work well. This method allows to more easily derive the qualitative features of the optimal mechanism. Using these qualitative features, we are then able to explicitly pin down the form of the optimal mechanism. 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),

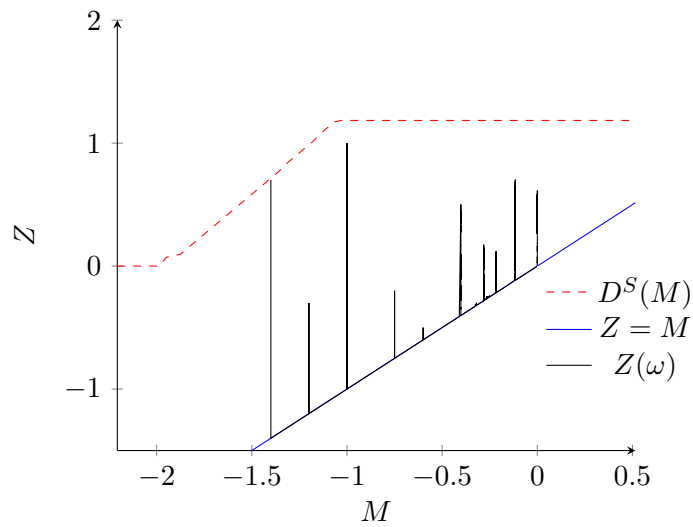


Figure 3: The dashed line gives the approval threshold as a function of M^X and the solid straight line marks the 45 degree line where $X = M^X$. The dashed line is initially constant in M^X during the stationary regime while it decreases in M^X for the incentivization regime. The lines coming up from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

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 use of Lagrangian techniques allows us to convert our constrained problem SM (the primal problem) into an unconstrained form (the dual problem). The key technical difficulty lies in the fact that when checking DP , we must consider all possible quitting rules τ' which A might use. For an arbitrary stopping rule (τ, d_τ) , we might try to solve for the optimal τ' which A would use. However, even for simple time-dependent stopping rules, solving for τ' is difficult and cannot be calculated in closed-form. Given the richness of the set of available (τ, d_τ) , which may be history-dependent, solving for τ' 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.

In order to accommodate these complications, we will define a relaxed problem over which our Lagrangian approach will prove useful. We will limit the set of DP constraints to consider only those quitting rules of a particular class, which we call threshold quitting rules.

Definition 11. A uses a **threshold quitting rule** at X_i if he quits at time $\tau(X_i) := \tau_{\leq}(X_i)$ (We drop the \leq in the subscript for notational convenience).

The payoffs to A of quitting early are equal to those of rejection. Therefore, A evaluates the (τ, d_τ) when following quitting rule $\tau(X_i)$ as equivalent to the mechanism $(\tau \wedge \tau(X_i), d(X_i))$ where we define $d(X_i) := d_\tau \mathbf{1}(\tau \leq \tau(X_i))$. We further restrict attention to a finite number of such quitting rules. Let $\mathcal{T}_N = \{X_i\}_{i=0}^N$ such that $X_0 = 0$ and $X_{i+1} = X_i + \frac{X}{N}$ for some $\underline{X} \in \mathbb{R}_-$ so that the solution to two-sided commitment problem starting at $Z_0 = \frac{-\phi}{\sigma} \underline{X}$ would be immediate rejection.

In order to apply our Lagrangian technique, we define a relaxed problem in which we only impose that A cannot profitably quit early when restricted use a finite number of threshold quitting rules. This problem is formally defined as

$$\begin{aligned}
[RSM_N] : & \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_0] \\
& \text{subject to } \forall X_i \in \mathcal{T}_N \\
RDP(X_i) : & \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} (d_\tau(X_i) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_0] \leq \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_0].
\end{aligned}$$

We note that because 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 .

Let us define the set of RDP constraints which are binding when using the optimal (τ, d_τ) for RSM_N as

$$\mathcal{B}_N = \{X_i \in \mathcal{T}_N : \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} (d_\tau(X_i) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | X_0] = \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | X_0]\}.$$

We will write $\mathcal{B}_N = \{X^1, \dots, X^{|\mathcal{B}_N|}\}$ which are ordered from largest to smallest (we will generally use superscripts to refer to binding constraints).

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 τ may also be a complex history-dependent quitting rule. We should also note that we are not restricting the solution of RSM_N to be a threshold policy. Instead, we are only checking that A has no incentive to deviate to a threshold quitting rule rather than obediently following R 's proposed mechanism.

We can now use Lemma 16 from the Appendix in order to transform our primal problem RSM_N into its corresponding dual problem by constructing an associated Lagrangian with Lagrange multipliers $\{\lambda(X_i)\}_{i=0}^N \in \mathbb{R}_-^{N+1}$

$$\begin{aligned} \mathcal{L} = & \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_0] \\ & + \sum_{i=0}^N \lambda(X_i) (\mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} (d_\tau(X_i) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_0] - \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_0]). \end{aligned}$$

For an appropriate choice of $\{\lambda(X_i)\}_{i=0}^N$, the solution to the associated Lagrangian will solve the primal problem RSM_N and will have complementary slackness conditions

$$\forall X_i \in \mathcal{T}_N, \lambda(X_i) (\mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} (d_\tau(X_i) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_0] - \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_0]) = 0.$$

This implies that we can rewrite the Lagrangian using only the binding constraints:

$$\begin{aligned} \mathcal{L} = & \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_0] \\ & + \sum_{j=1}^{|\mathcal{B}_N|} \lambda(X^j) (\mathbb{E}[e^{-r(\tau \wedge \tau(X^j))} (d_\tau(X^j) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_0] - \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_0]). \end{aligned}$$

This dual version of the problem drastically simplifies the analysis since it allows us to study what is effectively a single decision-maker problem. And while the selection of appropriate multipliers $\{\lambda(X_i)\}_{i=0}^N$ is difficult, the qualitative properties that we will derive from the analysis of the Lagrangian for arbitrary multipliers will allow us to pin down the form of the optimal solution.

We will decompose the problem into the time before the first binding quitting rule X^1 has been reached and the time after the first quitting rule has been reached (i.e., $\tau(X^1)$). If A is truly indifferent between quitting and continuing at $\tau(X^1)$, then his continuation value at $\tau(X^1)$ should be zero. We denote the continuation value for R of the mechanism which delivers a continuation value of zero to A when the evidence level is X_t by $H_N(X_t)$. This is defined formally as

$$\begin{aligned} [H_N(X_t)] : & \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_t] \\ & \text{subject to } RDP(X_i) \quad \forall X_i \in \mathcal{T}_N \\ PK(0) : & \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_t] - \frac{c}{r} = 0, \end{aligned}$$

which is similar to our original problem RSM_N except for an added promise keeping constraint $PK(0)$ which ensures that the expected utility of A for continuing until R rejects is zero.

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 the first binding constraint X^1 is reached. Moreover, we see that A will truly be indifferent between quitting and continuing at $\tau(X^1)$.

Lemma 2. *The solution to RSM_N is a static threshold approval policy until $\tau(X^1)$. The continuation mechanism at $\tau(X^1)$ is the solution to $H_N(X^1)$.*

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 $H_N(X^1)$ is independent of the history up until time $\tau(X^1)$. What happens after $\tau(X^1)$ is completely bundled into the value $H_N(X^1)$ and therefore doesn't affect the choice of R before $\tau(X^1)$ except through the value of $H_N(X^1)$. This also implies that whenever X^1 is reached, A is indifferent between quitting and continuing to experiment. This property comes from the particular form of the optimal stopping rule and is not true of all stopping rules that R could use. For example, if R was using a deterministic stopping rule $\tau = T$ for some $T \in \mathbb{R}_+$, then it could be that A 's $RDP(X^1)$ constraint was binding in expectation at $t = 0$, but, when X^1 is first reached, A could have a strictly positive or negative continuation value. The stationarity of the optimal stopping rule prior to $\tau(X^1)$ is key for proving the indifference of A at $\tau(X^1)$.

We must now solve for the optimal mechanism in the second regime (i.e. that which solves the problem which deliver $H_N(X^1)$). The proof of Lemma 3 mirrors that of Lemma 2. The difference is that when we are trying to solve for the optimal mechanism which deliver R payoff $H_N(X^1)$ (and A a payoff of zero), we have added a promise keeping constraint to the RDP constraints. This additional constraint can be incorporated in a Lagrangian similar to the RDP constraints and, by repeated application of the arguments used in Lemma 2, we can show the optimal mechanism to be a decreasing threshold in M_t^X .

Lemma 3. *As $N \rightarrow \infty$, the limit of the stopping mechanisms which solve $H_N(X^1)$ is given by $\tau = \inf\{t : X_t \geq \underline{B}(M_t^X)\} \wedge \tau(b^*(B_A^*) \vee b^*(-\frac{\sigma}{\phi}Z_0))$ and $d_\tau = \mathbf{1}(X_\tau = \underline{B}(M_t^X))$.*

The proof establishes that for each N , the stopping mechanism is (X_t, M_t^X) -Markov. This property, which we will see repeatedly in the next session as well, is crucial for establishing that our relaxed problem is a solution to the full problem. While the mechanism is history-dependent, the mechanism can be determined with only two state-variables, making the mechanism tractable for calculation and implementation.

In order to pin down the exact form of $B(M_t^X)$, we use complementary slackness as our grid of threshold quitting rules becomes increasingly fine. In addition to telling us the form of $B(M_t^X)$, Lemma 3 also states how long $B(M_t^X)$ will continue to decrease. Note that the role of decreasing B is to incentivize further experimentation. However, if the approval threshold descends below the myopic threshold of R at $X_t = \frac{-\sigma}{\phi}Z_0$ (where $Z_t = 0$), then R would rather cease experimenting than lower the approval threshold (since, if she lowers it, she will only approve at beliefs which give a negative utility). Moreover,

the argument we made in the illustration of the non-optimality of static threshold rules implies that it is never optimal to end experimentation when $B(M_t^X)$ is above R 's myopic threshold. These arguments imply that R will reject the project precisely when $B(M_t^X)$ is equal to R 's myopic threshold.

Having derived the solution to RSM_N , we need to check that the solution to the relaxed problem solves the full problem SM as $N \rightarrow \infty$. Because our mechanism is (X_t, M_t^X) -Markov, we can show that the best response of A will also be (X_t, M_t^X) -Markov; checking that all dynamic participation constraints is then relatively simple.

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

The optimal mechanism in Theorem 1 can also be written in belief-space, giving an approval threshold of $B_Z(M_t^Z) = \frac{\phi}{\sigma} B((M_t^Z - Z_0) \frac{\sigma}{\phi}) + Z_0$. Looking at the optimal mechanism in belief space, we can ask how the optimal mechanism depends on the initial belief Z_0 . Surprisingly, we find that $B_Z(\cdot)$ is independent of Z_0 , which stands in contrast to our model with two-sided commitment as well as other models with agency considerations, such as Henry and Ottaviani (2017), who find that the initial conditions do matter for determining the optimal mechanism for the principal. This invariance the initial belief is a common feature of single-decision maker problems and implies a dynamic consistency in the state variables (X_t, M_t^X) . In the case of Henry and Ottaviani (2017), however, their mechanism is exogenously restricted to be a threshold mechanism. Our result implies that dynamic consistency is restored when we allow the principal to consider all possible mechanisms.

Lemma 5. *The optimal approval threshold in belief-space $B_Z(M_t^Z)$ is independent of Z_0 and depends only on $M_t^Z := \min\{Z_s : s \leq t\}$.*

The most notable feature about the optimal mechanism is that the approval threshold is changing with M_t^X . For lower M_t^X , the lower approval threshold increases the probability of Type I error, in contrast to static threshold mechanisms (as in the case of two-sided commitment) in which the probability of error conditional upon approval is constant. This observation gives us predictions our model makes to an outside analyst how only sees the length of experimentation conditional on approval and, if the project is approved, the true state (i.e., the analyst cannot observe the realization of X_t). Our model predicts that the analyst will predict a higher probability of Type I error (i.e., approving a bad project) for projects which have taken a long time to be approved when compared to projects which were approved quickly. In many contexts this fits a natural

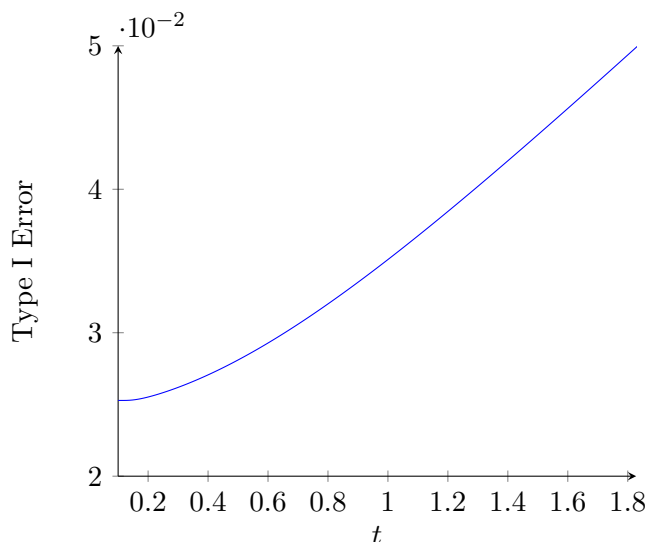


Figure 4: We see that the likelihood of Type I error given approval at time t is increasing.

intuition. For example, 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. As we will see in Section 6, this prediction will be matched empirically in data on FDA approval times and Type I error probabilities.

4.4 No Commitment

At first glance, the optimal mechanism under one-sided commitment seems to require a great deal of commitment: R agrees to permanently lower the approval threshold, even though she would be better off raising it back to its initial level if beliefs drift back up. We might naturally wonder how much R loses if she cannot commit to the optimal mechanism. To answer this question, we need to think about the exact details of the model without commitment. More specifically, we need to know the precise sequence of events when A stops experimenting: as we will show, these details are crucial for determining the equilibrium outcome. We introduce several different set-ups below:

- (I): A can irrevocably quit experimenting at any time t and R cannot approve after A has quit.
- (II): A can irrevocably quit experimenting at any time t and R can approve at any time after the agent quits.

- (III): A can temporarily stop experimenting at any time. When A is not experimenting, A pays no flow cost and R can approve at any time.

If we restrict attention to Markov Perfect Equilibrium (MPE) using the belief Z_t as the state variable (as is standard in the literature), then we see that the equilibrium has a static threshold structure. Where the threshold are exactly depends on the fine details of the model.

Proposition 2. *Under set-ups (I) and (II), when $a = 1$ there exists an MPE characterized by a pair (B, b) such that R approves only at time $\tau_{\geq}(B)$ and A quits at time $\tau(b)$. In set-up (I), $B > 0$ while in set up (II), $B = 0$ and $b = b^*(0)$ and A quits experimenting when $X_t \notin (b, B)$. The value of experimentation in the MPE to R is strictly less than under one- or two-sided commitment.*

Set-up (I) corresponds to the model of Kolb (2016) and Henry and Ottaviani (2017) and the corresponding result follows directly from Kolb (2016). Note that in set-up (II), 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 MPE with only Z_t as a state variable is with loss. One natural justification for the restriction to MPE is that they are “simple” enough to be implementable in real world situations and minimize history dependence. There a large number of history dependent equilibria, many of which may sound implausible. Surprisingly, we show that a complex structure is not necessary for finding principal optimal equilibrium: if we only slightly expand the state space to be (X_t, M_t^X) , then the optimal mechanism under one-sided commitment can be implemented as an MPE without commitment.

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

The intuition behind the proof is quite simple. Suppose that A expects R to follow the mechanism as outlined in Theorem 1 and R expects A to continue experimenting until R approves or $X_t = b^*(B_A^*) \wedge b^*(-\frac{\phi}{\sigma}Z_0)$. Then R has no incentive to approve early (if approving early were a profitable deviation, then she could implement it in the mechanism with commitment and still satisfy all DP constraints) or reject early (since rejection is always suboptimal). Moreover, A has no incentive to cease experimenting early since his continuation value is always weakly positive. Moreover, if A ceases experimentation at

$X_t = B(M_t^X)$ until R approves, R also has no incentive to delay approval. The key to the proof is that neither player ever has a strictly negative continuation value and thus has no incentive to deviate from the prescribed equilibrium actions.

Proposition 3 tells us that our solution to the no-commitment case is the preferred equilibrium for R . We can also ask what other payoffs are generated by various equilibria. Our techniques from the case with commitment can be useful in generating the Pareto frontier of the equilibrium set. More formally, consider the problem of solving the principal's problem so that it respects DP constraints and delivers at least W utility to A :

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

subject to

$$DP : \sup_{\tau'} \mathbb{E}[e^{-r(\tau \wedge \tau')} (d_\tau \mathbf{1}(\tau \leq \tau') \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})] \leq \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})]$$

$$PK : \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})] = W$$

It is straightforward to show that the solution to the above problem will take the form $\tau = \inf\{t : X_t \geq B(M_t^X)\} \wedge \tau(b^*(B_A^*) \vee b^*(-\frac{\sigma}{\phi} Z_0))$ and $d_\tau = \mathbf{1}(X_\tau \geq B(M_\tau^X))$ where $B(M_t^X)$ is defined as

$$B(M_t^X) = \begin{cases} B & M_t^X \in [b^*(B), 0] \\ \underline{B}(M_t^X) & M_t^X \in [b^*(B_A^*) \vee b^*(-\frac{\sigma}{\phi} Z_0), b^*(B)]. \end{cases}$$

for some $B \in [B_A^* \vee -\frac{\sigma}{\phi} Z_0, B^1]$ (i.e., all B between the principal's first best B^1 and the max of the A -optimal threshold and R 's myopic threshold $-\frac{\sigma}{\phi} Z_0$). Thus, by adjusting the initial threshold B (which is equivalent to adjusting W), we can map out the Pareto frontier of the equilibrium set. Additionally, these mechanisms are Pareto optimal after every history, giving them a high degree of robustness to renegotiation.

5 Asymmetric Information

In many principal-agent situations the agent may have private information about the state. For example, if A is a start-up and R a venture capitalist deciding when to invest, it is likely that A is more informed than R about the start-up's profitability. In the

case of drug companies, the company may have acquired information about the drug during the R&D phase or during animal or foreign clinical trials which are not directly observable by the FDA. This information is valuable to R , as it will allow them to shorten experimentation time and make more informed decisions. In this section we study how to elicit this private information and look at what distortions it introduces. As we will see, under one-sided commitment the optimal mechanisms are qualitatively different when compared to the symmetric information case. When A reports a higher prior on θ , the optimal mechanism may entail giving him a *fast-track* to approval, where he is given a low initial approval threshold with the caveat that he may be thrown out of the fast-track if the outcomes of the trial go poorly and face a more stringent burden of proof to meet for approval.

Intuitively, R would like to set a lower approval threshold when A reports a higher prior. However, since A prefers a lower approval threshold than R , this introduces incentives for types with lower priors to misreport their type. In order to restore incentives, R will seek to “punish” outcomes which lower types find more likely. Due to the different beliefs of R when evaluating the mechanism for a high type and a low type A evaluating the same mechanism, R can “back-load” punishments in such a way as to reduce low type incentives while minimizing ex-ante distortions for R . In the optimal mechanism, this punishment comes in the form of being thrown out of the fast-track: R , knowing that A has a lower prior, views the chances that the fast-track is revoked (which entails an inefficient increase in the approval threshold from R ’s perspective) to be lower than A does when A has a lower prior and has deviated.

We model asymmetric information by allowing for the agent’s starting belief π_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 ℓ) 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 $\mathbb{P}(Z_i)$ be the ex-ante probability of type Z_i .

Formally, 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 θ , i.e., his the signal $s = \pi_A$. We will focus on the case when $\pi_A \in (0, 1)$, although the model can fit the case where $\pi_A \in \{0, 1\}$.⁸ 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 $m \in \mathcal{M} = \{h, \ell\}$

⁸In many applications, A is likely to be more but not perfectly informed-e.g., drug companies have more information but do not know for sure that their drug is good.

to R , after which R can publicly commit to a mechanism. Finally, at time $t = 0_+$ the public news process begins.

We now redefine a stopping mechanism to account for the need to elicit the private information of the agent. By the Revelation Principle, we focus on direct mechanisms in which A reports his type to R . This will result in R offering a *menu* of stopping mechanisms from which A can choose by reporting his type.

Definition 12. A *stopping mechanism* is a menu $\{(\tau^i, d_\tau^i)\}_{i=h,\ell}$ such that $(\tau^i, d_\tau^i) \in \mathbb{T} \times \mathbb{D}$ and R implements (τ^i, d_τ^i) when she receives message $m = i$.

With asymmetric information, R would like to approve h types quicker than ℓ types. However, incentive compatibility constraints will come into conflict with this goal: 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 her 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. This limit to the punishment R can deliver to ℓ will be a key determinant of the optimal mechanism under one-sided commitment.

When we are considering the effects of A misreporting his type, 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

We again begin by briefly studying the case with two-sided commitment. With the introduction of private information, we must also ensure that the mechanism R designs provides incentives for each type of A to correctly declare their type, which is given in our definition of incentive compatibility.

Definition 13. A *stopping mechanism under two-sided commitment* is *incentive compatible* if for all i, k ,

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

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

$$\begin{aligned} & \sup_{(\tau^i, d_\tau^i)_{i=\ell, h}} \sum_{i=\ell, h} \mathbb{E}[e^{-r\tau^i} d_\tau^i \frac{e^{Z_{\tau^i}} - 1}{1 + e^{Z_{\tau^i}}} | Z_i] \cdot \mathbb{P}(Z_i) \\ & \text{subject to } \forall i, k \\ & P(Z_i) : \quad \mathbb{E}[e^{-r\tau^i} (d_\tau^i \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_i] \geq \frac{c}{r} \\ & IC(Z_i, Z_k) : \quad \mathbb{E}[e^{-r\tau^k} (d_\tau^k \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_i] \leq \mathbb{E}[e^{-r\tau^i} (d_\tau^i \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_i]. \end{aligned}$$

The introduction of IC constraints mean that R , when designing the mechanism for Z_i , must consider both the distribution over outcomes given Z_i and the distribution over outcomes given Z_k . We might suspect that this difference in the distribution over outcomes may give R room to introduce non-stationary distortions⁹, surprisingly we find that it is in fact optimal to still use static threshold mechanisms.

Proposition 4. *The optimal mechanism under two-sided commitment is a menu of static-threshold stopping rules.*

The assumption of two-sided commitment is somewhat similar to the model of Guo (2016); in her bandit framework, the agent reports his private information about the state and the principal commits to a policy which describes which bandit arm to pull. The agent has no ability to quit the mechanism early (in her main framework, this is without loss because the agent prefers more experimentation than the principal). As in her model, we see that static-threshold rules are optimal.

With two-sided commitment, we find that the h type's $IC(Z_h, Z_\ell)$ constraint is often binding. Because R can make ℓ commit to experiment past the threshold at which ℓ 's expected continuation value becomes negative, R can punish ℓ 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 ℓ from claiming to be h .

Proposition 5. *If $b_h \neq -\infty$, then $IC(Z_h, Z_\ell)$ is binding.*

5.2 One-Sided Commitment

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

⁹See McClellan (2017) for an example where this is optimal.

dynamic version of incentive compatibility in a similar manner as we defined the dynamic participation constraints. Incentive constraints for type i 's value of reporting to be type k 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 can now define the proper notion of incentive compatibility.

Definition 14. *A stopping mechanism under one-sided commitment is **dynamically incentive compatible** if for all i, k ,*

$$\sup_{\tau' \in \mathbb{T}} \mathbb{E}[e^{-r(\tau^k \wedge \tau')} (d_\tau^k \mathbf{1}(\tau < \tau') \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_i] \leq \mathbb{E}[e^{-r\tau^i} (d_\tau^i \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_i].$$

We include the *sup* over $\tau' \in \mathbb{T}$ in the incentive constraint to convey the fact that i is comparing correctly declaring his type to be i to the payoff he could get from reporting to be type k and potentially quitting early. This introduces a much richer set of deviations each type could take than in the standard *IC* constraints (as in our problem with two-sided commitment). These types of double deviations (misreporting one's type and quitting early) will play an important role in determining the optimal mechanism.

We can then write the mechanism design problem with asymmetric information as

$$\begin{aligned} [AM] : \quad & \sup_{(\tau^i, d_\tau^i)_{i=\ell, h}} \sum_{i=\ell, h} \mathbb{E}[e^{-r\tau^i} d_\tau^i \frac{e^{Z_\tau^i} - 1}{1 + e^{Z_\tau^i}} | Z_i] \cdot \mathbb{P}(Z_i) \\ & \text{subject to} \quad \forall i = h, \ell \text{ and } k \neq i \\ & DP(Z_i) : \quad \sup_{\tau'} \mathbb{E}[e^{-r(\tau^i \wedge \tau')} (d_\tau^i \mathbf{1}(\tau^i \leq \tau') \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_i] \leq \mathbb{E}[e^{-r\tau^i} (d_\tau^i \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_i] \\ & DIC(Z_i, Z_k) : \quad \sup_{\tau'} \mathbb{E}[e^{-r(\tau^k \wedge \tau')} (d_\tau^j \mathbf{1}(\tau^k \leq \tau') \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_i] \leq \mathbb{E}[e^{-r\tau^i} (d_\tau^i \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_i]. \end{aligned}$$

We will proceed by analyzing the problem type-by-type. Unlike many mechanism design problems, there is no clear answer to which are the relevant constraints.¹⁰ In fact, different combinations of binding *DIC* constraints may bind depending on the specific values of Z_h, Z_ℓ . We begin by looking at what the optimal mechanism for h is when *DIC*(Z_ℓ, Z_h) is binding and *DIC*(Z_h, Z_ℓ) is slack.¹¹ This conjecture is in contrast with the case of two-sided commitment in which h 's *IC* was binding. However, these

¹⁰Other standard mechanism design features also fail to hold here; for example, the absence of transfers precludes standard “no distortion at the top” results.

¹¹Note that it is possible, if Z_ℓ is low enough, that ℓ would always prefer to quit immediately. However such a case would immediately revert back to the symmetric information model with initial belief $Z_0 = Z_h$.

two environments differ considerably. Under two-sided commitment, R could dissuade ℓ from imitating h (even when h 's approval threshold was lower) by decreasing h 's rejection threshold which, due the higher beliefs of h , would be more costly for ℓ than h . However, under one-sided commitment such a “threat” to ℓ is no longer credible since ℓ can always quit prematurely and take his outside option. Therefore, if the expected time for h to be approved is lower in one-sided commitment (as R would like), then it seems natural to think that ℓ 's DIC will bind. As we will show later, this intuition is correct if Z_h is high enough.

Let V_ℓ be the utility that ℓ gets from truthfully declaring his type (this is determined by R through his choice of ℓ 's mechanism, but for now we can treat it as fixed.). Then our problem of determining the optimal high type mechanism is given by

$$[AM^h]: \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_h]$$

subject to $DP(Z_h)$,

$$DIC(Z_\ell, Z_h, V_\ell) : \sup_{\tau'} \mathbb{E}[e^{-r(\tau \wedge \tau')} (d_\tau \mathbf{1}(\tau \leq \tau') \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_\ell] \leq \frac{c}{r} + V_\ell.$$

Intuitively, we should expect that h 's DP constraints will not be binding as long ℓ has not found it strictly optimal to quit experimenting: since h has a higher belief about the state being good, he will ascribe a higher probability to approval than ℓ would. The lower belief on ℓ 's part means that for low enough X_t , the symmetric information mechanism for h will induce ℓ to quit immediately. Let $(\tau_h^{SM}, d_{\tau, h}^{SM})$ be the stopping rule for the symmetric mechanism given h and define b^{SM} to be the highest X_t such that $(\tau_h^{SM}, d_{\tau, h}^{SM})$ starting at X_t would induce ℓ to quit immediately-i.e.,

$$b^{SM} := \max\{X_t : \max_X \mathbb{E}[e^{-r(\tau_h^{SM} \wedge \tau(X_i))} (d_{\tau, h}^{SM}(X) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | X_t, M_t^X = X_t, Z_\tau = Z_\ell + \frac{\phi}{\sigma} X_\tau] = \frac{c}{r}\}.$$

It is straightforward to see that such a b^{SM} exists; for example, if X_t is such that the optimal approval threshold for h is given by $\underline{B}_h(M_t^X)$, then ℓ will find it optimal to quit immediately since $\underline{B}_h(M_t^X) \geq \underline{B}_\ell(M_t^X)$. Upon $\tau(b^{SM})$, using the symmetric mechanism for h is optimal: the optimal mechanism for AM^h will promise ℓ and h types who have continued to this point some (weakly positive) continuation utilities and will respect all DP constraints for h . Since the symmetric mechanism for h doesn't have such promised continuation values and does respect all DP constraints, it's value will yield an upper bound on the original mechanism at time $\tau(b^{SM})$.

Therefore, we define a relaxed problem in which we drop all DP constraints and all but a finite number of DIC constraints and restrict the process to stop at $X_t = b^{SM}$ with continuation value to R equal to his payoff from the symmetric mechanism for h at X_t (which we will denote $SM^h(X_t)$). Formally, our relaxed problem can be written as

$$\begin{aligned}
[RAM_N^h] : & \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r(\tau \wedge \tau(b^{SM}))} (d_\tau(b^{SM}) \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} \mathbf{1}(\tau \leq \tau(b^{SM})) + \mathbf{1}(\tau > \tau(b^{SM})) SM^h(b^{SM})) | Z_h] \\
& \text{subject to } \forall X_i \in \mathcal{T}_N \cup \{b^{SM}\} \\
RDIC_\ell(X_i) : & \mathbb{E}[e^{-r(\tau \wedge \tau(X_i) \wedge \tau(b^{SM}))} (d_\tau(X_i) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} \mathbf{1}(\tau \leq \tau(b^{SM})) + \frac{c}{r}) | Z_\ell] \leq V_\ell + \frac{c}{r}.
\end{aligned}$$

We can decompose the first time ℓ would optimally quit and the continuation mechanism from this time onward.

$$\begin{aligned}
[H_N^h(X_t)] : & \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r(\tau \wedge \tau(b^{SM}))} (d_\tau(b^{SM}) \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} \mathbf{1}(\tau \leq \tau(b^{SM})) + \mathbf{1}(\tau > \tau(b^{SM})) SM^h(b^{SM})) | Z_t] \\
& \text{subject to } \forall X_i \in \{X_j \in \mathcal{T}_N \cup \{b^{SM}\} : X_j < X_t\} \\
RPK(0) : & \mathbb{E}[e^{-r(\tau \wedge \tau(X_i) \wedge \tau(b^{SM}))} (d_\tau(X_i) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} \mathbf{1}(\tau \leq \tau(b^{SM})) + \frac{c}{r}) | Z_t - \Delta_Z] \leq \frac{c}{r}.
\end{aligned}$$

We start by verifying that our relaxed problem RAM_N^h is truly a relaxed problem and yields a higher value to R than AM^h . This follows from our previous discussion on b^{SM} and the fact that we have dropped DP constraints and a number of DIC constraints.

Lemma 6. *The solution to RAM_N^h is an upper-bound on AM^h .*

With this in hand, we can begin the analysis of RAM_N^h . Using similar arguments to that of Lemma 2, we show that the optimal solution until the first binding constraint X^1 is to use a threshold rule. The proof is not a straightforward application of the arguments for the symmetric information case since the expectation in our constraint set are taken with respect to a different distribution (i.e., ℓ 's beliefs) than our objective function. Despite this difference, we show that threshold mechanisms until X^1 are still optimal.

Lemma 7. *The solution to RAM^h is given by a stationary approval threshold policy until the first binding constraint X^1 . At X^1 , the continuation mechanism solves $H_N^h(X^1)$.*

At the first binding constraint X^1 , R must be using a stopping rule which induces ℓ to weakly prefer to quit. There are many ways in which the optimal mechanism could induce

ℓ to quit while still providing incentives for h to experiment (e.g., there exists thresholds such that, due ℓ 's lower belief, ℓ would quit while h would continue to experiment). If the first binding constraint is b^{SM} , then the symmetric information mechanism will induce ℓ to quit.

Because A cares about the state, it is possible that R could dissuade ℓ from misreporting to be h by giving him too low of an approval threshold (i.e., threatening him with early approval). We are interested in a situation in ℓ prefers to have a lower threshold; In order to rule out such threat of early approval, we will assume that at b^{SM} , ℓ would still prefer a lower approval threshold. More specifically, we assume that there is only one threshold which would leave ℓ indifferent between continuing to experiment and quitting.

Assumption 1. For each $b > b^{SM}$ and $X > b$, there exists a unique B_Z such that $\tilde{V}(B_Z, Z_\ell + \frac{\phi}{\sigma}b, Z_\ell + \frac{\phi}{\sigma}X) = 0$.

If we consider payoffs such that A prefers approval regardless of the state (i.e., $a > 0$) then such an assumption clearly holds. Alternatively, we can also rule out such threats with a low approval threshold without this assumption on b^{SM} as long as A is allowed to continue experimenting (and delay approval or rejection until A desires to do so) after R approves (that is A 's payoff to approval is at X_t is $\max_{B,b} \tilde{V}(B, b, Z_\ell + \frac{\phi}{\sigma}X)$ rather than $\frac{e^{Z_\tau+a}}{1+e^{Z_\tau}}$).¹² In such a situation in which A controls both experimentation and approval once R has signed off on approval, A can always get at least his preferred threshold B_A^* , making threats of “too-early” approval by R non-credible and will ensure that A 's utility is weakly decreasing in the approval threshold and therefore there is a unique approval threshold which leaves the agent indifferent with taking his outside option.¹³

It is important to note that while the continuation mechanism at $\tau(X^1)$ must induce ℓ to quit, the payoff relevant beliefs for R are those of h . By inducing ℓ to quit, the mechanism may be setting a stricter approval policy that R would like to given that the true beliefs are h . If this is the case, then R would like to “relax” the overly stringent approval policy over time while making sure to do in such a way as to not violate the earlier incentives for ℓ to quit. We verify that this intuition is correct in the following lemma. R loosens the threshold by decreasing the threshold as M_t^X : as ℓ 's beliefs get lower, a lower approval threshold is needed to ensure that ℓ found it optimal to quit at X^1 . First, though, we define some notation:

¹²It is easy to show that such a payoff assumption won't change the results of Section 4.

¹³In the case of drug companies, this means that the regulator can not forbid the companies from running additional trails after approval but before putting the drug on the market, which is a reasonable assumption.

$$\begin{aligned}
b_i^*(B) &:= b^*(B; Z_i) \\
B_{i,A}^* &:= B_A^*(Z_i) \\
\underline{B}_i(X) &:= \min\{B > B_{i,A}^* : b_i^*(B) = X\}.
\end{aligned}$$

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

Lemma 8. *When $DIC(Z_h, Z_\ell)$ is slack, the optimal mechanism which solves $H_N^h(X_t)$ when the current evidence is X_t is given by a dynamic threshold policy $\tau = \inf\{t : X_t \geq B_h(M_t^X)\} \wedge \tau(b^*(B_{h,A}^*) \vee b_h^*(-\frac{\sigma}{\phi} Z_h))$ and $d_\tau = \mathbf{1}(\tau = \tau(B_h(M_t^X)))$ where*

$$B_h(M_t^X) = \begin{cases} \underline{B}_\ell(M_t^X) & M_t^X \in [b^{SM}, X^1], \\ B_h^{SM}(M_t^X) & M_t^X < b^{SM}. \end{cases}$$

where $B_h^{SM}(M_t^X)$ is symmetric information mechanism for h .

Combining Lemmas 7 and 8, we summarize the optimal mechanism for h below.

Lemma 9. *When $DIC(Z_h, Z_\ell)$ is slack, the optimal mechanism which solves h is given by, for some (b_h^1, B_h^1) , the policy $\tau = \inf\{t : X_t \geq B^h(M_t^X)\} \wedge \tau(\underline{b}(Z_h))$ and $d_\tau = \mathbf{1}(X_\tau = B^h(M_t^X))$ where*

$$B^h(M_t^X) = \begin{cases} B_h^1 & \text{if } M_t^X \in [b_h^1, 0) \\ \underline{B}_\ell(M_t^X) & M_t^X \in [b^{SM}, b_h^1) \\ B_h^S(M_t^X) & M_t^X < b^{SM}. \end{cases}$$

where $B_h^S(M_t^X)$ is symmetric information threshold function for h .

The approval threshold $B^h(M_t^X)$ is pinned down by (b_h^1, B_h^1) . Note that when $b_h^1 > b_\ell^*(B_h^1)$, then $\underline{B}_\ell(b_h^1) > B_h^1$, implying that the approval threshold takes a jump when M_t^X crosses b_h^1 . This features distinguishes it from our symmetric information mechanism and, as we will see, the mechanism for ℓ . The reason behind this jump follows from our discussion at the outset of the section 5.2: b_h^1 acts a failure threshold which, if reached, moves into a punishment phase in order to lower the incentives of ℓ to misreport his type.

This second stage of the optimal mechanism has a continuous and monotonically decreasing (in M_t^X) threshold, which consists of two regimes:

- **Punishment Regime:** After exiting the initial stationary regime, the mechanism may enter a “punishment” regime in which the approval threshold jumps up. The fact that $B^h(M_t^X) = \underline{B}_\ell(M_t^X)$ ensures that ℓ does indeed want to quit, but uses the minimal increase in the approval threshold necessary. R would like to decrease the approval threshold as quickly as possible, but must satisfy PK constraints which ensure that ℓ did indeed find it optimal to quit at τ' . We show that the optimal means to do this is by using the stopping time $\tau = \inf\{t : X_t \geq \underline{B}_\ell(M_t^X)\}$.
- **Symmetric Mechanism Regime:** At a certain point, employing the symmetric mechanism for type h is enough to deliver the appropriate punishments for ℓ and is optimally chosen by R at $M_t^X = b^{SM}$. In this sense, the distortions introduced by asymmetric information dissipate over time since conditional on no approval before $M_t^X = b^{SM}$, eventually R will be able to implement her best mechanism under one-sided commitment absent any distortions from private information.

Because R eventually moves to the symmetric mechanism for h , the rejection threshold is the same as in the symmetric information case, which implies that there is no distortion from private information at the end of experimentation. The $DIC(Z_\ell, Z_h)$ constraint does not cause R to reject earlier than would be optimal in the absence of incentive constraints. This comes about because the ℓ always has more pessimistic beliefs than h and thus there is always a way to deliver incentives for ℓ to quit that don't involve rejection. Interestingly, this implies that the placement of distortions are non-monotonic in M_t^X for h : they increase upon entering the punishment phase but then are gradually reduced over time until they are eliminated upon entering the second stationary regime, in which R uses the best mechanism he could get absent private information.

Having found what the optimal mechanism is for h , we must also solve for the optimal mechanism for ℓ . Suppose that the mechanism must deliver utility V_ℓ to the low type correctly declaring his type. Then the mechanism design problem is given by

$$[AM_\ell] : \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_\ell]$$

subject to

$$DP(Z_\ell) : \sup_{\tau' \in \mathbb{T}} \mathbb{E}[e^{-r(\tau \wedge \tau')} (d_\tau \mathbb{1}(\tau' > \tau) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_\ell] \leq \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_\ell]$$

$$PK(V_\ell) : \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_\ell] \geq V_\ell.$$

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 ℓ is qualitatively the same as the symmetric information mechanism, which turns out to be correct.

Lemma 10. *The optimal mechanism for ℓ when $DIC(Z_h, Z_\ell)$ is slack is given by the stopping rule $\tau = \inf\{t : X_t \geq B_\ell(M_t^X)\} \wedge \tau(b_r)$ and $d_\tau = \mathbb{1}(X_\tau = B_\ell(M_\tau^X))$ which is defined as*

$$B_\ell(M_t^X) = \begin{cases} B_\ell^1, & M_t^X \in [b_\ell^*(B_\ell^1), 0), \\ \underline{B}_\ell(M_t^X), & M_t^X \in [b_r, b_\ell^*(B_\ell^1)). \end{cases}$$

where B_ℓ^1 is such that B_ℓ^1 is less than it would be in the symmetric information case.

Lemma 10 leaves open the possibility that R rejects at $b_r > b_\ell^*(B_\ell^1)$ -i.e., the second part of $B_\ell(M_t^X)$ is never reached. Whenever A and R 's payoffs are aligned over terminal payoffs (i.e., $a = -1$), then R will never reject early and $b_r = b_\ell^*(B_{A,\ell}^*)$. However, this may not be true when $a = 1$. Although there are no DIC constraints to consider, R , when designing ℓ 's mechanism, does consider how increasing the utility of ℓ when truthfully reporting weakens the incentives for ℓ to misreport himself to be h . Therefore, R will decrease the approval threshold lower than she would prefer; if this decrease is large enough, R may prefer to not enter the incentivization regime and instead reject at b_r . However, it is straightforward to show that if $\mathbb{P}(Z_\ell)$ is large enough and $Z_\ell > b_\ell^*(\frac{-\sigma}{\phi} Z_\ell)$, then we will have $b_\ell^1 = b^*(B_\ell^1)$. This will also always be the case when $Z_\ell > 0$.

Having derived the form of each mechanism for each type separately, we can now formally state the optimal mechanism.

Theorem 2. *When $DIC(Z_\ell, Z_h)$ is binding and $DIC(Z_h, Z_\ell)$ is slack, the optimal mechanism is given the mechanisms of Lemmas 9 and 10. If the optimal mechanism for h is not equal to the symmetric information case and $b_h^1 = b^*(B_\ell^1)$, then $B_h^1 \leq B_\ell^1$ and $b_\ell^1 \leq b_h^1$; moreover, $B_h^1 < B_\ell^1$ implies $b_\ell^1 < b_h^1$*

The mechanism displays a number of interesting characteristics:

- **Low Type Monotonicity:** The mechanism for ℓ closely resembles that of the symmetric mechanism in that the approval threshold will only drift downwards.
- **High Type Jump:** When $b_h^1 > b_\ell^*(B_h^1)$ (which we verify later will be the case when Z_h is high enough), then the approval threshold for h takes a jump upwards

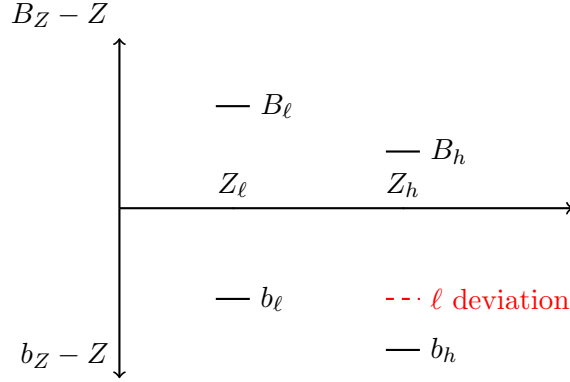


Figure 5: If the initial regime thresholds were such that $B_\ell^1 > B_h^1$ and $b_\ell^1 > b_h^1$, then ℓ report himself to be h and quit early, effectively giving himself a lower approval threshold.

when X_t reaches b_h^1 for the first time, after which it is monotonically decreasing in M_t^X . This jump upward occurs in order to provide enough punishment for a deviating ℓ type to prefer to quit immediately.

- **Some Distortion At The Start:** Until the first time $X_t = b^{SM}$, the approval threshold will differ than R 's optimal mechanism under symmetric information.
- **No Distortion At The End:** R never rejects at a belief higher than she would under symmetric information for h and for ℓ as well if $b_r < b_\ell^*(B_\ell^1)$.

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

We refer to the mechanism given to h , when not equal to his symmetric information mechanism, 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 b_h^1 . If the failure threshold is reached first, then the trial is declared a failure and the agent is thrown out of the fast-track. However, instead of rejecting, R allows h to immediately begin experimenting again, only now h is given a higher approval threshold.

This fast-track mechanism illustrates the trade offs that must be made under one-sided commitment: in order to grant h a lower approval threshold, R must deter deviations by ℓ by increasing the failure threshold. This lower approval threshold is more

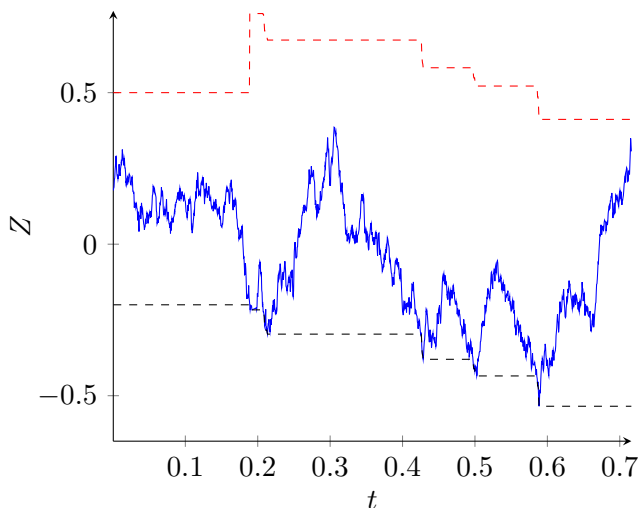


Figure 6: The approval threshold for h (the upper dashed line) starts off low, but, when X_t crosses the failure threshold (the lower dotted line), the approval threshold jumps up and drifts down as M_t^X decreases.

likely to be reached by h than ℓ , which allows R to profitably back load distortions in the “failure” threshold.

We now illustrate some of the ideas behind why the optimal mechanism must take this nested form. For ℓ , when declaring himself to be ℓ or h , his utility is completely determined by the initial static thresholds. Suppose that $V_\ell > 0$; then we know that $b_1^\ell = b_\ell^*(B_1^\ell)$ -i.e., R keeps the initial threshold fixed until the point at which ℓ is first indifferent between ceasing and continuing experimentation. We note 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 ℓ), ℓ must gain from experimenting longer on the low end of beliefs when claiming to be h . But since ℓ 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 ℓ from claiming to be h . If $B_1^h < B_1^\ell$, then it must be that $b_h < b_\ell$. Otherwise ℓ could profitably deviate by claiming to be h and quitting when beliefs drift down from initial beliefs by b_ℓ . In this way, ℓ is able to maintain the same quitting threshold as truthfully declaring himself to be ℓ 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 ℓ . This incentivization regime decreases the

approval threshold enough to keep ℓ indifferent. Since h has a higher belief than ℓ after observing X_t , h (when reporting to be ℓ) will still have positive continuation value when in ℓ 's incentivization regime, creating incentives for h to imitate ℓ . 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 6. *For each Z_ℓ , $\exists \bar{Z}$ 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. Moreover, if ℓ 's utility is strictly positive, then $b_h^1 > b^*(B_h^1)$.*

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. As we will show, under the assumption that $Z_\ell < 0$ ¹⁴, 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 ℓ 's mechanism by inducing early rejection.

Lemma 11. *The optimal mechanism for ℓ 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*

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

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

When both DIC constraints are binding, the problem to determine h 's mechanism is only modified by 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 and 8.

Lemma 12. *When $DIC(Z_h, Z_\ell)$ is binding, there exists (b_h^1, B_h^1, B_h^2) such that the optimal mechanism for h is given by $\tau = \inf\{t : X_t \geq B^h(M_t^X)\} \wedge \tau(-\frac{\sigma}{\phi} Z_h)$*

$$B^h(M_t^X) = \begin{cases} B_h^1 & \text{if } M_t^X \geq b_h^1, \\ \underline{B}_\ell(M_t^X) & \text{if } M_t^X \in [b_\ell^*(B_h^2), b_h^1], \\ B_h^2 & \text{if } M_t^X \in [b_h^*(B_h^2), b_\ell^*(B_h^2)), \\ \underline{B}_h(M_t^X) & \text{if } M_t^X \leq b_h^*(B_h^2), \end{cases}$$

and $d_\tau = \mathbb{1}(X_\tau = B^h(M_t^X))$.

¹⁴The assumption that $Z_\ell < 0$ is reasonable given our application: Over 90% of all drugs that begin a clinical trial fail to be approved (see FDA (2017)).

In all, we summarize the optimal mechanism in this case below:

Theorem 3. *The optimal mechanism for ℓ when $Z_\ell < 0$ and $DIC(Z_h, Z_\ell)$ is binding is given by Lemma 11 and the optimal mechanism for h is given by Lemma 12.*

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 ℓ early (i.e.- $b_r > \underline{b}(0)$) in order to lower incentives for h to imitate ℓ . 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 ℓ 's incentive to quit).

Interestingly, unlike h 's mechanism when $DIC(Z_\ell, Z_h)$ is binding, ℓ 's mechanism does not qualitatively change much even when $DIC(Z_h, Z_\ell)$ is binding. The fast-track feature of h 's mechanism comes from the backloading of punishments by R , whereas the distortions in ℓ 's mechanism from $DIC(Z_h, Z_\ell)$ come through in early rejection of the project. This tells us that the different priors in the constraint set matter and it is important whether they are greater or less than the prior of the objective function.

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_h^1, b_h^1) and (B_ℓ^1, b_r) . When $DIC(Z_h, Z_\ell)$ is binding, we have consider three parameters each for h, ℓ : (B_h^1, b_h^1, B_h^2) and $(B_\ell^1, b_\ell^1, b_r)$. 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 $j_i(X, M, b_r)$ (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 $Z = Z_i + \frac{\phi}{\sigma}X$ 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

$$\begin{aligned}
j_i(X, M^X) = & \underbrace{\Psi(B_{i,Z}(M^Z), M^X, Z) \frac{e^Z - e^{Z-B_{i,Z}(M^Z)}}{1 + e^Z}}_{\text{Expected utility from approval before } Z_t = M^Z} \\
& + \underbrace{\psi(B_{i,Z}(M^Z), M, Z) \frac{e^Z + e^{Z-M^Z}}{1 + e^Z}}_{\text{Discounted prob. } M^Z \text{ hit first}} \cdot \underbrace{j_i(M^X, M^X)}_{\text{Continuation value at } M^Z}.
\end{aligned}$$

where the mapping from X, M^X, B to X, M^Z, B_Z is understood. Thus if we can calculate $j_i(M^X) := j_i(M^X, M^X)$, the value of $j_i(X, M^X)$ follows immediately. In order to calculate $j_i(M^X)$, we use the principle of normal reflection¹⁵: $\frac{\partial j_i(X, M^X)}{\partial M^X} \Big|_{X=M^X} = 0$. We can then take the derivative with respect to M^X to get

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

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

$$j'_i(M^X) = j_i(M^X) \left[\frac{1}{1 + e^{M^X}} - \psi_b \right] - \frac{e^{M^X} - e^{M^X - B_{i,Z}(M^Z)}}{1 + e^M} \Psi_b, \quad (1)$$

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

Proposition 7. *The value of experimentation to R for type i when the current evidence level is M_t^X and the minimum is M_t^X is given the unique solution to $j'_i(M_t^X)$.*

For ℓ , we know that the approval mechanism is strictly decreasing in M and so this equation gives the value of the incentivization for ℓ . The function j_h also corresponds to h 's optimal mechanism when h 's continuation value at $X_t = M_t^X$ is equal to zero. We still need to derive the value of the mechanism to R when in the punishment regime.

¹⁵See Peskir and Sharyaev (2006) for a derivation.

By a similar derivation, we can show that value in the punishment regime (which we call j_h^p) is given the same differential equation as in equation 1, but now with boundary condition at b_h^2 of

$$j_h^p(b_h^1) = \Psi(B_{h,Z}^2, b_{h,Z}^2, M^Z) \frac{e^{M^Z} - e^{M^Z - B_{h,Z}^2}}{1 + e^{M^Z}} + \psi(B_{h,Z}^2, b_{h,Z}^2, M^Z) \frac{e^{M^Z} - e^{M^Z - b_{h,Z}^2}}{1 + e^{M^Z}} j_2(b_{h,Z}^2).$$

where $B_{h,Z}^2$ and $b_{h,Z}^2$ are the translations of B_h^2 and b_h^2 into belief space. When $DIC(Z_h, Z_\ell)$ is binding, we need to also consider the value to h of ℓ 's optimal mechanism when in the incentivization regime, which we will call v_h^ℓ . By a similar argument as for $j(X, M^X)$, we get a differential equation for $v_h^\ell(M^X)$ as

$$\begin{aligned} \frac{dv_h^\ell(M^X)}{dM^X} &= \frac{\phi}{\sigma} \left([1 - \psi_B B'_{h,Z}(M^Z) - \psi_b] \left(\frac{c}{r} + v(M^Z) \right) \right. \\ &\quad \left. - [\Psi_B B'(M) + \Psi_b] \frac{e^{M^Z + \Delta_z} (1 + \frac{c}{r}) + (a + \frac{c}{r}) e^{M - B_{h,Z}(M)}}{1 + e^{M^Z + \Delta_z}} \right) \end{aligned}$$

with boundary condition $v_h^\ell(b_r) = 0$.

Similarly, we must find the value to h from the beginning of punishment regime in the mechanism for h , which we denote v_h^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 $v_h^h(M) = \Psi(B_h^2, b_h^2, M) \frac{e^{M(1+\frac{c}{r})+(a+\frac{c}{r})} e^{M-B_h^2}}{1+e^M} + \psi(B_h^2, b_h^2, M) \frac{c}{r} \frac{e^{M+e^{M-b_h^2}}}{1+e^M}$.

This allow us to write out the mechanism design problem as

$$[QD]: \quad \max_{(B_i, b_i, \underline{b}_i)_{i=h,\ell}} \sum_{i=h,\ell} \Psi(B_{Z,i}, b_{i,Z}, Z_i) \frac{e^{Z_i} - e^{Z_i - B_{i,Z}}}{1 + e^{Z_i}} + \psi(B_{i,Z}, b_{i,Z}, Z_i) \frac{e^{Z_i} + e^{Z_i - b_{i,Z}}}{1 + e^{Z_i}} j_i(b_i, \underline{b}_i)$$

subject to $\forall i = h, \ell$ and $k \neq i$

$$\begin{aligned} DIC(Z_i, Z_k) &: \Psi(B_{i,Z}, b_{i,Z}, Z_i) \frac{e^{Z_i} (1 + \frac{c}{r}) + (a + \frac{c}{r}) e^{Z_i - B_i}}{1 + e^{Z_i}} + \psi(B_{Z,i}, b_i, Z_i) \frac{e^{Z_i} + e^{Z_i - b_i}}{1 + e^{Z_i}} v_i^i(b_i, \underline{b}_i) \\ &\geq \Psi(B_{k,Z}, b_{k,Z}, Z_k) \frac{e^{Z_i} (1 + \frac{c}{r}) + (a + \frac{c}{r}) e^{Z_k - B_{k,Z}}}{1 + e^{Z_i}} + \psi(B_{k,Z}, b_{k,Z}, Z_k) \frac{e^{Z_i} + e^{Z_k - b_{k,Z}}}{1 + e^{Z_i}} v_i^k(b_k, \underline{b}_k) \end{aligned}$$

$$DP(Z_i) : b_i \geq b_i^*(B_i).$$

where v_h^k is defined as above, v_ℓ^k .

5.4 Comparative Statics

While all of the previous section allows for general a , we now restrict attention to the case when $a = 1$ so that A always prefers immediate approval. In this case, the misalignment of R and A 's preferences is particularly severe, making it more difficult for R to elicit A 's private information. As we will see, this difference in preferences opens up a number of interesting comparative statics.

In the symmetric information case, increasing the cost c unambiguously hurts R , since it induces R to provide more incentivization and reject at a higher beliefs (i.e., $\frac{\partial b^*(-\frac{\phi}{\sigma}Z_0)}{\partial c} > 0$). However, with asymmetric information this is no longer the case. Additional costs may be of use as a screening device. When c becomes small, it becomes increasingly harder for R to induce ℓ to quit while still inducing h to keep experimenting. Taking the limit as $c \rightarrow 0$, we get that the private information of A is not used at all.

Proposition 8. *Under both one- and two-sided commitment, as $c \rightarrow 0$ the optimal mechanisms for h, ℓ converge to value of the single-decision maker problem for R with prior $\mathbb{P}(Z_h)\pi_h + (1 - \mathbb{P}(Z_H))\pi_\ell$.*

With asymmetric information and the absence of monetary transfers, costly experimentation provides a tool for screening of types, as detailed in the following proposition. This result can speak to the debate on who should fund drug trials (drug companies or government agencies), providing a reason for requiring the companies by requiring them to have some “skin in the game” and making it easier to elicit any private information the companies may have.

Proposition 9. *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 .*

To understand the idea behind this proposition, we, consider the limiting cases of c and suppose that $\pi_h \approx 1$ and $\pi_\ell \approx 0$ and $\mathbb{P}(Z_\ell)$ is large. As $c \rightarrow 0$, the value of the optimal mechanism converges to that of the principal-optimal symmetric information problem with prior $\mathbb{P}(Z_h)\pi_h + \mathbb{P}(Z_\ell)\pi_\ell$: 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 ℓ . R could find a mechanism which rejects ℓ 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 θ . 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 θ . The following proposition shows that asymmetric information is in fact better for R .

Proposition 10. *Let π_0 be the prior of R . Then the value to R of optimal mechanism under asymmetric information in which A learns θ perfectly is higher than the value to R of the optimal mechanism under symmetric information*

6 FDA Drug Approval

Our model provides a theoretical justification for why approval standards may change over time. A related question, outside of the normative viewpoint adopted in the previous sections, is whether or not we see such changes in practice. To consider this question, we look more closely at FDA approval standards. We will study the relationship between the length of clinical trials and the probability of Type I error and will show that, as in our model, longer experimentation leads to more Type I error. First though, we describe the process which a new drug must go through to get approved.

Once a new drug has been patented, the drug company must apply for a Investigational New Drug (IND) application with the FDA before beginning trials on human subjects. Companies can, however, perform tests on animals and overseas without the FDA's approval, which can motivate why drug companies hold private information about the drug. Once the IND has been approved by the FDA, the companies can begin human-subject clinical trials. The trials are typically conducted in three phases and cumulatively can last for many years (5.76 years on average in our data, but can take up to 20 years). Once the company has completed their trials, they submit a New Drug Application (NDA) with the FDA, at which point the FDA reviews the results of the clinical trials. After the FDA reviews the companies application, it either approves or fails to approve the drug. However, over 90% of all NDA applications are approved, indicating that companies understand and have met the standards needed for approval.

In addition to the normal approval process, the FDA has a number of programs for expedited approval. Beginning in 1997, the FDA began using a process for approvals, which it calls the Fast Track designation whose "purpose is to get important new drugs to the patient earlier." Prior to the IND application, the sponsoring company can apply

for the Fast Track designation. After consultation with the company, the FDA will grant the designation if the drug “is intended to treat a serious condition AND nonclinical or clinical data demonstrate the potential to address unmet medical need.” However, the fast-track designation is not always permanent: if the data from the clinical trial shows that the drug no longer meets the criteria for the Fast Track designation, it may be removed after which the company can continue to perform clinical trials, albeit without the benefits of the Fast Track. The FDA expanded its expedited approval programs by creating the Accelerated Approval and Breakthrough Therapy designation, both of which have similar features as the Fast Track designation. These attributes, expedited approval and removal from the Fast Track upon negative news from the clinical trail, fit the qualitative features of Section 5.2.

In contrast to the standard single-decision maker problem, our model predicts that the FDA should use a dynamically moving approval threshold. Unfortunately, we cannot directly observe belief at which the FDA is approving the drugs. The standard model, with a single constant approval threshold, should have the same probability of Type I error regardless of the length of the clinical trial. This will allow us to test the asymmetric information model. For ℓ types (those who do not receive the Fast Track), the approval threshold is decreasing and hence the same increasing probability of Type I error as in Section 4.3 holds. At first glance, the problem for h types is more complicated, since the approval threshold is non-monotonic. However, we can note that as long as h is in the fast track (i.e., the failure threshold has not been reached), the approval threshold is constant over time. Therefore, our model predicts that drugs in the fast track should have a constant Type I error probability over time while drugs in the standard approval process should have an increasing probability of Type I error.

Our data below consists of 370 drugs approved between 1987-2013, of which 65 were either withdrawn or received a BBW. 79 of the drugs were approved under an expedited approval process (Fast Track or Accelerated Approval). Much of the data was compiled from Gilchrist (2016), Frank et al. (2014) and Carpenter et al. (2008) while the rest of the data was taken by searching the FDA website directly¹⁶. Table 1 gives some descriptive statistics of the data and Figure 7 gives a histogram over the length of testing prior to NDA submission.¹⁷ Table 1 shows that, as intended, the approval time we can confirm that the length of clinical trials under Expedited Approval is lower then that under the Standard Approval pathway; performing a two-sample t-test, we can confirm that this

¹⁶ From Gilchrist (2016), we get a measure of how long each drug took from the initial IND application to NDA submission. We then use Frank et al. (2014) and Carpenter et al. (2008) to compile a measure of Type I error.

¹⁷All results presented below are robust to dropping outliers.

Table 1: Summary Statistics

	Standard Approval	Expedited Approval
Mean	6.135	5.143
Median	5.317	4.890
Maximum	26.21	14.997
Minimum	1.011	0.076
Standard Deviation	3.348	2.684
N	339	79

difference is statistically significant ($p = 0.0145$).

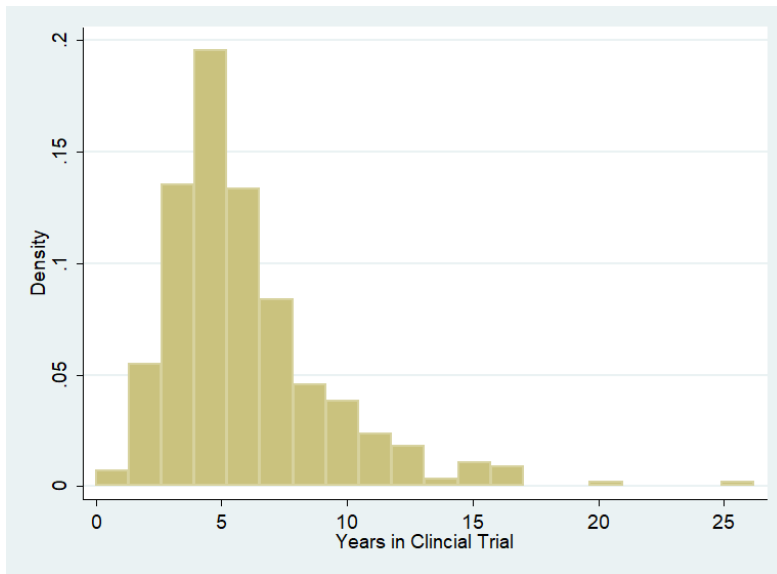


Figure 7: This histogram shows the length of clinical trials across our sample.

There are several ways in which we could measure Type I error. The most obvious is to count whenever a drug is withdrawn from the market for safety concerns. However, some drugs are not pulled from the market but receive Black Box Warnings (BBW) which appear on the label of the drug and advise of negative side-effects and changes in the prescribed application of the drug. These events signify a failure of the approval process to correctly determine the safety of a new drug. Following Frank et al. (2014)

and Carpenter et al. (2008), we say that a drug experiences Type I error if it is either withdrawn from the market or receives a post-market entry BBW.

When looking at the effect of trial length on Type I error, it is important to control for the type of drug: it may be that drugs treating cancer are judged by a different standard than those treating allergies. In order to control for this, we classify each drug according to its Anatomical Therapeutic Chemical (ATC) code, which is used by the World Health Organization to categorize drugs by which organ and system they treat as well as their pharmacological and chemical properties. We used this to group drugs in to 14 categories, which we then use as fixed effects to control for the drug class in our regressions.

In addition to drug category fixed effects, we can also control for a number of other factors in the regression. While our interest is in the effect of trial length on Type I error, we might wonder whether the length of the FDA review process after NDA submission has an impact on the probability of Type I error. Ex-ante, it is reasonable to think that the length of the review is related to the length of the trial and the probability of Type I error. Additionally, we also control for the size of the company (including a dummy variable for whether or not the sponsoring company was one of the 25 largest) as well as the approval year (to pick up any time trends in the data). Running the regression, some of our observations are lost due to lack of variability in the error indicator within drug class (for some classes there were no Type I errors). The results are presented in [Table 2](#).

Table 2: Regression Results

	SP	EP
Trial Length	0.104*** (0.043)	0.0001 (0.176)
Review Time	0.074 (0.080)	-1.448 (1.311)
Top 25 Company	0.440 (0.324)	-0.195 (0.876)
Approval Year	-0.055** (0.024)	-0.030 0.079
Drug Class Fixed Effects	Yes	Yes
N	309	61

Standard deviation in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The results exactly match what our model predicts: we see that under the standard pathway, drugs that took longer in clinical trials were more likely to experience a Type I error while those clinical trial length had no impact in the expedited approval programs. Our model provides a justification for why such a relationship may actually be optimal.

7 Extensions

In the previous sections we have assumed a limited set of controls for R (i.e, to approve or not) and very specific utility functions/diffusion process. We now explore how each of features shapes our results and show that the qualitative form of optimal policies is robust to many of our model’s details.

7.1 Transfers

One natural extension is to allow for some form of transfers. In many settings, unbounded transfers are infeasible. Instead, we will focus on the case where R may share the costs

of experimenting with A . In our FDA example, this can be done via subsidies or grants. Formally, we let w_t be the transfers made at time t and $u(w_t)$ be the agent's utility from transfers w_t . The utility of R is given by

$$\mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} - \int_0^\tau e^{-rt} w_t | Z_0],$$

and the utility to A is given by

$$\mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - a}{1 + e^{Z_\tau}} - \int_0^\tau e^{-rt} (\frac{c}{r} + u(w_t)) | Z_0].$$

If we extend our Lagrangian analysis to this setup, we can see that the choice of w will depend on the accumulated Lagrange multipliers. Because the choice of w_t does not affect the evolution of X_t or Z_t , w_t will be determined at each time t by the first-order condition

$$1 = - \sum_{i=1}^N \lambda(X_i) \mathbb{1}(M_t^X \leq X_i) u'(w_t).$$

Interestingly, the amount of the costs borne by R is, conditional on M_t^X , independent of the current level of X_t . Since $-\sum_{i=1}^N \lambda(X_i) \mathbb{1}(M_t^X \leq X_i)$ is increasing in M_t^X , this will imply that the optimal α is also increasing as M_t^X decreases. This seems natural; as the the firm becomes more pessimistic, R is willing to take on more costs in order to prolong experimentation and to reduce the need to decrease the approval threshold.

While it might seem strange at first that R takes on more costs as the beliefs about the project gets worse, this is exactly what happens in many real-world settings. If we think about for which projects the government provides subsidies, they are often for research which has less of a chance of being successful but the government values more than the agent controlling the project.

Proposition 11. *The optimal mechanism with cost sharing depends only on M_t^X and $w_t = \tilde{w}(M_t^X)$ for some decreasing function \tilde{w} .*

7.2 General Markov Process

It is natural to wonder how the results of the model depend on the particular framework used. For example, how does the payoff structure determine the optimal mechanism? Does the exact specification of the diffusion process qualitatively determine the optimal strategy? To answer these questions, we generalize the symmetric information model to

allow for a wide range of utility functions and stochastic processes. We can show that the optimal mechanism retains the same form as in Section 4.3.

Let X_t be a one-dimensional diffusion process on $\mathcal{I} \subseteq \mathbb{R}$ (where \mathcal{I} is the interval $[\underline{a}, \bar{a}]$ and \underline{a}, \bar{a} are possibly $-\infty, \infty$ respectively) which solves the stochastic differential equation

$$dX_t = \mu(X_t)dt + \sigma(X_t)dB_t, \quad (2)$$

for some Borel functions $\mu : \mathcal{I} \rightarrow \mathbb{R}$, $\sigma : \mathcal{I} \rightarrow \mathbb{R}$ and a given X_0 . We assume μ, σ are such that equation 2 has a unique (weak) solution. For an arbitrary mechanism (τ, d_τ) , let R 's utility be given by

$$\mathbb{E}[e^{-r\tau}g(X_\tau, d_\tau)|X_0],$$

and A 's utility be given by

$$\mathbb{E}[e^{-r\tau}f(X_\tau, d_\tau)|X_0].$$

As before, we want to explore the dynamics of the optimal mechanism when A cannot sign binding long-term contracts (one-sided commitment). If A chooses to quit at X_t , he receives his outside option, $f(X_t, 0)$, which is equal to his payoff from rejection and may depend on X_t . The mechanism design problem for R can then be written as

$$[GSM] : \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau}g(X_\tau, d_\tau)|X_0]$$

subject to

$$DP : \sup_{\tau'} \mathbb{E}[e^{-r(\tau \wedge \tau')}f(X_{\tau \wedge \tau'}, d_\tau \mathbb{1}(\tau \leq \tau'))|X_0] \leq \mathbb{E}[e^{-r\tau}f(X_\tau, d_\tau)|X_0]$$

In order to use the techniques as sketched in Section 4.3, we need to place several assumptions on the f, g .

Assumption 2. *We assume that for $w \in \{f, g\}$, the following are satisfied:*

- 1. *Pure delay is sub-optimal: For $\alpha \in \{0, 1\}$, we have*

$$\mathbb{E}[e^{-r(\tau \geq (B) \wedge \tau(b))}w(X_{\tau \geq (B) \wedge \tau(b)}, \alpha)|X_t] \leq w(X_t, \alpha).$$

- 2. *Decision Threshold: If $X' > X$, then $w(X, 1) \geq w(X, 0)$ implies $w(X', 1) > w(X', 0)$.*

- 3. *Continuity:* $w(X, \alpha)$ is continuous in X for each $\alpha \in \{0, 1\}$.
- 4. *There exists \underline{X} such that the solution for any $X_0 < \underline{X}$ to*

$$\sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} g(X_\tau, d_\tau) | X_0]$$

subject to

$$\mathbb{E}[e^{-r\tau} f(X_\tau, d_\tau) | X_0] \geq f(X_0, 0)$$

is to stop and reject immediately.

- 5. *A prefers approval whenever R does:* $g(X, 1) > g(X, 0) \Rightarrow f(X, 1) > f(X, 0)$.
- 6. $\lim_{x \rightarrow \infty} \mathbb{E}[e^{-r\tau(x)} | y] w(x, 1) = 0$ for any $y \in \mathcal{I}$.

While the list of assumptions may seem long, each part of the assumption is generally very mild and will fit many other models besides the one we have analyzed so far. Parts 1 – 3 are straightforward and are satisfied in most stopping problems considered in the literature. Part 2 implies that there is a myopic cutoff point for R , above which R approves and below which R rejects, which we denote X_{my}^R -i.e., $g(X_{my}, 1) = g(X_{my}, 0)$ (and define X_{my}^A similarly). Part 4 ensures that there is a lower bound below it is too costly for R to incentivize A to continue. Part 5 ensures that R has an incentive to stop at some point if X_t goes too low and Part 6 is merely a technical condition needed to ensure payoffs do not diverge.

Next, we need to place some assumptions on the preferences of the players in relation to their optimal approval thresholds. As before, we will look at situations in which A prefers “less” experimentation (i.e., a lower approval threshold) than R does. Define A ’s utility to static threshold mechanism which stops whenever $X_t \notin [b, B]$ and takes decision d^a at threshold a as

$$\tilde{V}(B, d^B, b, d^b, X) = \mathbb{E}[e^{-r(\tau_{\geq(B)} \wedge \tau(b))} f(X_{\tau_{\geq(B)} \wedge \tau(b)}, d_{\tau_{\geq(B)} \wedge \tau(b)}) | X],$$

where $d_{\tau_{\geq(B)} \wedge \tau(b)} = d^B \mathbf{1}(\tau_{\geq(B)} < \tau(b)) + d^b \mathbf{1}(\tau(b) < \tau_{\geq(B)})$. We define $\hat{V}(B, b, X) := \tilde{V}(B, 1, b, 0, X)$ and similarly for $\hat{J}(B, b, X)$.

Let us suppose that the optimal mechanism were restricted to reject at $\tau(b_r)$. We can then ask what the preferred static approval threshold would be for R and A . As in Section 4, we will want to impose that R prefers a higher approval threshold than A does. We define $B_R^*(b_r), B_A^*(b_r)$ to be functions which give us the optimal approval thresholds of R and A respectively:

$$B_R^*(b_r) := \operatorname{argmax}_B \mathbb{E}[e^{-r(\tau_{\geq(B)} \wedge \tau_{\leq(b_r)})} g(X_{\tau_{\geq(B)} \wedge \tau_{\leq(b_r)}}), \mathbf{1}(\tau(B) < \tau(b_r)))]$$

$$B_A^*(b_r) := \operatorname{argmax}_B \mathbb{E}[e^{-r(\tau_{\geq(B)} \wedge \tau_{\leq(b_r)})} f(X_{\tau_{\geq(B)} \wedge \tau_{\leq(b_r)}}), \mathbf{1}(\tau(B) < \tau(b_r)))]$$

This allows us to formally state our assumption that R prefers a higher threshold. We assume that the utility with respect to a static approval threshold is strictly single-peaked and that R 's preferred approval threshold is greater than that of A .

Assumption 3. *The following assumptions on the utility of R and A hold:*

- $\widehat{V}(B, b, X), \widehat{J}(B, b, X)$ are both strictly single-peaked in B and continuous in all arguments.
- Given rejection at b_r , R 's preferred approval threshold is higher than that of A : $B_R^*(b_r) > B_A^*(b_r)$.

We can then write down functions analogous to b^*, \underline{B} as defined previously:

$$b^*(B) := \operatorname{argmax}_b \widehat{V}(B, b, B - \epsilon)$$

$$\widetilde{B}(b, X) := \{B > \max_b B_A^*(b) : \widehat{V}(B, b, X) = f(X, 0)\}$$

$$\underline{B}(X) := \lim_{\delta \rightarrow 0} \widetilde{B}(X + \delta, X).$$

We now make our final assumption, a condition on \widetilde{B} , which will be used to ensure that \underline{B} exists.

Assumption 4. $\widetilde{B}(b, X)$ is continuously differentiable.

With this in hand, we can show that state a generalization of Theorem 1 for general payoff functions and stochastic processes.

Theorem 4. *Under Assumptions 2-4, the solution to GSM is given by $\tau = \inf\{t : X_t \geq B(M_t^X)\} \wedge \tau(b_r)$ (for some b_r) and $d_\tau = \mathbf{1}(X_t \geq B(M_t^X))$ where the approval threshold is given by*

$$B(M_t^X) = \begin{cases} B^1 & \text{if } M_t^X \geq b^*(B^1) \\ \underline{B}(M_t^X) & \text{if } M_t^X \leq b^*(B^1) \end{cases}$$

Theorem 4 illustrates how we can expand the results of Section 4.3 to more general payoff structures (e.g., allowing A to have state-dependent utility or R to bear some cost of experimentation) and allows us to easily state the optimal mechanisms for a number of standard environments outside of the experimentation/learning framework considered up until now. For example, we can consider a real-option game similar to that of Grenadier et al. (2016).

Example 1. *Suppose that A is tasked with running a project and R is an outside investor who can invest in the project. The value of the project, if invested in, is given by X_t , which solves the stochastic differential equation*

$$dX_t = \mu X_t dt + \sigma X_t dB_t,$$

The payoffs to R is zero if he rejects the project at $X_t - K$ (for some $K \in \mathbb{R}_+$) if he invests in the project while the payoff to A is $\beta X_t + L_1$ for some $\beta \in \mathbb{R}_+$ and $L_1 \in \mathbb{R}$ if R invests and $L_2 \in \mathbb{R}_+$ if R rejects or A quits.

Our general model can also be used to model a manager's decision of whether or not to promote an agent, in which case it is realistic for the agent's outside option may depend on the beliefs about his type (something not capture in the model of Section 3).

Example 2. *Suppose that R is a manager deciding whether or not to promote an agent A . The agent pays a flow cost c until he is either promoted or let go. The agent's type is either $\{\theta_h, \theta_\ell\}$ and R only wants to hire a θ_h type. Both R and A have the same belief about the agent's type¹⁸ and they learn about the agent's type by observing a Brownian motion X_t with type-dependent drift*

$$dX_t = \mu_\theta dt + \sigma dB_t.$$

If A is promoted, he receives a payoff of 1 while if he is let go or he quits, he receives a payoff of $f(X_t, 0)$ (we can interpret this as the outside wage he will receive given the market's belief about his type, which may depend on the information revealed over the course of the game). If R promotes a θ type, she receives a utility $g_\theta(X_t, 1)$ (we allow her utility to depend on the agent's type as well as her beliefs at the time of promotion; this can capture situations in which how the A is viewed by other employees at the time of promotion determines the payoff to approving A).

¹⁸While it is often the case in real-life that A will have more information about his type than R , we can view the type here as being indicative of the productivity match between R and A , in which case the symmetric information assumption is more innocuous.

If the primitives of this problem satisfy Assumptions 2-4, then Theorem 4 implies that R will use a decreasing approval threshold over time. This means that if A is higher quickly, he is more likely to be a good type than if he took a long time to be promoted, which fits a natural intuition on type inference in these situations that single decision-maker models fail to capture.

We can then also apply the model to study a lobbying situation. The decreasing threshold then corresponds to a gradual decrease in R 's demands, something we naturally see in many real life negotiations.

Example 3. Let R be a company lobbying with a politician A over the supply of some good. The politician wants to be seen as proactive and derives a utility of 1 whenever a deal is made. The degree of public support X_t determines the payment A can offer R for the good (let's assume when public support is X_t , A can offer R X_t). R derives a utility $u_R(X_t) - K$ when a deal is reached and A offers the maximum possible when public support is X_t . Additionally, A can exert costly effort (with a flow cost c) to rally public support for the deal, so that public support evolves according to the the diffusion process

$$dX_t = \mu(X_t)dt + \sigma(X_t)dB_t.$$

We can then interpret "approval" as R agreeing to a deal when offered X_t . The approval threshold is analogous to the demand of R . Under Assumptions 2-4, Theorem 4 then implies that R 's optimal negotiating strategy is to slowly decrease his demand if public support decreases.

8 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 show how

these results can be generalized to a large class of payoff functions and diffusion processes, allowing us to explore their implications in a number of other economic settings, such as promotion or lobbying models.

We also apply the model to the case when the agent has private information. The optimal solution may take the form of a fast-track mechanism: high types are offered a low starting approval threshold, but if the evidence becomes too unfavorable, 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. Using data on FDA approval decisions and Type I error, we show that the predictions of our optimal model shows that agency considerations can explain the empirical relationship between Type I error and the length of clinical trials. Additionally, we show how our fast-track mechanism matches many features of FDA expedited approval programs.

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] Carpenter D, Zucker EJ, Avorn J. Drug Review “Deadlines and Subsequent Safety Problems. *New England Journal of Medicine.*” 2008;358 (13) :1354-61.

- [8] Chan, Hing Chi Jimmy, Alessandro Lizzeri, Wing Suen, and Leeat Yariv. “Deliberating Collective Decisions.” (2015).
- [9] Dayanik, Savas, and Ioannis Karatzas. “On the optimal stopping problem for one-dimensional diffusions.” *Stochastic Processes and their Applications* 107.2 (2003): 173-212.
- [10] DiMasi, Joseph. “Innovation in the Pharmaceutical Industry: New Estimates of R&D Costs.” (2014)
- [11] Dokuchaev, N.G. “Optimal stopping of stochastic processes in a problem with constraints.” *Th. Probab. Appl.* 41 (1997), N 4. 761-768.
- [12] FDA. “The Drug Development Process - Step 3: Clinical Research.” U S Food and Drug Administration Home Page. Office of the Commissioner, 05 May 2017. Web. 14 June 2017.
- [13] Fong, Kyna. “Evaluating skilled experts: Optimal scoring rules for surgeons.” Working Paper, 2009.
- [14] Garfagnini, Umberto. “Delegated experimentation.” Working paper, 2011.
- [15] Gomes, Renato, Daniel Gottlieb, and Lucas Maestri. “Experimentation and project selection: Screening and learning.” *Games and Economic Behavior* 96 (2016): 145-169.
- [16] Grenadier, Steven R., Andrey Malenko, and Nad Malenko. “Timing Decisions in Organizations: Communication and Authority in a Dynamic Environment.” *American Economic Review* (Forthcoming).
- [17] Guo, Yingni. “Dynamic delegation of experimentation.” *American Economic Review* (2016).
- [18] Halac, Marina, Navin Kartik, and Qingmin Liu. “Optimal contracts for experimentation.” *The Review of Economic Studies* (2016): rdw013.
- [19] Halac, Marina, Navin Kartik, and Qingmin Liu. “Contests for experimentation.” *Journal of Political Economy*, Forthcoming (2014).
- [20] Henry, Emeric, and Marco Ottaviani. “Research and the Approval Process: The Organization of Persuasion.” Working Paper, (2017).

- [21] Hörner, Johannes, and Larry Samuelson. “Incentives for experimenting agents.” *The RAND Journal of Economics* 44, no. 4 (2013): 632-663.
- [22] Keller, Godfrey, Sven Rady, and Martin Cripps. “Strategic experimentation with exponential bandits.” *Econometrica* 73, no. 1 (2005): 39-68.
- [23] Keller, Godfrey, and Sven Rady. “Strategic experimentation with Poisson bandits.” *Theoretical Economics* 5, no. 2 (2010): 275-311.
- [24] Kolb, Aaron. “Strategic Real Options.” Working Paper (2016).
- [25] Kruse, Thomas, and Philipp Strack. “Optimal stopping with private information.” *Journal of Economic Theory* 159 (2015): 702-727.
- [26] Madsen, Erik. “Optimal project termination with an informed agent.” Working Paper (2016)
- [27] Matveev, A., and V. Yakubovich. “Nonconvex problems of global optimization: linear-quadratic control problems with quadratic constraints.” *Dynamics and Control* 7.2 (1997): 99-134.
- [28] McClellan, Andrew. “Dynamic Project Standards with Adverse Selection.” Working Paper (2017)
- [29] Moscarini, Giuseppe, and Lones Smith. “The optimal level of experimentation.” *Econometrica* 69, no. 6 (2001): 1629-1644.
- [30] Orlov, Dmitry, Andrzej Skrzypacz, and Pavel Zryumov. “Persuading the Regulator To Wait.” Working Paper.
- [31] Peskir, Goran, and Albert Shiryaev. *Optimal stopping and free-boundary problems*. Birkhäuser Basel, 2006.
- [32] Sannikov, Yuliy. “A continuous-time version of the principal-agent problem.” *The Review of Economic Studies* 75, no. 3 (2008): 957-984.
- [33] Strulovici, Bruno. “Learning while voting: Determinants of collective experimentation.” *Econometrica* 78, no. 3 (2010): 933-971.

Appendices

A Properties of V, J

Lemma 13. $V(B, b, X), J(B, b, X)$ are single-peaked in B and, for a fixed b , we have $\operatorname{argmax}_B J(B, b, X) \geq \operatorname{argmax}_B V(B, b, X)$.

Proof. The single-peaked property follows from Lemma 1 of Chan et. al (2016).

We argue that for a fixed b , the optimal threshold B for A is lower than that of R . For R , the the optimal threshold solves the first-order condition:

$$\Psi_B(1 - e^{-B}) + \Psi e^{-B} = 0,$$

and the optimal threshold for A satisfies

$$\Psi_B(1 + \frac{c}{r} + (a + \frac{c}{r})e^{-B}) - (a + \frac{c}{r})\Psi e^{-B} + \psi_B \frac{c}{r} = 0.$$

We note that the derivative of the first-order condition with respect to a is

$$(\Psi_B - \Psi)e^{-B} < 0,$$

and thus by the implicit function theorem and the second-order condition, the optimal threshold is decreasing in a . Therefore it is enough to prove the claim for $a = -1$, in which case A 's first-order condition is

$$\Psi_B(1 + \frac{c}{r} + (-1 + \frac{c}{r})e^{-B}) - (-1 + \frac{c}{r})\Psi e^{-B} + \psi_B \frac{c}{r} = 0. \quad (3)$$

Let $\Delta = B - b$. The derivative of the first-order condition with respect to c is

$$\begin{aligned} & \Psi_B(1 + e^{-B}) - \Psi e^{-B} + \psi_B \\ &= \Psi \left(\frac{R_1 e^{-R_1 \Delta} - R_2 e^{-R_2 \Delta}}{e^{-R_1 \Delta} - e^{-R_2 \Delta}} (1 + e^{-B}) - e^{-B} + \frac{R_2 e^{-\Delta} - R_1 e^{-\Delta}}{e^{-R_1 \Delta} - e^{-R_2 \Delta}} \right) \end{aligned}$$

Thus the above is negative if and only if

$$(R_1 e^{-R_1 \Delta} - R_2 e^{-R_2 \Delta})(1 + e^{-B}) - e^{-B}(e^{-R_1 \Delta} - e^{-R_2 \Delta}) + R_2 e^{-\Delta} - R_1 e^{-\Delta} \quad (4)$$

At $\Delta = 0$, this is equal to

$$e^{-B}(R_1 - R_2) < 0.$$

If we take the derivative of equation 4 with respect to Δ when equation 4 is equal to zero, we have

$$\begin{aligned} & (-R_1^2 e^{-R_1 \Delta} + R_2^2 e^{-R_2 \Delta})(1 + e^{-B}) + e^{-B}(R_1 e^{-R_1 \Delta} - R_2 e^{-R_2 \Delta}) - R_2 e^{-\Delta} + R_1 e^{-\Delta} \\ &= (R_1(1 - R_1)e^{-R_1 \Delta} + R_2(R_2 - 1)e^{-R_2 \Delta})(1 + e^{-B}) + e^{-B}((R_1 - 1)e^{-R_1 \Delta} - (R_2 - 1)e^{-R_2 \Delta}) \\ &= R_1 R_2 (e^{-R_1 \Delta} - e^{-R_2 \Delta})(1 + e^{-B}) + e^{-B} e^{-B} ((R_1 - 1)e^{-R_1 \Delta} - (R_2 - 1)e^{-R_2 \Delta}) \\ &< 0. \end{aligned}$$

Therefore, we know that equation 4 is always negative. Therefore, increasing c decreases the right-hand side of equation 3. By the implicit function theorem and the second-order condition, we have that A 's optimal B is decreasing in c . Since the R and A optimal thresholds are equal when $c = 0$, it must be that A 's optimal B is lower than that of R . \square

Lemma 14. \tilde{V} satisfies single crossing of 0 with respect to X .

Proof. Suppose that $\exists X^1 < X^2$ such that $\tilde{V}(B, b, X^1) = \tilde{V}(B, b, X^2) = 0$. Then for any $X \in (X^1, X^2)$, we have

$$\begin{aligned} \tilde{V}(B, b, X^1) &= \mathbb{E}[e^{-r\tau}(d_\tau + \frac{c}{r})|X] - \frac{c}{r} \\ &= \mathbb{E}[e^{-r\tau(X^1)} \mathbf{1}(\tau(X^1) < \tau(X^2))(\tilde{V}(B, b, X^1) + \frac{c}{r})|X] \\ &\quad + \mathbb{E}[e^{-r\tau(X^2)} \mathbf{1}(\tau(X^1) > \tau(X^2))(\tilde{V}(B, b, X^2) + \frac{c}{r})|X] - \frac{c}{r} \\ &= \mathbb{E}[e^{-r\tau(X^1)} \mathbf{1}(\tau(X^1) < \tau(X^2)) \frac{c}{r} |X] + \mathbb{E}[e^{-r\tau(X^2)} \mathbf{1}(\tau(X^1) > \tau(X^2)) \frac{c}{r} |X] - \frac{c}{r} \\ &< 0. \end{aligned}$$

\square

B General Optimal Stopping Properties

We now present several general properties of single-decision optimal stopping problems which will prove useful in our analysis.

Lemma 15. *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

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

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 τ 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 $Z'_1 = \sup_{(\tau, d_\tau)} \{Z \in \partial\mathcal{C} : Z \leq Z_0\}$ is the highest boundary of \mathcal{C} which is below Z_0 and $Z'_2 = \inf_{\mathcal{C}} \{Z \in \partial\mathcal{C} : Z \geq Z_0\}$ 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 . \square

Lemma 16 (Duality). *Let $\{\phi_i\}_{i=1}^n$ and Φ be bounded \mathcal{F}_t^X -measurable functions and define*

$$C := \{(\tau, d_\tau) : \mathbb{E}[\phi_i(\tau, \omega, d_\tau)|Z_0] \leq 0 \forall i = 1, \dots, n\}.$$

Suppose that $\exists(\tau, d_\tau)$ such that $\mathbb{E}[\phi_i(\tau, \omega, d_\tau)|Z_0] < 0 \forall i = 0, \dots, N$ and that the optimal solution to $\sup_{(\tau, d_\tau) \in C} \mathbb{E}[\Phi(\tau, \omega, d_\tau)|Z_0]$ is such that $\mathbb{P}(\tau > 0) = 1$. Then there is no duality gap-i.e.,

$$\sup_{(\tau, d_\tau) \in C} \mathbb{E}[\Phi(\tau, \omega, d_\tau) | Z_0] = \inf_{\lambda \in \mathbb{R}_-^{N+1}} \sup_{(\tau, d_\tau)} \mathbb{E}[\Phi(\tau, \omega, d_\tau) | Z_0] + \sum_{i=0}^N \lambda_i \mathbb{E}[\phi_i(\tau, \omega, d_\tau) | Z_0].$$

Moreover, the infimum is obtained by some finite $\lambda^* \in \mathbb{R}_-^{N+1}$. Additionally, (τ, d_τ) is a solution to $\sup_{(\tau, d_\tau) \in C} \mathbb{E}[\Phi(\tau, \omega, d_\tau) | Z_0]$ if and only if it is a solution to $\sup_{(\tau, d_\tau)} \mathbb{E}[\Phi(\tau, \omega, d_\tau) | Z_0] + \sum_{i=0}^N \lambda_i^* \mathbb{E}[\phi_i(\tau, \omega, d_\tau) | Z_0]$ and complementary slackness conditions hold:

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

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

Lemma 17. Let $G(\pi_t, d_t) = d_t(\alpha_1 \pi_t + \alpha_2) + \alpha_3$ and $\alpha_3 \geq 0$. If the solution to

$$V(\pi_t) = \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} G(\pi_\tau, d_\tau) | \pi_0],$$

is a static threshold mechanism with approval at both the upper threshold π_B and the lower threshold π_b , then $\pi_B = \pi_b$.

Proof. By standard arguments, V solves the differential equation $rV(\pi) = \phi\pi^2(1 - \pi^2)V''(\pi)$. Then $V(\pi) > 0$, which implies that $V''(\pi) \geq 0$. Let $\beta = \frac{\pi - \pi_b}{\pi_B - \pi_b}$. Then

$$\begin{aligned} \alpha_1 \pi_0 + \alpha_2 + \alpha_3 &\leq V(\pi_0) = V(\beta\pi_b + (1 - \beta)\pi_B) \\ &\leq \beta V(\pi_b) + (1 - \beta)V(\pi_B) \\ &= \beta(\alpha_1 \pi_b + \alpha_2 + \alpha_3) + (1 - \beta)(\alpha_1 \pi_B + \alpha_2 + \alpha_3) \\ &= \alpha_1 \pi_0 + \alpha_2 + \alpha_3, \end{aligned}$$

which implies that immediate approval is optimal. □

C Symmetric Information

C.1 Two-Sided Commitment

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:

$$\frac{\Psi_b(1 - e^{-B})}{\frac{c}{r}[\psi_b(1 + e^{-b}) - \psi e^{-b}] + \Psi_b[(1 + \frac{c}{r} + (a + \frac{c}{r})e^{-B})]} = \frac{\Psi_B(1 - e^{-B}) + e^{-B}\Psi}{a\Psi e^{-B} + \Psi_B(1 + \frac{c}{r} + (a + \frac{c}{r})e^{-B}) + \psi_B(1 + e^{-b})\frac{c}{r}}$$

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

where $B < B^{FB}$ and $P(\tau, d_\tau)$ is binding. If $b = -\infty$, then $B = B^{FB}$.

Proof. We start by proving the conditions of Lemma 16 are met. To see this, note that the stopping policy $\tau = \epsilon$ and $d_\tau = 1$ will keep the participation constraint slack for ϵ small enough. The other conditions of Lemma 16 are easily checked.

By applying Lemma 16, we can use a Lagrangian in order to turn the primal problem:

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

subject to

$$P : \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_0] - \frac{c}{r} \geq 0,$$

into the dual problem

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

By Lemma 15, 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

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

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

D Symmetric Information with One-Sided Commitment

D.1 Proof of Lemma 1

Proof. Let (τ, d_τ) be a mechanism which satisfies all dynamic participation constraints. Suppose that it did not satisfy a *DP* constraint-i.e., $\exists \tau'$ such that

$$\begin{aligned} & \mathbb{E}[e^{-r\tau}(d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_0] - \mathbb{E}[e^{-r(\tau \wedge \tau')} (d_\tau \mathbf{1}(\tau' \leq \tau) + \frac{c}{r})|Z_0] < 0 \\ \Rightarrow & \mathbb{E}[e^{-r\tau}(d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})\mathbf{1}(\tau > \tau')|Z_0] < \frac{c}{r}. \end{aligned}$$

Therefore, there must be a history $h_{\tau'}$ such that at τ' we have $\mathbb{E}[e^{-r\tau}(d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_{\tau'}, h_{\tau'}] < \frac{c}{r}$, a contradiction of the fact that all dynamic participation constraints hold. Therefore all (τ, d_τ) which satisfy dynamic participation constraints also satisfy *DP* constraints.

Next, let us consider a mechanism $(\tilde{\tau}, \tilde{d}_\tau)$ which satisfies *DP*. We will construct a new mechanism, $(\hat{\tau}, \hat{d}_\tau)$ which satisfies all dynamic participation constraints and gives the same payoff to R as $(\tilde{\tau}, \tilde{d}_\tau)$. If $(\tilde{\tau}, \tilde{d}_\tau)$ satisfies all dynamic participation constraints, we are done. If some dynamic participation constraints are violated, we claim that it must happen only on a zero probability set. Define, for some small ϵ , $\Gamma = \{h_t \in \mathbb{H}_t : \mathbb{E}[e^{-r\tau}(d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_t, h_t] \leq \frac{c}{r} - \epsilon\}$ to be the set of histories such that A 's continuation value is at least ϵ worse than quitting immediately. Let us define $\tau' = \inf\{t : h_t \in \Gamma\}$ to be the first time the history is in Γ . Then we know that

$$\begin{aligned} \mathbb{E}[e^{-r(\tau \wedge \tau')} (d_\tau \mathbf{1}(\tau \leq \tau') \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_0] &= \mathbb{E}[\mathbf{1}(\Omega \setminus \Gamma) e^{-r(\tau \wedge \tau')} (d_\tau \mathbf{1}(\tau \leq \tau') \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_0] \\ &+ \mathbb{E}[\mathbf{1}(\Gamma) e^{-r(\tau \wedge \tau')} (d_\tau \mathbf{1}(\tau \leq \tau') \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_0] \\ &= \mathbb{E}[\mathbf{1}(\Omega \setminus \Gamma) e^{-r\tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_0] \\ &+ \mathbb{E}[\mathbf{1}(\Gamma) e^{-r(\tau \wedge \tau')} (d_\tau \mathbf{1}(\tau \leq \tau') \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_0]. \end{aligned}$$

Because $(\tilde{\tau}, \tilde{d}_\tau)$ satisfies *DP*, we know that

$$\begin{aligned}
\mathbb{E}[e^{-r\tilde{\tau}}(\tilde{d}_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_0] &\geq \mathbb{E}[\mathbf{1}(\Omega \setminus \Gamma)e^{-r\tilde{\tau}}(\tilde{d}_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_0] \\
&\quad + \mathbb{E}[\mathbf{1}(\Gamma)e^{-r(\tilde{\tau} \wedge \tau')} (d_\tau \mathbf{1}(\tilde{\tau} \leq \tau') \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_0] \\
\Rightarrow \mathbb{E}[\mathbf{1}(\Gamma)e^{-r\tilde{\tau}}(\tilde{d}_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_0] &\geq \mathbb{E}[\mathbf{1}(\Gamma)e^{-r(\tilde{\tau} \wedge \tau')} (d_\tau \mathbf{1}(\tilde{\tau} \leq \tau') \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_0] \\
&= \mathbb{E}[\mathbf{1}(\Gamma)e^{-r\tau'} \frac{c}{r}|Z_0],
\end{aligned}$$

where the final line holds since a dynamic participation constraint can only be violated if $\tilde{\tau}$ has not been reached. We note that

$$\mathbb{E}[\mathbf{1}(\Gamma)e^{-r\tilde{\tau}}(\tilde{d}_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_0] = \mathbb{E}[\mathbf{1}(\Gamma)\mathbb{E}[e^{-r\tilde{\tau}}(\tilde{d}_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_{\tau'}, h_{\tau'}]|Z_0].$$

By definition of Γ , we have that for each $h_{\tau'}'$ in B ,

$$\begin{aligned}
\mathbb{E}[e^{-r\tilde{\tau}}(\tilde{d}_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_{\tau'}, h_{\tau'}'] &< e^{-r\tau'} \frac{c}{r} \\
\Rightarrow \mathbb{E}[\mathbf{1}(\Gamma)\mathbb{E}[e^{-r\tilde{\tau}}(\tilde{d}_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_{\tau'}, h_{\tau'}']|Z_0] &< \mathbb{E}[\mathbf{1}(\Gamma)e^{-r\tau'} \frac{c}{r}|Z_0].
\end{aligned}$$

a contradiction. Therefore, it cannot be that Γ has strictly positive probability and this must hold for all ϵ .

Suppose that $(\tilde{\tau}, \tilde{d}_\tau)$ does violate some dynamic participation constraints. Then we know that the set of histories Γ such that the constraints are violated has probability zero. Therefore, we can specify $(\hat{\tau}, \hat{d}_\tau)$ to be equal to $(\tilde{\tau}, \tilde{d}_\tau)$ on the set of all histories not in Γ and $(\hat{\tau}, \hat{d}_\tau)$ to reject at the first time that a dynamic participation constraint is violated. With probability one, the outcome from $(\hat{\tau}, \hat{d}_\tau)$ is the same as $(\tilde{\tau}, \tilde{d}_\tau)$ and thus they must yield the same payoffs. \square

D.2 Proof of Theorem 1

Theorem 1 and supporting Lemmas are special cases Theorem 4 and supporting Lemmas (Assumption 2 clearly holds; Assumption 3 follows by Lemma 13). We therefore defer to proof to Section I.

Lemma 5. *The optimal approval threshold in belief-space $B_Z(M_t^Z)$ is independent of Z_0 and depends only on $M_t^Z := \min\{Z_s : s \leq t\}$.*

Proof. Let us now consider the optimal mechanism as defined in belief space. Suppose that at two different initial beliefs Z_a and Z_c , the optimal mechanism called for different initial approval thresholds (say B_a^1 and B_c^1 such that $B_a^1 > B_c^1$). Let $b_i^1 = b_Z^*(B_i^1)$ and suppose that $b_a^1 < Z_c$. Define $J_i(z)$ be the utility of the mechanism under Z_i at $\tau(z)$. Note that switching over to the mechanism for $J_c(b_a^1)$ is admissible for $J_a(b_a^1)$ since it lowers the approval threshold, which slackens DP constraints. Then we know that

$$\begin{aligned} & \mathbb{E}[e^{-r\tau} d_\tau(b_a^1) \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} + e^{-r\tau(b_a^1)}(1 - d_\tau(b_a^1))J_a(b_a^1)|Z_a] \\ & \geq \mathbb{E}[e^{-r\tau} d_\tau(b_a^1) \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} + e^{-r\tau(b_a^1)}(1 - d_\tau(b_a^1))J_c(b_a^1)|Z_a], \end{aligned}$$

which, since $J_i(b_a^1)$ is independent of the history prior to $\tau(b_a^1)$, implies that $J_c(b_a^1) \leq J_a(b_a^1)$. But, since the mechanism at $J_a(b_a^1)$ is admissible with respect to $J_c(b_a^1)$ (since by definition it satisfies all DP constraints and DP constraints prior to $\tau(b_a^1)$ will not be violated since $B_c^1 > B_a^1$), we also know that

$$\begin{aligned} & \mathbb{E}[e^{-r\tau} d_\tau(b_a^1) \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} + e^{-r\tau(b_a^1)}(1 - d_\tau(b_a^1))J_c(b_a^1)|Z_c] \\ & \geq \mathbb{E}[e^{-r\tau} d_\tau(b_a^1) \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} + e^{-r\tau(b_a^1)}(1 - d_\tau(b_a^1))J_a(b_a^1)|Z_c], \end{aligned}$$

which implies that $J_c(b_a^1) = J_a(b_a^1)$.

It is then without loss to assume that at b_a^1 , both Z_a and Z_c use the same mechanism. Treating this as a continuation value upon $\tau(b_a^1)$, we can see that the initial choice of Z_i must satisfy

$$\max_B \Psi \frac{e^{Z_i}(1 - e^{-B})}{1 + e^{Z_i}} + \psi \frac{e^{Z_i}(1 - e^{-b})}{1 + e^{Z_i}} J_a(b_a^1).$$

Taking the first-order condition, we can easily show that the choice of B_i^1 is independent of Z_i and thus Z_a and Z_c must use the same mechanism.

Finally, we consider the case where $b_a^1 > Z_c$. Consider the continuation mechanism for Z_i at time $\tau(Z_c)$. Since the mechanism for $J_a(Z_c)$ satisfies all DP constraints, it is admissible with respect to $J_c(Z_c)$; thus we must have $J_a(Z_c) \leq J_c(Z_c)$. Since the mechanism for $J_c(Z_c)$ also satisfies DP constraints, it is admissible with respect to $J_a(Z_a)$ to switch to the mechanism for $J_c(Z_c)$ at $\tau(Z_c)$, which tells us

$$\begin{aligned} & \mathbb{E}[e^{-r\tau} d_\tau(Z_c) \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} + e^{-r\tau(b_a^1)}(1 - d_\tau(Z_c))J_a(Z_c)|Z_a] \\ & \geq \mathbb{E}[e^{-r\tau} d_\tau(b_a^1) \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} + e^{-r\tau(b_a^1)}(1 - d_\tau(b_a^1))J_a(Z_c)|Z_a], \end{aligned}$$

and implies that $J_a(Z_c) = J_c(Z_c)$. Therefore, we can replace the mechanism for $J_a(Z_c)$ with that of $J_c(Z_c)$ and both solutions will be identical up to the minimum M_t^Z . \square

E Symmetric Information with No Commitment

E.1 Proof of Proposition 2

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. \square

E.2 Proof of Proposition 3

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

- Experiment until τ^* .
- 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 τ^* . If she doesn't approve, then the agent quits experiment at time τ^* forever. Since no new learning occurs, R has a strict incentive to approve immediately at τ^* since $Z_{\tau^*} > 0$. Suppose R had a profitable deviation τ' such that $\tau' \leq \tau^*$ and there is some history such

that τ' approves strictly sooner than τ^* .¹⁹ Than τ' will not violate any *DP* constraints (approving sooner would only slacken the DP constraints), contradicting the optimality of τ^* . 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 τ^* 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, (τ^*, d_τ^*) is indeed an equilibrium. \square

F Asymmetric Information with Two-Sided Commitment

F.1 Proof of Proposition 4

Proof. Let $V_i \geq 0$ be the expected utility to type i from truthfully declaring his type. Then we can write the problem for determining type i 's mechanism to be

$$\begin{aligned} & \sup_{(\tau, d_\tau)} \mathbb{E}\left[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} \mid Z_i\right] \\ & \text{subject to (for } k \neq i) \\ & PK : \mathbb{E}\left[e^{-r\tau} \left(d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}\right) \mid Z_i\right] - \frac{c}{r} \geq V_i \\ & IC(Z_k, Z_i) \mathbb{E}\left[e^{-r\tau} \left(d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}\right) \mid Z_k\right] - \frac{c}{r} \leq V_k \end{aligned}$$

The only complication that prevents the applications of Lemma 16 and Lemma 15 to reach our desired conclusion that static threshold mechanisms are optimal is the fact that the *IC* expectation is taken with respect to Z_k . However, the same argument as in Lemma 19 allows us to convert the expectation into one with respect to Z_i , allowing us to reach our desired conclusion. \square

¹⁹*R* will never find it profitable to reject earlier

F.2 Proof of Proposition 5

Proof. Let (b_h, B_h) be the thresholds used in h 's mechanism and suppose that $b_h \neq -\infty$ and R 's symmetric information mechanism for Z_h doesn't involve immediate rejection. First, we will show that we cannot have $IC(Z_h, Z_\ell)$ slack. Suppose that it were. It is easy to see that we must then have $IC(Z_\ell, Z_h)$ binding. We will show that we can reduce b_h and increase R 's utility while still satisfying $IC(Z_\ell, Z_h)$ and $IC(Z_h, Z_\ell)$. Suppose that $b_h > b_h^*(B_h)$ (where b_h^* is as defined in Section 5.2). By Lemma ??, we can decrease b_h to b'_h such that $\tilde{V}_X(B_h, b_h, Z_h) = \tilde{V}_X(B_h, b'_h, Z_h)$ and $\tilde{V}_X(B_h, b_h, Z_h) > \tilde{V}_X(B_h, b'_h, Z_h)$ (where $\tilde{V}_X(B, b, Z)$ is the corresponding threshold utility in X_t -space given (b, B) and initial belief Z_ℓ). Since R 's utility is decreasing in b , (b_h, B_h) cannot have been optimal. If $b_h < b_h^*(B_h)$, then we can decrease b_h slightly and increase R 's utility while decreasing ℓ 's utility since $b_h^*(B_h) < b_\ell^*(B_\ell)$. This is admissible since $IC(Z_h, Z_\ell)$ is not binding by assumption and $IC(Z_\ell, Z_K)$ and $P(Z_\ell)$ imply that $P(Z_h)$ is slack:

$$0 \leq \tilde{V}_X(B_\ell, b_\ell, Z_\ell) = \tilde{V}_X(B_h, b_h, Z_\ell) \leq \tilde{V}_X(B_h, b_h, Z_h).$$

Therefore, we cannot have $IC(Z_h, Z_\ell)$ slack.

Finally, if R 's optimal symmetric information mechanism for h involves immediate rejection, then so does R 's optimal symmetric information mechanism for ℓ . Therefore, R can achieve the value of his problem without IC constraints (an upper-bound on the problem including IC constraints) by rejecting immediately, in which case $IC(Z_h, Z_\ell)$ binds. □

G Asymmetric Information with One-Sided Commitment

G.1 Proof of Lemmas 7 and 8

Rather than directly prove Lemmas 7 and 8, we instead solve a generalization of $AM^{h,\gamma}$ which will be useful in proving Lemma 12. Fix an arbitrary $\gamma \in \mathbb{R}$ (the proof of Lemmas 7 and 8 follow from letting $\gamma = 0$); we define the problem $AM^{h,\gamma}$ as

$$[AM^{h,\gamma}] : \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau}(1+\gamma) - (a\gamma - 1)}{1 + e^{Z_\tau}} + \gamma \frac{c}{r}) | Z_h]$$

subject to $DP(Z_h)$, $DIC(Z_\ell, Z_h, V_\ell)$.

Let (τ^{SM}, d^{SM}) be the solution to $SM^\gamma(Z_h)$ where $SM^\gamma(Z_h)$ is the symmetric information problem with prior Z_h when R 's payoffs depend on γ as above. We define $b^{SM, \gamma}$ to be the X_t such that, when ℓ has belief $Z_t = Z_\ell + \frac{\phi}{\sigma} X_t$, he would quit immediately if R proposed h 's symmetric mechanism $SM^\gamma(X_t)$ (the symmetric mechanism starting at $X_t, M_t^X = X_t$):

$$b^{SM, \gamma} := \max\{X_t : \sup_{\tau'} \mathbb{E}[e^{-r(\tau_h^{SM} \wedge \tau')} (d_{\tau, h}^{SM} \mathbf{1}(\tau \leq \tau') \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | X_t, Z_t = Z_\ell + \frac{\phi}{\sigma} X_t] = \frac{c}{r}\}.$$

We begin by proving solving a relaxed problem $RAM_N^h(\gamma)$ defined as

$$\begin{aligned} [RAM_N^{h, \gamma}] : \quad & \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r(\tau \wedge \tau(b^{SM, \gamma}))} (d_\tau(b^{SM, \gamma})) (\frac{e^{Z_\tau} (1 + \gamma) + (a\gamma - 1)}{1 + e^{Z_\tau}} + \gamma \frac{c}{r}) \mathbf{1}(\tau \leq \tau(b^{SM, \gamma})) \\ & + \mathbf{1}(\tau > \tau(b^{SM, \gamma})) SM^{h, \gamma}(b^{SM, \gamma}) | Z_h] \\ & \text{subject to } \forall X_i \in \mathcal{T}_N \cup \{b^{SM, \gamma}\} \\ RDIC_\ell(X_i) : \quad & \mathbb{E}[e^{-r(\tau \wedge \tau(X_i) \wedge \tau(b^{SM, \gamma}))} (d(X_i)) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_\ell] \leq V_\ell + \frac{c}{r}. \end{aligned}$$

Define an analogous version of $H^h(X_t)$ by

$$\begin{aligned} [H_N^{h, \gamma}(X_t)] : \quad & \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r(\tau \wedge \tau(b^{SM, \gamma}))} (d_\tau(b^{SM, \gamma})) \frac{e^{Z_\tau} (1 + \gamma) + (a\gamma - 1)}{1 + e^{Z_\tau}} + \gamma \frac{c}{r}) \mathbf{1}(\tau \leq \tau(b^{SM, \gamma})) \\ & + \mathbf{1}(\tau > \tau(b^{SM, \gamma})) SM^{h, \gamma}(b^{SM, \gamma}) | Z_t] \\ & \text{subject to } \forall X_i \in \{X_j \in \mathcal{T}_N \cup \{b^{SM, \gamma}\} : X_j < X_t\} \\ RPK(0) : \quad & \mathbb{E}[e^{-r(\tau \wedge \tau(X_i) \wedge \tau(b^{SM, \gamma}))} (d_\tau(X_i)) \mathbf{1}(\tau \leq \tau(b^{SM, \gamma})) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_t - \Delta_Z] \leq \frac{c}{r}. \end{aligned}$$

Lemma 18. $RAM_N^{h, \gamma}$ is an upper-bound on $AM^{h, \gamma}$.

Proof. First, we claim that the continuation function for $AM^{h, \gamma}$ at $\tau(b^{SM, \gamma})$ is bounded above by $SM^{h, \gamma}$. This is clear, since any mechanism admissible with respect to $AM^{h, \gamma}$ must satisfy all DP constraints when starting at $b_h^{SM}(\gamma)$ and thus is admissible with respect to $SM^{h, \gamma}$. Moreover, since we have dropped some constraints, for any (τ, d_τ) admissible with respect to $AM^{h, \gamma}$, we have that $(\tau \wedge \tau(b^{SM, \gamma}), d_\tau(b^{SM, \gamma}))$ is admissible with respect to $RAM_N^{h, \gamma}$. The result then follows immediately. \square

Lemma 19. *The solution to $RAM_N^{h,\gamma}$ is given by a stationary approval threshold until the first binding constraint X^1 . At X^1 , the continuation mechanism solves $H_N^{h,\gamma}(X^1)$.*

Proof. We face the new complication in solving $RAM_N^{h,\gamma}$ in that the expectations are taken with respect to different priors, which makes the arguments from Lemma 23 inapplicable. However, we can note that

$$\begin{aligned}
& \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} (d_\tau(X_i) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_h - \Delta_z] \\
&= \frac{e^{Z_h - \Delta_z}}{1 + e^{Z_h - \Delta_z}} \mathbb{E}[e^{-r(\tau \wedge \tau(Z_i))} (\mathbf{1}(\tau' \leq \tau(X_i)) d'_\tau + \frac{c}{r}) | \theta = H] \\
&+ a \frac{1}{1 + e^{Z_h - \Delta_z}} \mathbb{E}[e^{-r(\tau \wedge \tau(Z_i))} (d_\tau(X_i) + \frac{c}{r}) | \theta = L] \\
&= \frac{1 + e^{Z_h}}{1 + e^{Z_h - \Delta_z}} \left(\frac{e^{Z_h}}{1 + e^{Z_h}} \mathbb{E}[e^{-r(\tau' \wedge \tau(X_i))} e^{-\Delta_z} (d_\tau(X_i) + \frac{c}{r}) | \theta = H] \right. \\
&\left. + a \frac{1}{1 + e^{Z_h}} \mathbb{E}[e^{-r(\tau \wedge \tau(Z_i))} (d_\tau(X_i) + \frac{c}{r}) | \theta = L] \right) \\
&= \frac{1 + e^{Z_h}}{1 + e^{Z_h - \Delta_z}} \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} (d_\tau(X_i) + \frac{c}{r}) \frac{e^{-\Delta_z} e^{Z_\tau \wedge \tau(X_i)} + a}{1 + e^{Z_{\tau' \wedge \tau(X_i)}}} | Z_h].
\end{aligned}$$

With this change of expectation, we can now apply Lemma 23 in order to conclude that the optimal mechanism consists of a static approval threshold until the first binding constraint X^1 is reached or $b^{SM,\gamma}$.

All that is left is to show that the continuation mechanism at $\tau(X^1)$ is the solution to $H_N^{h,\gamma}(X^1)$. This is immediate if $X^1 = b^{SM,\gamma}$. Therefore suppose that $X^1 > b^{SM,\gamma}$. As noted in Lemma 23, the optimal mechanism from $\tau(X^1)$ onward is independent of the history up to $\tau(X^1)$. Therefore, by complementary slackness at X^1 , we know that for $X_i < X^1$, we have

$$\mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} (d_\tau(X_i) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | X_0] \leq V_\ell + \frac{c}{r}. \quad (5)$$

Let $\tau[h_{\tau(X^1)}](X_i)$ be the threshold quitting time of X_i given $X_0 = X^1$. Since the mechanism at $\tau(X^1)$ is independent of the history until $\tau(X^1)$, we know that

$$\begin{aligned}
& \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} (d_\tau(X_i) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | X_0] \\
&= \mathbb{E}[\mathbf{1}(\tau \leq \tau(X^1)) e^{-r\tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | X_0] \\
&+ \mathbb{E}[\mathbf{1}(\tau > \tau(X^1)) e^{-r\tau(X^1)} \mathbb{E}[e^{-r(\tau[h_{\tau(X^1)}] \wedge \tau[h_{\tau(X^1)}](X_i))} (d_\tau(X_i) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | X^1] | X_0] \\
&= V_\ell + \frac{c}{r} - \frac{c}{r} \mathbb{E}[e^{-r\tau(X^1)} \mathbf{1}(\tau > \tau(X^1)) | X_0] \\
&+ \mathbb{E}[\mathbf{1}(\tau > \tau(X^1)) e^{-r\tau(X^1)} | X_0] \mathbb{E}[e^{-r(\tau[h_{\tau(X^1)}] \wedge \tau[h_{\tau(X^1)}](X_i))} (d_\tau(X_i) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | X^1],
\end{aligned}$$

which, with equation 5, implies that

$$\mathbb{E}[e^{-r(\tau[h_{\tau(X^1)}] \wedge \tau[h_{\tau(X^1)}](X_i))} (d_\tau(X_i) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | X^1] \leq \frac{c}{r}$$

and thus the expected continuation value to ℓ of continuing until $\tau(X_i)$ is weakly negative at X^1 . Therefore, the continuation mechanism at $\tau(X^1)$ is admissible with respect to $H_N^{h,\gamma}(X^1)$ and thus $H_N^{h,\gamma}(X^1)$ is weakly higher than the value of the optimal mechanism at $\tau(X^1)$. Suppose that using the mechanism (τ^H, d_τ^h) which solves $H_N^{h,\gamma}(X^1)$ yielded a strictly higher value to R than the continuation mechanism at $\tau(X^1)$. Then replacing the continuation mechanism at $\tau(X^1)$ with (τ^H, d_τ^H) would give a value of

$$\begin{aligned}
& \mathbb{E}[\mathbf{1}(\tau \leq \tau(X^1)) e^{-r\tau} (d_\tau \frac{e^{Z_\tau}(1 + \gamma) + (a\gamma - 1)}{1 + e^{Z_\tau}} + \gamma \frac{c}{r}) | Z_h] + \mathbb{E}[\mathbf{1}(\tau > \tau(X^1)) e^{-r\tau(X^1)} H^h(\gamma, X^1) | Z_h] \\
&\geq \mathbb{E}[\mathbf{1}(\tau \leq \tau(X^1)) e^{-r\tau} (d_\tau \frac{e^{Z_\tau}(1 + \gamma) + (a\gamma - 1)}{1 + e^{Z_\tau}} + \gamma \frac{c}{r}) | Z_h] \\
&+ \mathbb{E}[\mathbf{1}(\tau > \tau(X^1)) e^{-r\tau(X^1)} \mathbb{E}[e^{-r\tau[h_{\tau(X^1)}]} (d_\tau \frac{e^{Z_\tau}(1 + \gamma) + (a\gamma - 1)}{1 + e^{Z_\tau}} + \gamma \frac{c}{r}) | Z_h^1 = Z_h + \frac{\phi}{\sigma} X^1] | Z_h] \\
&= \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau}(1 + \gamma) + (a\gamma - 1)}{1 + e^{Z_\tau}} + \gamma \frac{c}{r}) | Z_h].
\end{aligned}$$

Moreover, the stopping rule given by replacing the continuation mechanism at $\tau(X^1)$ with (τ^H, d_τ^H) yields a value to ℓ when quitting at X_i of

$$\begin{aligned}
& \mathbb{E}[\mathbf{1}(\tau \leq \tau(X^1))e^{-r(\tau \wedge \tau(X_i))}(d_\tau(X_i)\frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_\ell] \\
& + \mathbb{E}[\mathbf{1}(\tau > \tau(X^1))e^{-r\tau(X^1)}\mathbb{E}[e^{-r(\tau^H \wedge \tau[h_{\tau(X^1)}](X_i))}(d_\tau\frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_\ell^1]|Z_\ell] \\
& = V_\ell + \frac{c}{r} - \frac{c}{r}\mathbb{E}[e^{-r\tau(X^1)}\mathbf{1}(\tau > \tau(X^1))|Z_\ell^1] \\
& + \mathbb{E}[\mathbf{1}(\tau > \tau(X^1))e^{-r\tau(X^1)}\mathbb{E}[e^{-r(\tau^H \wedge \tau[h_{\tau(X^1)}](X_i))}(d_\tau\frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})|Z_\ell^1]|Z_\ell] \\
& \leq V_\ell + \frac{c}{r}.
\end{aligned}$$

where $Z_\ell^1 = Z_\ell + \frac{\phi}{\sigma}X^1$ is belief of ℓ at X^1 . Thus, switching to (τ^H, d_τ^H) at $\tau(X^1)$ is admissible with respect to $RSM_N^{h,\gamma}$. Therefore, the optimal mechanism at $\tau(X^1)$ must solve $H_N^{h,\gamma}(X^1)$. \square

Lemma 20. *The limit as $N \rightarrow \infty$ of the mechanisms which solve $H_N^{h,\gamma}(X^1)$ is given by $\tau = \inf\{t : X_t \geq \underline{B}_\ell(M_t^X)\}$ and $d_\tau = \mathbf{1}(X_\tau = \underline{B}_\ell(M_\tau^X))$.*

Proof. Repeated application of Lemma 23 yields the conclusion that the approval rule is a sequence of thresholds $\{B_h^i\}$ such that the approval changes from B_h^i to B_h^{i+1} when the i th binding *RDIC* constraint is reached. Applying Lemma 19, we know that at each binding constraint $X^j > b^{SM,\gamma}$, the continuation mechanism solves $H_N^{h,\gamma}(X^j)$.

Let us consider the limit as $N \rightarrow \infty$. By the same arguments as in Lemma 25, we get that as long as *RPK* constraints are binding, the upper approval threshold must converge to $\underline{B}_\ell(M_t^X)$ as $N \rightarrow \infty$. All that is left to show is that *RPK* constraints are binding until $M_t^X = b^{SM,\gamma}$. Consider a fixed N and suppose that $\exists X^j > b^{SM,\gamma}$ such that $\exists X_i \in \mathcal{T}_N \in (b^{SM,\gamma}, X_j)$ and all *RPK* constraints prior to $b^{SM,\gamma}$ are slack in $H_N^{h,\gamma}(X^j)$. In this case, the complementary slackness conditions in our Lagrangian amount to solving

$$\begin{aligned}
& \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r(\tau \wedge \tau(b^{SM,\gamma}))}[(d_\tau(b^{SM,\gamma})\frac{e^{Z_\tau}(1 + \gamma) + (a\gamma - 1)}{1 + e^{Z_\tau}} + \gamma\frac{c}{r})\mathbf{1}(\tau \leq \tau(b^{SM,\gamma})) \\
& + \mathbf{1}(\tau > \tau(b^{SM,\gamma}))SM^{h,\gamma}(b^{SM,\gamma})]|Z_t].
\end{aligned}$$

The solution to this consists of only an approval threshold at B (since $SM^{h,\gamma}(b^{SM,\gamma}) \geq 0$, R never benefits from rejecting prior to $b^{SM,\gamma}$). We claim that B must be equal to the initial approval threshold in the symmetric mechanism $B_{SM}^1(\gamma)$.

We consider the corresponding relaxed version to $SM^{h,\gamma}(b^{SM,\gamma})$ in which we drop all DP constraints except threshold quitting rules in \mathcal{T}_N . This problem (whose value we call $RSM_N^h(\gamma)(X_t)$ when starting at X_t) yields the Lagrangian

$$\begin{aligned} \mathcal{L} = & \sup_{\tau, d_\tau} \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} (1 + \gamma) + \gamma \frac{c}{r}) \\ & + \sum_{j=1}^{|\mathcal{B}_N|} \lambda(X^j) (e^{-r(\tau \wedge \tau(X^j))} (d_\tau(X^j) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) - e^{-r\tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r})) | b^{SM,\gamma}]. \end{aligned}$$

With the dual-problem, we can apply the principle of optimality at $X_t > b^{SM,\gamma}$ to conclude that the optimal symmetric mechanism solves

$$\begin{aligned} \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r(\tau \wedge \tau(b^{SM,\gamma}))} & [(d_\tau(b^{SM,\gamma}) \frac{e^{Z_\tau} (1 + \gamma) + (a\gamma - 1)}{1 + e^{Z_\tau}} + \gamma \frac{c}{r}) \mathbf{1}(\tau \leq \tau(b^{SM,\gamma})) \\ & + \mathbf{1}(\tau > \tau(b^{SM,\gamma})) RSM_N^h(\gamma)(b^{SM,\gamma})] | Z_t]. \end{aligned}$$

which is identical to that the problem prior to $\tau(b_{SM}^1(\gamma))$ when we let $N \rightarrow \infty$ (the same arguments applied in Theorem 1 show that $RSM_N^\gamma \rightarrow SM^h(\gamma)$). Therefore, the approval threshold which solves $H_N^h(X^j)$ must be equal to $B_{SM}^1(\gamma)$. \square

Lemma 21. *Let (τ^N, d_τ^N) solve $RAM_N^{h,\gamma}$ and $(\tau, d_\tau) := \lim_{N \rightarrow \infty} (\tau^N, d_\tau^N)$. Then (τ, d_τ) solves $AM^{h,\gamma}$.*

Proof. Since $RAM_N^{h,\gamma}$ is an upper-bound on $AM^{h,\gamma}$, all that is necessary to verify is that (τ, d_τ) is admissible with respect to $AM^{h,\gamma}$. First, we need to show that h has no incentive to quit early. This is immediate after $\tau(b^{SM,\gamma})$ since the symmetric mechanism, which satisfies all DP constraints, is used. Before $\tau(b^{SM,\gamma})$, the continuation value to h is always positive since $\underline{B}_\ell(M_t^X) < \underline{B}_h(M_t^X)$.

We also need to check that ℓ 's DIC constraint truly does hold. To this end, we claim that the optimal quitting rule ℓ could use is a threshold quitting rule. Note that the optimal mechanism is Markov with respect to (X_t, M_t^X) and thus, by the principle of optimality, ℓ 's optimal quitting rule will also be Markov with respect to (X_t, M_t^X) and thus can be expressed as $\tau' = \inf\{t : X_t \geq \kappa(M_t^X)\} \wedge \tau(b_A)$ for some function κ and constant b_A (where we take $\kappa(M_t^X) > B(M_t^X)$ to imply never quitting at any $X_t \in (M_t^X, B(M_t^X))$). However, since $B(M_t^X) \leq \underline{B}_\ell(M_t^X)$, it is never optimal ℓ to

quit at $\kappa(M_t^X) > B(M_t^X)$. Therefore, the optimal quitting rule must be equivalent to $\tau' = \tau(b_A)$ for some b_A . □

G.2 Proof of Lemma 10

Proof. As before, we define a relaxed version of AM^ℓ to be

$$[RAM_N^\ell]: \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_\ell]$$

subject to $\forall X_i \in \mathcal{T}_N$

$$RDP_\ell(X_i): \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} (d_\tau \mathbf{1}(\tau < \tau(X_i)) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_\ell] \leq \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_\ell]$$

$$PK_\ell(V_\ell): \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_\ell] \geq V_\ell + \frac{c}{r}.$$

We claim that the value of RAM_N^ℓ is (for some value γ) equivalent to

$$[H_N^{\ell, \gamma}]: \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} (1 - \gamma) - \gamma \frac{c}{r}) | Z_\ell^1]$$

subject to $RDP_\ell(X_i) \forall X_i \in \mathcal{T}_N, X_i < X^1$.

By Lemma 16, we know that the value of RAM_N^ℓ is equal to

$$\begin{aligned} & \inf_{(\gamma, \lambda_\ell) \in \mathbb{R}_-^{N+2}} \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} (1 - \gamma) + (-a\gamma - 1)}{1 + e^{Z_\tau}} - \gamma \frac{c}{r}) | Z_\ell] + \gamma V_\ell \\ & + \sum_{i=0}^N \lambda_\ell(X_i) (\mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} (d_\tau \mathbf{1}(\tau < \tau(X_i)) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_\ell] - \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_\ell]) \\ & = \inf_{\gamma \in \mathbb{R}_-} \inf_{\lambda_\ell \in \mathbb{R}_-^{N+1}} \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} (1 - \gamma) + (-a\gamma - 1)}{1 + e^{Z_\tau}} - \gamma \frac{c}{r}) | Z_\ell] + \gamma V_\ell \\ & + \sum_{i=0}^N \lambda_\ell(X_i) (\mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} (d_\tau \mathbf{1}(\tau < \tau(X_i)) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_\ell] - \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_\ell]) \\ & = \inf_{\gamma \in \mathbb{R}_-} H_N^{\ell, -\gamma} + \gamma V_\ell. \end{aligned}$$

We can apply Theorem 4 to $H_N^{\ell, \gamma}$, yielding the desired form of the optimal mechanism. □

G.3 Proof of Theorem 2

Proof. Suppose that $b_\ell^1 = b_\ell^*(B_\ell^1)$. Because ℓ receives zero expected utility conditional on reaching b_h^1 , ℓ '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 ℓ gets from claiming to be h is bounded above by $\max_b \tilde{V}(B_h^1, b, Z_\ell)$. Because $b_\ell^1 = b_\ell^*(B_\ell^1)$, then the utility ℓ gets from truthfully reporting his type is given by $\max_b \tilde{V}(B_\ell^1, b, Z_\ell)$. By Assumption 1 and the fact that \tilde{V} is single-peaked in B , we know that for $b > b^{SM}$ we have $\max_b \tilde{V}(B_h^1, b, Z_\ell) < \max_b \tilde{V}(B_\ell^1, b, Z_\ell)$, which contradicts $DIC(Z_\ell, Z_h)$ binding.

Now suppose that $b_\ell^1 \leq b^*(-\frac{\sigma}{\phi}Z_\ell)$. Again let $B_h^1 > B_\ell^1$. Now consider the alternative mechanism in which ℓ 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}_ℓ^1). Because this is not optimal, we must have (for $b_{\ell,Z}^1 = Z_\ell + \frac{\phi}{\sigma}b_\ell^1$ and $B_{\ell,Z}^1 = Z_\ell + \frac{\phi}{\sigma}B_\ell^1$ and similar for $\tilde{B}_{\ell,Z}^1, \tilde{b}_{\ell,Z}^1$)

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

Thus the probability of approval when $\theta = H$ in the stationary regime is higher for ℓ when reporting ℓ 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 $\theta = L$ in the stationary regime is higher for ℓ when reporting ℓ rather than h . But since the expected costs are lower in the stationary regime for ℓ 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 $b_\ell^1 < b_h^1$ in order to not violate $DIC(Z_\ell, Z_h)$; if we had $b_h^1 \leq b_\ell^1$, then ℓ could choose to report h and quit if the evidence reaches b_ℓ^1 . This deviation is identical to lowering the approval standard for ℓ , which strictly increases utility for ℓ .

Finally, we argue that if $b_h^1 \neq b_\ell^1$, then it must be that h is getting his symmetric information mechanism. By the arguments in Lemmas 7 and 8, R only changes the approval threshold when a constraint is binding. For $X_i < b_\ell^1$, it is easy to see that $RDIC$ constraints are slack and thus the approval threshold must be constant until b_h^{SM} . Since there is no change in the approval threshold at $\tau(b_h^{SM})$, this will be equal to h 's symmetric mechanism. □

G.4 Proof of Proposition 6

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 ($\mathbb{P}(d_\tau = 1|H, h)$) 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 reject with negligible probability as $b_h^*(-\frac{\sigma}{\phi}Z_h) \rightarrow -\infty$. For a fixed Z_ℓ , we will have $\mathbb{P}(d_\tau = 1|H, h) > \mathbb{P}(d_\tau = 1|H, \ell)$.

Let $\rho_i = \mathbb{E}[e^{-r\tau^i}|\theta = H]$ be the expected discounted time till τ when $\theta = H$. As $Z_h \rightarrow \infty$, we have that $\mathbb{E}[e^{-r\tau}(d_\tau + \frac{c}{r})|Z_h] \approx \rho_i(1 + \frac{c}{r})$ since $\mathbb{P}(d_\tau = 1|H) \rightarrow 1$. Similarly, as $Z_h \rightarrow \infty$, $\rho_i \approx J(\tau, d_\tau, Z_h)$.

Suppose that we solve the optimal mechanism dropping $DIC(Z_h, Z_\ell)$. Since R could always offer the ℓ 's mechanism to h , we must have $\rho_h \geq \rho_\ell$. This will imply that $DIC(Z_h, Z_\ell)$ is slack.

Finally, we argue that the fast-track mechanism is optimal. Take $\pi_h = 1$ and suppose that ℓ 's utility V_ℓ is strictly positive and, for the sake of contradiction, that $b_h^1 = b^*(B_h^1)$. Let $J(b_h^1)$ be the continuation value to R of the optimal mechanism at b_h^1 . We will show that $b_h^1 = b^*(B_h^1)$ is suboptimal by constructing a two stage mechanism determined by some b^1 with initial approval threshold $B(b^1; V_\ell)$ (equal to the approval threshold which gives ℓ utility V_ℓ when ℓ 's continuation value at b^1 is zero) which jumps up to $\underline{B}_\ell(b^1)$ at b^1 , where B^1 is set so that the expected utility to ℓ from (b^1, B^1) is equal to V_ℓ . The utility of this mechanism to R is given by

$$\Psi(B'(b^1; V_\ell), b^1, X_0) + \psi(B'(b^1; V_\ell), X_0, b^1, X_0)[\Psi(\underline{B}_\ell(b^1), b_h^1, b^1) + \psi(\underline{B}_\ell(b^1), b_h^1, b^1)J(b_h^1)]. \quad (6)$$

Let $\Psi^1 := \Psi(\underline{B}_\ell(b^1), b^1, X_0)$ and $\Psi^2 := \Psi(\underline{B}_\ell(b^1), b_h^1, b^1)$, which similar notation for ψ . Because $b_h^1 = b^*(B_h^1)$, we know that $B'(b^1; V_\ell)|_{b^1=b_h^1} = 0$. If we take the derivative of equation 6 with respect to b^1 , we have

$$\begin{aligned} & \Psi_b^1 + \psi_b^1[\Psi^2 + \psi^2 J(b_h^1)] + \psi^1[\Psi_x^2 + \psi_x^2 J(b_h^1)] \\ & + B'(b^1; V_\ell) \left(\Psi_B^1 + \psi_B^1[\Psi^2 + \psi^2 J(b_h^1)] \right). \end{aligned} \quad (7)$$

where Ψ_x and ψ_x are the derivatives with respect to X_0 . When evaluated at $b^1 = b_h^1$, straightforward calculations show that the above expression is equal to zero. Taking the derivative of 7 and using the fact that $B'(b^1; V_\ell) = 1$, we get

$$\begin{aligned} & \Psi_{bb}^1 + \psi_{bb}^1[\Psi^2 + \psi^2 J(b_h^1)] + 2\psi_b^1[\Psi_x^2 + \psi_x^2 J(b_h^1)] + \psi^1[\Psi_{xx}^2 + \psi_{xx}^2 J(b_h^1)] \\ & + B'(b^1; V_\ell)''(b^1) \left(\Psi_B^1 + \psi_B^1[\Psi^2 + \psi^2 J(b_h^1)] \right). \end{aligned} \quad (8)$$

Note that $\Psi_B^1 + \psi_B^1[\Psi^2 + \psi^2 J(b_h^1)] < 0$ (otherwise it would be profitable to increase B_h^1 , keeping b_h^1 fixed, which would still be incentive compatible since ℓ 's utility is strictly decreasing in B) and that $B''(b^1; V_\ell) < 0$ (since increasing b_h^1 decreases ℓ 's utility and therefore B must decrease to keep ℓ 's utility fixed at V_ℓ). Taking $X_0 \rightarrow B_h^1$, we have that $\Psi_{bb}, \psi_{bb}, \psi_b \rightarrow 0$; therefore it must be that increasing b^1 will increase R 's utility will keeping ℓ 's utility fixed. Since this does better than using $b_h^1 = b_\ell^*(B_h^1)$, the fast-track mechanism must be strictly optimal. \square

G.5 Proof of Lemma 11

Proof. Since h will always have a higher belief than ℓ , we can conjecture that h will never quit as long ℓ still finds it optimal to experiment. This leads us to define a relaxed problem in which we assume that h will not quit early:

$$\begin{aligned} [RAM_\ell^{BIND}] : & \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_\ell] \\ & \text{subject to } PK_\ell(V_\ell), RDP_\ell(X_i) \forall X_i \in \mathcal{T}_N, \\ RDIC_h(V_h) : & \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_h] \leq V_h + \frac{c}{r}. \end{aligned}$$

Using the arguments of Lemma 10 and a similar change of expectation as in Lemma 7, we can see that the solution to RAM_ℓ^{BIND} is a sequence of static stopping thresholds $\{B^i\}_{m=1}^M$ (where M is the number of binding RDP constraints; if $M = 0$, there is a single stopping threshold B^1) and which moves from B^i to B^{i+1} when the i th binding RDP constraint is reached, as well as a rejection threshold b_r .

We need to argue that R approves whenever X_t reaches the stopping threshold B_i . It is clear that in order for RDP to be satisfied when the threshold is B_M , R cannot be rejecting at both B_M and b_r . Suppose that R approves at b_r (with beliefs Z_r). Then for approval to be optimal, it must be that

$$\frac{e^{Z_r} - 1}{1 + e^{Z_r}} - \left(\left(\sum_{j=0}^{|\mathcal{B}_N|} \lambda_\ell(X^j) + \lambda_\ell \right) \frac{e^{Z_r} + a}{1 + e^{Z_r}} - \lambda_h \frac{e^\Delta e^{Z_r} + a}{1 + e^{Z_r}} \right) > 0,$$

where $\lambda_\ell, \lambda_h, \lambda(X^j)$ are the Lagrange multipliers in our dual problem for RAM_ℓ^{BIND} associated with the $PK(V_\ell)$, $RDIC_h(V_h)$ and $RDP_\ell(X_i)$ constraints respectively. Because $Z_r < Z_\ell < 0$, this implies that $\frac{e^{Z_r}-1}{1+e^{Z_r}} < 0$ and thus

$$\left(\sum_{j=0}^{|\mathcal{B}_N|} \lambda_\ell(X^j) + \lambda_\ell\right) \frac{e^{Z_r} + a}{1 + e^{Z_r}} - \lambda_h \frac{e^\Delta e^{Z_r} + a}{1 + e^{Z_r}} < 0.$$

Suppose that R rejects at B_M and that $M \geq 1$. This means that at X^M , R is only approving at belief b_r which give negative utility and hence would be better off rejecting immediately (which can only weaken h 's DIC constraint). Now suppose that no RDP constraints are binding. Then the optimal mechanism stops at $\tau_{\geq}(B) \wedge \tau(b)$ for some (b, B) . If there is approval at B and rejection at b , we are done. Suppose instead that R rejects at B and approves at b . Because approval at b yields negative utility to R , R would be better off approving immediately with probability equal to ℓ 's expected utility of

$$\mathbb{E}[e^{-r\tau(b)}(\mathbb{1}(\tau(b) \leq \tau_{\geq}(B)) \frac{e^{Z_r} + a}{1 + e^{Z_r}} + \frac{c}{r}) | Z_\ell] - \frac{c}{r}.$$

Because $DIC(Z_\ell, Z_h)$ holds, this satisfies $DIC(Z_h, Z_\ell)$ (since ℓ will receive lower expected utility in h 's mechanism than h will). Moreover, this will yield a higher utility for R . To see this, let Q_b, Q_B be defined as

$$\begin{aligned} Q_b &:= \mathbb{E}[e^{-r(\tau_{\geq}(B) \wedge \tau(b))} \mathbb{1}(\tau(b) < \tau_{\geq}(B)) | Z_\ell], \\ Q_B &:= \mathbb{E}[e^{-r(\tau_{\geq}(B) \wedge \tau(b))} \mathbb{1}(\tau(b) > \tau_{\geq}(B)) | Z_\ell] \end{aligned}$$

Let $Z_r = Z_\ell + \frac{\phi}{\sigma} b$. Then we the utility of stopping immediately (divided by $1 + \frac{c}{r}$) is given by

$$\frac{e^{Z_\ell} - 1}{1 + e^{Z_\ell}} \left(Q_b \frac{e^{Z_r} + \frac{a + \frac{c}{r}}{1 + \frac{c}{r}}}{1 + e^{Z_r}} + \frac{c}{r + c} (Q_B - 1) \right) > \frac{e^{Z_\ell} - 1}{1 + e^{Z_\ell}} Q_b \frac{e^{Z_r} + \frac{a + \frac{c}{r}}{1 + \frac{c}{r}}}{1 + e^{Z_r}}$$

while the utility from the original mechanism was

$$Q_b \frac{e^{Z_r} - 1}{1 + e^{Z_r}}$$

Thus the utility of stopping immediately is higher if

$$\frac{e^{Z_\ell} - 1}{1 + e^{Z_\ell}} \frac{e^{Z_r} + \frac{a + \frac{c}{r}}{1 + \frac{c}{r}}}{1 + e^{Z_r}} \geq \frac{e^{Z_r} - 1}{1 + e^{Z_r}}$$

which holds since $\frac{a + \frac{c}{r}}{1 + \frac{c}{r}} < 1$ and $\frac{e^{Z_r} - 1}{1 + e^{Z_r}} < 0$.

Finally, we need to rule out approval at both B and b . Our Lagrangian for RAM_ℓ^{BIND} when the final binding RDP constraint has been reached is

$$\begin{aligned} & \mathbb{E}[e^{-r\tau} (d_\tau [\frac{e^{Z_r} - 1}{1 + e^{Z_r}} - \left(\left(\sum_{j=0}^{|\mathcal{B}_N|} \lambda_\ell(X^j) + \lambda_\ell \right) \frac{e^{Z_r} + a}{1 + e^{Z_r}} - \lambda_h \frac{e^\Delta e^{Z_r} + a}{1 + e^{Z_r}} \right)]] \\ & - \frac{c}{r} \left(\sum_{j=0}^{|\mathcal{B}_N|} \lambda_\ell(X^j) + \lambda_\ell - \lambda_h \frac{e^\Delta e^{Z_r} + 1}{1 + e^{Z_r}} \right) | Z_t^M]. \end{aligned}$$

Approval at b_r (corresponding to belief Z_r) implies that

$$\frac{e^{Z_r} - 1}{1 + e^{Z_r}} - \left(\left(\sum_{j=0}^{|\mathcal{B}_N|} \lambda_\ell(X^j) + \lambda_\ell \right) \frac{e^{Z_r} + a}{1 + e^{Z_r}} - \lambda_h \frac{e^\Delta e^{Z_r} + a}{1 + e^{Z_r}} \right) > 0$$

Because $\frac{e^{Z_r} - 1}{1 + e^{Z_r}} < 0$, we must have $\sum_{j=0}^{|\mathcal{B}_N|} \lambda_\ell(X^j) + \lambda_\ell - \lambda_h \frac{e^\Delta e^{Z_r} + 1}{1 + e^{Z_r}} < 0$. Therefore by Lemma 17, we know that waiting is strictly suboptimal and R could do better by approving immediately, a contradiction. \square

G.6 Proof of Lemma 12

Suppose that $DIC(Z_h, Z_\ell)$ is binding and let V_h be the utility promised to h under the optimal mechanism. The design problem for h 's optimal mechanism is similar to that in Lemma 18 expect for the addition of a promise keeping constraint for h :

$$\begin{aligned} [AM^h]: & \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_r} - 1}{1 + e^{Z_r}} | Z_h] \\ & \text{subject to } DP(Z_h), DIC(Z_\ell, Z_h, V_\ell) \\ PK(V_h): & \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_r} + a}{1 + e^{Z_r}} + \frac{c}{r}) | Z_h] \geq V_h + \frac{c}{r}. \end{aligned}$$

Proof. We define a relaxed problem, dropping quitting rules except for threshold quitting rules in \mathcal{T}_N as

$$[RAM_N^h]: \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_h]$$

subject to $\forall X_i \in \mathcal{T}_N$

$$RDP_h(X_i): \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} (d_\tau(X_i) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_h] \leq \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_h],$$

$$RDIC_\ell(X_i): \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} (d_\tau(X_i) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_\ell] \leq V_\ell + \frac{c}{r},$$

$$PK(V_h): \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_h] \geq V_h + \frac{c}{r}.$$

By the same arguments as in Lemma 10, $\exists \gamma < 0$ such that $RAM_h = H^h(\gamma, V_h)$ where

$$[H_N^h(\gamma, V_h)]: \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} (d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} - \gamma(d_\tau + \frac{c}{r})) | Z_h] + \gamma V_h$$

subject to $RDP_h(X_i) \forall X_i \in \mathcal{T}_N$,

$$RDIC(Z_\ell, Z_h): \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} (d_\tau(X_i) \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r}) | Z_\ell] \leq V_\ell + \frac{c}{r}.$$

By applying Lemmas 18 - 20, we get our desired for the optimal stopping mechanism (where B_h^2 is the initial level set for $SM^h(\gamma)$) as $N \rightarrow \infty$.

□

G.7 Proof of Proposition 7

Proof. Let $j_i(X, M^X)$ be given by

$$j_i(X, M^X) = \Psi(B_{i,Z}(M), M, Z_i) \frac{e^{Z_i} - e^{Z_i - B_{i,Z}(M)}}{1 + e^{Z_i}} + \psi(B_{i,Z}(M), M, Z_i) \frac{e^{Z_i} + e^{Z_i - M}}{1 + e^{Z_i}} j_i(M),$$

which is the solution to the Dirichlet problem (dropping i subscripts)

$$\mathbb{L}_X j(X, M^X) = rj(X, M)$$

$$j(B(M), M) = B(M)$$

$$j(M, M) = j(M)$$

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

$$j'(M^X) = \frac{\phi}{\sigma} \left(j(M^X) \left[\frac{1}{1 + e^{M^Z}} - \psi_B B'_Z(M) - \psi_b \right] - \frac{e^{M^Z} - e^{M^Z - B'_Z(M^Z)}}{1 + e^{M^Z}} [\Psi_B B'_Z(M^Z) + \Psi_b] \right), \quad (9)$$

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

We first argue that there is a unique solution to the differential equation for $j'(M)$. Lipschitz continuity of the RHS of equation 9 is clear and continuity in M follows from the continuity of $\underline{B}(M), \underline{B}'(M)$. Therefore, the Picard-Lindelof Theorem implies a unique solution.

Now we want to argue that $j(X_0, M_0^X) = \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{B_Z(M_\tau^Z)} - 1}{1 + e^{B_Z(M_\tau^Z)}} | Z_0, M_0]$ where $\tau = \inf\{t : X_t \geq B(M_t^X)\} \wedge \tau(-\frac{\phi}{\sigma} Z_i)$. By applying Ito's Lemma to $j(X, M^X)$, we have

$$\begin{aligned} e^{-rt} j(X_t, M_t^X) &= j(Z_0, M_0^X) + \int_0^t e^{-rs} \left[\sigma \frac{\partial j(X_s, M_s^X)}{\partial X} dB_s + \frac{\partial j(X_s, M_s^X)}{\partial X_s} \mu(X_s) ds \right. \\ &\quad \left. + \frac{\partial^2 j(X_s, M_s^X)}{\partial X^2} \frac{\sigma^2}{2} ds - rj(X_s, M_s^X) ds + \frac{\partial j(X_s, M_s^X)}{\partial M_s^X} dM_s^X \right] \\ &= j(X_0, M_0^X) + S_t + \int_0^t e^{-rs} [\mathbb{L}_X j(X_s, M_s^X) - rj(X_s, M_s^X)] ds, \end{aligned}$$

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

$$S_t = \int_0^t e^{-rs} \frac{\partial j(X_s, M_s^X)}{\partial X_s} dB_s,$$

which is a continuous local martingale.

We now note that $\mathbb{L}_X j(Z_s, M_s^X) - rj(X_s, M_s^X) = 0$ for all $X_s \in (M_s^X, B(M_s^X))$. Therefore, we can reduce the above equation for $e^{-rt} j(X_t, M_t^X)$ to

$$e^{-r\tau} j(X_\tau, M_\tau^X) = j(X_0, M_0^X) + S_\tau,$$

When the process is stopped, the value $j(X_\tau, M_\tau^X)$ is always equal to $\mathbb{1}(X_\tau \geq B(M_t^X)) \frac{e^{B_Z(M_\tau^Z)} - 1}{1 + e^{B_Z(M_\tau^Z)}}$. Therefore, we have that

$$e^{-r\tau} \mathbb{1}(X_\tau \geq B^X(M_t^X)) \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} = e^{-r\tau} j(X_\tau, M_\tau^X) = j(X_0, M_0^X) + S_\tau.$$

Taking expectations of both sides, we have

$$\mathbb{E}[e^{-r\tau} \mathbf{1}(X_\tau = B(M_t^X)) \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_0] = j(X_0, M_0^X) + \mathbb{E}[S_t | Z_0].$$

It follows from Doob's optimal sampling theorem that $\mathbb{E}[S_t | Z_0, M_0^X] = 0$. Noting that $\mathbb{E}[e^{-r\tau} d_\tau \frac{e^{B_Z(M_\tau^X)} - 1}{1 + e^{B_Z(M_\tau^X)}} | Z_0, M_0^X] = \mathbb{E}[e^{-r\tau} \mathbf{1}(X_\tau \geq B(M_t^X)) \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_0]$, we can conclude that $j(X_0, M_0^X) = \mathbb{E}[e^{-r\tau} d_\tau \frac{e^{B_Z(M_\tau^X)} - 1}{1 + e^{B_Z(M_\tau^X)}} | Z_0, M_0^X]$

□

G.8 Proof of Proposition 8

Proof. Consider the case of $c = 0$. Let $\alpha_i = \mathbb{E}[e^{-r\tau_i} \mathbf{1}(d_\tau^i = 1) | \theta = H]$ and $\beta := \mathbb{E}[e^{-r\tau_i} \mathbf{1}(d_\tau^i = 1) | \theta = L]$ be the discounted probability of approval for type Z_i when $\theta = H$ and $\theta = L$ (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 τ_h, τ_ℓ , 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 π_ℓ , we get $\alpha_\ell \geq \alpha_h$. Doing the same with π_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. This mechanism must maximize R 's utility subject to offering both types the same mechanism, which corresponds to R 's optimal solution with prior $\mathbb{P}(Z_h) \pi_h + \mathbb{P}(Z_\ell) \pi_\ell$. A straightforward application of the Theorem of the Maximum yields the statement of the proposition.

□

G.9 Proof of Proposition 9

Proof. Suppose that $\pi_h = 1$ and $\pi_\ell = 0$. We first 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 Proposition 8, 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. If $\mathbb{P}(Z_\ell) > \mathbb{P}(Z_h)$, then R 's utility will converge to zero.

Next, we want to show that for c large enough, we can devise a (suboptimal) approval rule such ℓ 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, for some small dt ,

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

We can find a large c and small dt such that this is solved by $X_c < 0$ (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 $\frac{\mathbb{P}(Z_h)}{2} > 0$. Therefore, the value for high c is higher than the value for low c . □

G.10 Proof of Proposition 10

Proof. Suppose that A knows the state perfectly and R uses the symmetric mechanism for $\pi = \mathbb{P}(Z_h)$. Then h will never have an incentive to quit early, since \underline{B} is increasing in Z . Moreover, by the same argument, ℓ will choose to quit early. Therefore, let us define (τ^h, d_τ^h) to be the same as the symmetric mechanism and (τ^ℓ, d_τ^ℓ) to be the same as the symmetric mechanism except that it rejects whenever ℓ would find it optimal to quit. This menu of mechanisms is clearly incentive compatible and using the (τ^h, d_τ^h) will yield the same distribution of approval times as the symmetric mechanism if $\theta = H$ is present, but, (τ^ℓ, d_τ^ℓ) will less approval than the symmetric mechanism if $\theta = L$. Therefore, we conclude that the value of the asymmetric mechanism using $(\tau^i, d_\tau^i)_{i=h,\ell}$ when A is informed of the state is higher than the value of the symmetric information mechanism when A has no private information since

$$\begin{aligned} \mathbb{E}[e^{-r\tau^{SM}} \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} | Z_0] &= \mathbb{P}(Z_h) \mathbb{E}[e^{-r\tau^{SM}} | \theta = H] - \mathbb{P}(Z_\ell) \mathbb{E}[e^{-r\tau^{SM}} | \theta = L] \\ &< \mathbb{P}(Z_h) \mathbb{E}[e^{-r\tau^h} | \theta = H] - \mathbb{P}(Z_\ell) \mathbb{E}[e^{-r\tau^\ell} | \theta = L] \end{aligned}$$

□

H Transfers

H.1 Proof of Proposition 11

Proof. We again define a relaxed problem by dropping keeping on threshold quitting rules. Our Lagrangian is then

$$\begin{aligned} & \mathbb{E}\left[e^{-r\tau} d_\tau \frac{e^{Z_\tau} - 1}{1 + e^{Z_\tau}} - \int_0^\tau e^{-rt} w_t dt \right. \\ & - \sum_{i=1}^N \lambda(X_i) \left(e^{-r\tau} \left(d_\tau \frac{e^{Z_\tau} + a}{1 + e^{Z_\tau}} + \frac{c}{r} \right) + \int_0^\tau e^{-r\tau} u(w_t) dt \right. \\ & \left. \left. - [e^{-r(\tau \wedge \tau(X_i))} (d_\tau(X_i) \frac{e^{Z_\tau} + ae^{Z_\tau}}{1 + e^{Z_\tau}} + \frac{c}{r}) + \int_0^{\tau \wedge \tau(X_i)} e^{-rt} u(w_t) dt] \right) \right] \end{aligned}$$

The first order condition for w_t is given by

$$1 = - \sum_{i=1}^N \lambda(X_i) \mathbf{1}(M_t^X \leq X_i) u'(w_t).$$

Plugging this into the Lagrangian, the solution to the relaxed problem will again be a threshold strategy where the threshold changes in M_t^X . The sufficiency of the relaxed problem follows from the arguments given for the case without transfers. \square

I General Markov Process

We now move to the model of Section 7.2 and begin by proving a useful Lemma on the agent's value function \widehat{V} .

Lemma 22. *Under Assumptions 2, for each (b, B) such that $B > X_{my}$, $\widehat{V}(B, b, X)$ satisfies strict single-crossing of $f(X, 0)$.*

Proof. Since $B > X_{my}$ and $f(B, 1) > f(B, 0)$, we know that $\widehat{V}(B, b, b) = f(b, 0)$ and $\widehat{V}(B, b, B) > f(B, 0)$. For the sake of contradiction. Suppose that $\exists X$ such that for some $X' < X''$ such that $X \in (X', X'')$ and we had

$$\widehat{V}(B, b, X') - f(X', 0) = \widehat{V}(B, b, X'') - f(X'', 0) = 0 < \widehat{V}(B, b, X) - f(X, 0).$$

We can rewrite $\widehat{V}(B, b, X)$ as

$$\begin{aligned}
\widehat{V}(B, b, X) &= \mathbb{E}[e^{-r\tau(X')} \widehat{V}(B, b, X') \mathbf{1}(\tau(X') < \tau(X'')) | X] + \mathbb{E}[e^{-r\tau(X'')} \widehat{V}(B, b, X'') \mathbf{1}(\tau(X'') < \tau(X') | X] \\
&= \mathbb{E}[e^{-r\tau(X')} f(X', 0) \mathbf{1}(\tau(X') < \tau(X'')) | X] + \mathbb{E}[e^{-r\tau(X'')} f(X'', 0) \mathbf{1}(\tau(X'') < \tau(X') | X] \\
&< f(X, 0)
\end{aligned}$$

where the last line follows from the fact that pure delay is suboptimal. This contradicts the assumption that $\widehat{V}(B, b, X) - f(X, 0) > 0$ and shows that we must have strict single crossing. \square

Before considering the problem *GSM*, we first consider the following constrained optimal stopping problem which is useful in the proof of Theorem 4 as well as Section 5.2:

$$\begin{aligned}
&\sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} g(X_\tau, d_\tau) | X_0] \\
&\text{subject to } \forall X_i \in \mathcal{T}_N \text{ such that } X_i \leq X_0 \text{ and } Y_m \in \mathcal{Y}_M \\
&RDP(X_i) : \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} f(X_{\tau \wedge \tau(X_i)}, d_\tau(X_i)) | X_0] \leq \mathbb{E}[e^{-r\tau} f(X_\tau, d_\tau) | X_0] \\
&RPK(X_i, Y_m) : \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} w_m(X_{\tau \wedge \tau(X_i)}, d_\tau(X_i)) | X_0] \leq Y_m.
\end{aligned} \tag{10}$$

where $\mathcal{T}_N \in \mathbb{R}^{N+1}$ and $\mathcal{Y}_M \in \mathbb{R}^{M+1}$. We can think of g as the utility function for R and f as the utility function for A , while w_m are a set of promise keeping constraints. We will assume that g, f, w are all bounded in X, d and that problem 10 satisfies all the conditions of Lemma 16. Let $H_N(X^1)$ be defined as the optimal mechanism without *RPK* constraints which makes sure that A weakly wants to continue experimenting.

$$\begin{aligned}
[H_N(X_t)] : &\sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} g(X_\tau, d_\tau) | X_t] \\
&\text{subject to } \forall X_i \in \mathcal{T}_N \cup \{X_t\} \text{ such that } X_i \leq X_t \\
&RDP(X_i) : \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} f(X_{\tau \wedge \tau(X_i)}, d_\tau(X_i)) | X_t] \leq \mathbb{E}[e^{-r\tau} f(X_\tau, d_\tau) | X_t] \\
&PK(X_t) : \mathbb{E}[e^{-r\tau} f(X_\tau, d_\tau) | X_t] \geq f(X_t, 0).
\end{aligned}$$

Let us define the set of *RDP* constraints which are binding when using the optimal (τ, d_τ) for *RSM* _{N} as

$$\mathcal{B}_N = \{X_i \in \mathcal{T}_N : RDP(X_i) \text{ or, for some } m \text{ } RPK(X_i, Y_m) \text{ is binding}\}.$$

We will assume write $\mathcal{B}_N = \{X^1, \dots, X^{|\mathcal{B}_N|}\}$ which are ordered from largest to smallest.

Lemma 23. *Let (τ, d_τ) be the solution to 10. Then the optimal stopping rule is a static threshold until X^1 is reached for the first time; if there are no RPK constraints, then the continuation mechanism at $\tau(X^1)$ solves $H_N(X^1)$ where $PK(X^1)$ binds.*

Proof. By Lemma 16, there exists a set of multipliers $(\lambda(X_0), \dots, \lambda(X_N)) \in \mathbb{R}_-^{N+1}$ and $(\gamma(K_0, X_0), \dots, \gamma(K_M, X_N)) \in \mathbb{R}_-^{(M+1)(N+1)}$ such that the solution to $\sup_{(\tau, d_\tau) \in \Delta_C} \mathbb{E}[e^{-r\tau} g(X_\tau, d_\tau) | X_0]$ is also a solution to

$$\begin{aligned} & \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} g(X_\tau, d_\tau) | X_0] \\ & - \sum_{j=1}^{|\mathcal{B}_N|} [\lambda(X^j) (\mathbb{E}[e^{-r(\tau \wedge \tau(X^j))} f(X_{\tau \wedge \tau(X^j)}, d_\tau(X^j)) | X_0] - \mathbb{E}[e^{-r\tau} f(X_\tau, d_\tau) | X_0]) \\ & + \sum_{m=0}^M \gamma(K_m, X^j) \mathbb{E}[e^{-r(\tau \wedge \tau(X^j))} w_m(X_{\tau \wedge \tau(X^j)}, d_\tau(X^j)) | X_0]]. \end{aligned}$$

We will argue that as long as X_t has not reached X^1 , the optimal policy must be a threshold policy where the process stops if $X_t \geq B^1$ for some $B^1 \geq 0$. To see this, define the value of the optimal stopping rule after crossing X^1 as

$$\begin{aligned} k^R(X^1) & := \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} (g(X_\tau, d_\tau) + \lambda(X^1) f(X_\tau, d_\tau)) \tag{11} \\ & + \sum_{j=2}^{|\mathcal{B}_N|} [\lambda(X^j) (e^{-r(\tau \wedge \tau[h_\tau(X^1)](X^j))} f(X_{\tau \wedge \tau(X^j)}, d_\tau(X^j)) - e^{-r\tau} f(X_\tau, d_\tau)) \\ & + \sum_{m=0}^M \gamma(K_m, X^j) e^{-r(\tau \wedge \tau[h_\tau(X^1)](X^j))} w_m(X_{\tau \wedge \tau(X^j)}, d_\tau(X^j))] | X^1 \\ & + \lambda(X^1) f(X^1, 0) + \sum_{m=1}^M \gamma(K_j, X^1) w_m(X^1, 0). \end{aligned}$$

Note that at X^1 if $\sum_{j=1}^M \gamma(K_j, X^1) w_m(X^1, 0) < 0$, then R receives a one-time loss of $\sum_{j=1}^M \gamma(K_j, X^1) w_m(X^1, 0)$ at exactly $\tau(X^1)$, which might make it optimal to stop immediately. Therefore, we allow the continuation value at $\tau(X^1)$ to be equal to the maximum of the payoff of continuing or stopping. This value $K^R(X^1)$ is given by

$$K^R(X^1) := \max\{k^R(X^1), \max_{d_\tau} g(X_\tau, d_\tau) + \sum_{j=1}^{|\mathcal{B}_N|} \sum_{m=1}^M \gamma(K_m, X^j) w_m(X^1, d_\tau)\}. \tag{12}$$

The solution to K^R is independent of the previous history up until X^1 is reached. Since K^R is finite, we can use Assumption 2 to apply Proposition 5.7 from Dayanik and Kaaratzas (2003) to conclude that an optimal stopping rule to the problem defined such that the game ends when X^1 is first reached, yielding the stopping value K^R . By the principal of optimality, we know that at time $\tau(X^1)$, the value of optimal mechanism will be equal to $K^R(X^1)$. Thus, treating $K^R(X^1)$ as the continuation value upon reaching X^1 for the first time and dropping the constants Y_m from the problem, we can rewrite equation 11 as

$$\begin{aligned} & \sup_{(\tau, d_\tau)} \mathbb{E}[\mathbf{1}(\tau \leq \tau(X^1)) [e^{-r\tau} g(X_\tau, d_\tau) + \sum_{j=1}^{|\mathcal{B}_N|} [\lambda(X^j) (e^{-r(\tau \wedge \tau(X^i))} f(X_{\tau \wedge \tau(X^j)}, d_\tau(X^j)) - e^{-r\tau} f(X_\tau, d_\tau))] \\ & + \sum_{m=0}^M \gamma(K_m, X^j) e^{-r(\tau \wedge \tau(X^j))} w_m(X_{\tau \wedge \tau(X^j)}, d_\tau(X^j))] + \mathbf{1}(\tau > \tau(X^1)) e^{-r\tau(X^1)} K^R(X^1) | X^0]. \end{aligned}$$

By applying Lemma 15, we can see that the optimal stopping policy takes a threshold form which stops when $X_t \geq B^1$ until $\tau(X^1)$. If the policy stops at $\tau(X^1)$, then we are done.

Therefore, suppose that the mechanism doesn't end at $\tau(X^1)$ and that $RDP(X^1)$ is binding. The optimal stopping rule from $\tau(X^1)$ onward is that which solves $K^R(X^1)$. Since this continuation mechanism does not depend on the history of play up until $\tau(X^1)$, it must be that the constraint expectations over f at $\tau(X^1)$

$$\mathbb{E}[e^{-r(\tau - \tau(X^1))} f(X_\tau, d_\tau) | X^1, h_{\tau(X^1)}],$$

are also independent of the history to $\tau(X^1)$ (and similarly for the other f constraint expectations). Let $\tau^1 := \tau[h_{\tau(X^1)}]$, $d^1 := d[h_{\tau(X^1)}]$ and $\tau^1(X) := \tau^1[h_{\tau(X^1)}](X)$, $d^1(X) := d^1[h_{\tau(X^1)}](X)$ be the continuation mechanism and continuation threshold quitting rule at $\tau(X^1)$. Then we can define the continuation value at $\tau(X^1)$ as

$$K^A(X^1) := \mathbb{E}[e^{-r\tau^1} f(X_{\tau^1}, d_{\tau^1}^1) | X^1].$$

We can further decompose this value into what happens before and after $\tau(X^1)$:

$$\begin{aligned} \mathbb{E}[e^{-r\tau} f(X_\tau, d_\tau) | X^1] &= \mathbb{E}[\mathbf{1}(\tau \leq \tau(X^1)) e^{-r\tau} f(X_\tau, d_\tau) | X^0] \\ &+ \mathbb{E}[\mathbf{1}(\tau > \tau(X^1)) e^{-r\tau(X^1)} K^A(X^1) | X^0], \end{aligned} \tag{13}$$

and similarly for any $X_i \leq X^1$

$$\begin{aligned}
& \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} f(X_{\tau \wedge \tau(X_i)}, d_\tau(X_i)) | X_0] \\
&= \mathbb{E}[\mathbf{1}(\tau \leq \tau(X^1)) e^{-r\tau} f(X_\tau, d_\tau) | X_0] \\
&+ \mathbb{E}[\mathbf{1}(\tau > \tau(X^1)) e^{-r\tau(X^1)} \mathbb{E}[e^{-r(\tau^1 \wedge \tau^1(X_i))} f(X_{\tau^1 \wedge \tau^1(X_i)}, d_\tau^1(X_i)) | X^1] | X_0].
\end{aligned} \tag{14}$$

Then we know by complementary slackness we know that for all X^j

$$\mathbb{E}[e^{-r\tau} f(X_\tau, d_\tau) | X_0] = \mathbb{E}[e^{-r(\tau \wedge \tau(X^j))} f(X_{\tau \wedge \tau(X^j)}, d_\tau(X^j)) | X_0].$$

Substituting in equations 13, 14, we get

$$\begin{aligned}
& \mathbb{E}[\mathbf{1}(\tau > \tau(X^1)) e^{-r\tau(X^1)} K^A(X^1) | X_0] \\
&= \mathbb{E}[\mathbf{1}(\tau > \tau(X^1)) e^{-r\tau(X^1)} \mathbb{E}[e^{-r(\tau^1 \wedge \tau^1(X^j))} f(X_{\tau^1 \wedge \tau^1(X^j)}, d_\tau^1(X^j)) | X^1] | X_0].
\end{aligned}$$

Then, using the fact that neither $K^A(X^1)$ nor

$$\mathbb{E}[e^{-r(\tau^1 \wedge \tau^1(X^j))} f(X_{\tau^1 \wedge \tau^1(X^j)}, d_\tau^1(X^j)) | X^1]$$

depend on the history up until $\tau(X^1)$ or the specific time of $\tau(X^1)$, we see that

$$\begin{aligned}
& K^A(X^1) \mathbb{E}[\mathbf{1}(\tau > \tau(X^1)) e^{-r\tau(X^1)} | X_0] \\
&= \mathbb{E}[e^{-r(\tau^1 \wedge \tau^1(X^j))} f(X_{\tau^1 \wedge \tau^1(X^j)}, d_\tau(X^j)) | X^1] \mathbb{E}[\mathbf{1}(\tau > \tau(X^1)) e^{-r\tau(X^1)} | X_0], \\
K^A(X^1) &= \mathbb{E}[e^{-r(\tau^1 \wedge \tau^1(X^j))} f(X_{\tau^1 \wedge \tau^1(X^j)}, d_\tau(X^j)) | X^1].
\end{aligned} \tag{15}$$

Evaluating this at $X^j = X^1$, we see that

$$K^A(X^1) = f(X^1, 0).$$

Thus, upon reaching X^1 , the expected continuation value to A is equal to the value of quitting at X^1 .

We now seek to show that the mechanism which solves $K^R(X^1)$ also solves $H_N(X^1)$ when there are no RPK constraints. Let (τ^H, d_τ^H) be the mechanism which solves H_N . Because it satisfies all RDP constraints, then we know that for all X_i ,

$$\lambda(X_i)\mathbb{E}[e^{-r(\tau^H \wedge \tau(X_i))} f(X_{\tau^H \wedge \tau(X_i)}, d_{\tau^H}^H(X_i)) - e^{-r\tau^H} f(X_{\tau^H}^H, d_{\tau^H}^H)|X^1] \geq 0.$$

Let $(\tau^K, d_{\tau^K}^K)$ be the continuation mechanism at $\tau(X^1)$. Since $(\tau^K, d_{\tau^K}^K)$ satisfies all *RDP* and *PK(0)* constraints, it is admissible with respect to H_N . Therefore, it must be that

$$\mathbb{E}[e^{-r\tau^H} g(X_{\tau^H}, d_{\tau^H}^H)|X^1] \geq \mathbb{E}[e^{-r\tau^K} g(X_{\tau^K}, d_{\tau^K}^K)|X^1]$$

Moreover, since $(\tau^H, d_{\tau^H}^H)$ satisfies all *RPD* constraints, we know that

$$\begin{aligned} & \mathbb{E}[\mathbf{1}(\tau < \tau(X^1))e^{-r\tau} g(X_{\tau}, d_{\tau}) + \mathbf{1}(\tau \geq \tau(X^1))e^{-r\tau(X^1)} \mathbb{E}[e^{-r\tau^K} g(X_{\tau^K}, d_{\tau^K}^K)|X^1]X_0] \\ & \geq \mathbb{E}[\mathbf{1}(\tau < \tau(X^1))e^{-r\tau} g(X_{\tau}, d_{\tau}) + \mathbf{1}(\tau \geq \tau(X^1))e^{-r\tau(X^1)} \mathbb{E}[e^{-r\tau^H} g(X_{\tau^H}, d_{\tau^H}^H)|X^1]X_0] \end{aligned}$$

which, using the fact that $\mathbb{E}[e^{-r\tau^H} g(X_{\tau^H}, d_{\tau^H}^H)|X^1]$ and $\mathbb{E}[e^{-r\tau^K} g(X_{\tau^K}, d_{\tau^K}^K)|X^1]$ are independent of the history up until $\tau(X^1)$, implies that

$$\mathbb{E}[e^{-r\tau^K} g(X_{\tau^K}, d_{\tau^K}^K)|X^1] \geq \mathbb{E}[e^{-r\tau^H} g(X_{\tau^H}, d_{\tau^H}^H)|X^1]$$

Therefore, it must be that $(\tau^K, d_{\tau^K}^K)$ solves H_N . Since complementary slackness implies that $\mathbb{E}[e^{-r\tau^K} f(X_{\tau^K}, d_{\tau^K}^K)|X^1] = f(X^1, 0)$, we know that *PK(0)* must be binding. \square

Analogously to the sketch of the proof in Section 4.3, we define a relaxed version of *GSM* to be *GRSM_N*, which is given by

$$\begin{aligned} [GRSM_N]: & \sup_{(\tau, d_{\tau})} \mathbb{E}[e^{-r\tau} g(X_{\tau}, d_{\tau})|X_0] \\ & \text{subject to } \forall X_i \in \mathcal{T}_N \\ & RDP(X_i): \mathbb{E}[e^{-r(\tau \wedge \tau(X_i))} f(X_{\tau \wedge \tau(X_i)}, d_{\tau}(X_i))|X_0] \leq \mathbb{E}[e^{-r\tau} f(X_{\tau}, d_{\tau})|X_0] \end{aligned}$$

where $\mathcal{T}_N = \{X_0 + \frac{n(X_0 - X)}{N}\}_{n=0}^N$.

Lemma 24. *Under Assumption 2, The solution to $GRSM_N$ is a static approval threshold policy until $\tau(X^1)$. Then continuation mechanism at $\tau(X^1)$ is the solution to $H_N(X^1)$ where *PK(0)* is binding.*

Proof. Let us suppose that at least one *RDP* constraint is binding (if no *RDP* constraints are binding, then the optimal solution must be a static threshold mechanism

with approval at B and rejection at b). By Lemma 23, the optimal stopping rule until time $\tau(X^1)$ is given by a static stopping rule $\tau = \inf\{t : X_t \geq B\}$ for some $B \geq X_0$. All that is left to derive is the decision rule.

First, let us suppose that R does not stop immediately at $\tau(X^1)$ and, for the sake of contradiction, that R rejected at B . By Lemma 23, the continuation value to A at $\tau(X^1)$ is equal to his outside option and thus A receives the same utility as if R rejected at X^1 : $V(\tau, d_\tau, X_0) = V(\tau_{\geq(B) \wedge \tau(X^1)}, 0, X_0)$. But this will violate $RDP(X_0)$ since by Assumption 2

$$\mathbb{E}[e^{-r(\tau_{\geq(B) \wedge \tau(b)})} f(X_{\tau_{\geq(B) \wedge \tau(b)}}, 0) | X_0] \leq f(X_0, 0).$$

Therefore, it must be that R is approving at $\tau_{\geq(B)}$ if $\tau_{\geq(B)} < \tau(X^1)$.

Now we must rule out the case in which R also stops at $X^1 < X_0$ (the constraint at X^1 will be binding if R acts at X^1). We can construct the Lagrangian corresponding to $GRSM_N$ as

$$\mathcal{L} = \sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} g(X_\tau, d_\tau) + \sum_{j=1}^{|\mathcal{B}_N|} \lambda(X^j) (e^{-r(\tau \wedge \tau(X^j))} f(X_{\tau \wedge \tau(X^j)}, d_\tau(X^j)) - e^{-r\tau} f(X_\tau, d_\tau)) | X_0].$$

Since we can now analyze the problem as a single-decision maker would, it is easy to see that

$$d_{\tau(X^1)} = 1 \iff g(X^1, 1) > g(X^0, 0) \Rightarrow g(B, 1) > g(B, 0) \iff d_{\tau_{\geq(B)}} = 1.$$

Therefore, if the optimal decision rule approves at X^1 , then it must approve at both B as well. But by Assumption 2, we know that immediate approval would be better for R than waiting to approve. Since this is also better for A , immediate approval would be admissible with respect to $GRSM_N$, contradicting the optimality of waiting until B or X^1 .

Finally, we consider the case in which R rejects at X^1 . Since rejection at X^1 is admissible with respect to $H_N(X^1)$, it must be that immediate rejection solves $H_N(X^1)$. Otherwise, if $H_N(X^1) > g(X^1, 0)$, then we could replace rejection at X^1 with the solution to $H_N(X^1)$. It is easy to see that this would satisfy all RDP constraints and would yield higher utility for R in the primal problem, contradicting the optimality of rejection at X^1 . □

Lemma 25. *Under Assumptions 2 and 3, as $N \rightarrow \infty$, the solution to $H_N(X^1)$ (when $PK(0)$ is binding) is given by $\tau = \inf\{t : X_t \geq \underline{B}(M_t^X)\} \wedge \tau(b_r)$ and $d_\tau = \mathbf{1}(X_t = \underline{B}(M_t^X))$ for some cutoff b_r .*

Proof. We now want to solve for the optimal mechanism which solves $H_N(X)$ when $PK(X)$ is binding. By applying Lemma 23 repeatedly, we see that the optimal mechanism will consist of a sequence of upper stopping thresholds $\{B^n\}_{n=1}^{|\mathcal{B}_N|}$ and one lower threshold b . Moreover, at each binding constraint $X^j > X^{|\mathcal{B}_N|}$, we must have PK binding at X^j . As the mechanism progresses, the current upper threshold depends only on the lowest binding threshold which has been reached; hence, we can write the upper threshold at time t as a step function $B_N(M_t^X)$ of the minimum of the X until time t .

Suppose that the state is $X_t = M_t^X = X^j$ and $X^j > X^{|\mathcal{B}_N|}$. By complementary slackness, we know that at X^j , we have

$$\mathbb{E}[e^{-r\tau} f(X_\tau, d_\tau) | X^j] = f(X^j, 0)$$

Therefore, in order to satisfy RDP at X^{j+1} , we must have approval at $B_N(X^{j+1})$, since, if R rejected at B_j ,

$$\begin{aligned} \mathbb{E}[e^{-r\tau} f(X_\tau, d_\tau) | X^j] &= \mathbb{E}[\mathbf{1}(\tau \leq \tau(X^{j+1})) e^{-r\tau} f(X_\tau, d_\tau) | X^j] \\ &\quad + \mathbb{E}[\mathbf{1}(\tau > \tau(X^{j+1})) e^{-r\tau(X^{j+1})} \mathbb{E}[e^{-r(\tau - \tau(X^{j+1}))} f(X_\tau, d_\tau) | X^{j+1}] | X^j] \\ &= \mathbb{E}[e^{-r(\tau(B_i) \wedge \tau(X^{j+1}))} f(X_{\tau(B_i) \wedge \tau(X^{j+1})}, 0) | X^j] \\ &< f(X^j, 0). \end{aligned}$$

a violation of RDP at X^j . Therefore, since R is approving at B and PK is binding, the threshold at X^j must be $\tilde{B}(X^{j+1}, X^j)$ (we rule out the possibility that $B_N(X^i) < \max_b B^* A(b)$ in Lemma 2).

We now turn to the relationship between X^j and X^{j+1} . Intuitively, we should not expect them to be far apart: Between time $\tau(X^j)$ and $\tau(X^{j+1})$, the optimal mechanism yields A 's outside option to A whenever $X_t = X^j$ (since, for A , the mechanism is equivalent to a static threshold with approval at B and rejection at b) and when $X_t = X^{j+1}$ (by complementary slackness). Therefore, when $X_t \in (X^{j+1}, X^j)$, the mechanism is, for A , equivalent to a static threshold mechanism (with thresholds (X^{j+1}, X^j)) giving him his outside option at both thresholds. Therefore, we should expect him to prefer taking his outside option at such X_t . Thus, if there is a constraint $X_i \in \mathcal{T}_N$ such that

$X_i \in (X^{j+1}, X^j)$, this would seem to violate $RDP(X_i)$. Formally we need to use the single-crossing of \widehat{V} to establish this.

At X^j , the agent's continuation value is equal to $\widehat{V}(\widetilde{B}(X^{j+1}, X^j), X^{j+1}, X^j)$. If there were a slack constraint $X_i \in (X^{j+1}, X^j)$, then since $\widehat{V}(\widetilde{B}(X^{j+1}, X^j), X^{j+1}, X) > f(X, 0)$ at X_{i+1} and $X \approx \widetilde{B}(X_j, X_i)$ (since f is increasing in d_τ) and $\widehat{V}(\widetilde{B}(X^{j+1}, X^j), X^{j+1}, X) = f(X, 0)$ at X^j and X^{j+1} , which implies violates strict single crossing of \widehat{V} and f . Therefore, no such X_i can exist and thus $X^{j+1} = X^j - \frac{1}{N}$.

We can also show that at the final binding quitting threshold $X^{|\mathcal{B}_N|}$, R must be rejecting. After time $\tau(X^{|\mathcal{B}_N|})$, there are no more binding constraints and the optimal stopping mechanism will solve

$$\sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau}(g(X_\tau, d_\tau) + \sum_{j=1}^{|\mathcal{B}_N|} \lambda(X^j) f(X_\tau, d_\tau) | X^{|\mathcal{B}_N|})].$$

By Lemma 15, the solution to this problem will be a pair of threshold (b, B) and, using arguments from Lemma 24, it must be that R approves at B and rejects at b (which implies that $X^i < b$ are binding).

We now turn to examine $\lim_{N \rightarrow \infty} B_N(M_t^X)$. Since $X^{|\mathcal{B}_N|} \in [\underline{X}, X^1]$, there exists a limit of $X^{|\mathcal{B}_N|}$ as $N \rightarrow \infty$. Let $\underline{X}(M_t^X) := \max\{X_i \in \mathcal{T}_N : X_i < M_t^X\}$. Since $B_N(M_t^X) = \widetilde{B}(\underline{X}(M_t^X), \underline{X}(M_t^X) - \frac{1}{N})$ and $\widetilde{B}(X_j, X_i)$ is continuously differentiable, $B_N(M_t^X)$ has uniformly bounded variation on $[\underline{X}, X_0]$ and is uniformly bounded; therefore, by Helly's Selection Theorem, it has a limit. Since RDP binding at X_i implies that it is binding at $X_{i+1} = X_i - \frac{1}{N}$, we can see that on $M_t^X > X^{|\mathcal{B}_N|}$ we have

$$\lim_{N \rightarrow \infty} B_N(M_t^X) = \widetilde{B}(\underline{X}(M_t^X), \underline{X}(M_t^X) - \frac{1}{N}) = \underline{B}(M_t^X)$$

□

Lemma 1. For $X \in [b^*(\max_b B_A^*(b)), X_{my}^A]$, there exists two function $\underline{B}_1, \underline{B}_2$ such that $\arg\max_b V(\underline{B}_1(X), b, X) = \arg\max_b V(\underline{B}_2(X), b, X) = X$ with \underline{B}_1 decreasing and \underline{B}_2 increasing.

Proof. We know that for each X , $\lim_{B \rightarrow 0} V(B, b, X) < f(X_t, 0)$; because $X > b^*(\max_b B_A^*(b))$, there exists $\underline{B}_1(X)$ which keeps A 's utility equal to zero. For the existence of $\underline{B}_2(X)$, note that for $B = X_{my}^A$, A utility of continuing is strictly negative: reaching B is equivalent to rejecting at X_{my}^A and b . Since delay is sub-optimal, it would be better to quit immediately. Therefore, by increasing B , there must exist some $\underline{B}_2(x) < \max_b B_A^*(b)$ such that A 's best utility against $\underline{B}_2(x)$ is equal to zero.

□

Lemma 2. $\lim_{N \rightarrow \infty} B_N(\cdot) = \underline{B}_1(\cdot)$.

Proof. We know that there exist two curves $\underline{B}_1(x)$ and $\underline{B}_2(x)$ such that $\underline{B}_1(x)$ is increasing and $\underline{B}_2(x)$ is decreasing, and at x , A is indifferent between experimenting and quitting against an approval threshold of $\underline{B}_i(x)$. We argue here that the optimal approval threshold must be to follow $\underline{B}_2(x)$.

Suppose that for large N the optimal choice of B switched from \underline{B}_1 to \underline{B}_2 (a similar argument will rule out a switch from \underline{B}_2 to \underline{B}_1); then $\exists X^1 > X^2$ such that $B(X^1) = \underline{B}_1(X^1) > B(X^2) = \underline{B}_2(X^2)$. Because we know that R 's preferences are single-peaked, the fact that $\underline{B}_2(X^2)$ was not chosen at X^2 implies that when faced with a lower value of $J(X^3)$ upon reaching X^3 for the first time, R 's preferred threshold must be below $\underline{B}(X^2)$. If R 's preferred threshold was above $\underline{B}_2(X^2)$, then she could her preferred threshold at $\tau(X^2)$ until $\tau(X^3)$ and preserve all DP constraints. Therefore, it must be that R 's preferred threshold is below $\underline{B}(X^2)$.

By continuity of R 's preferred threshold, we know that it will be roughly equal when evaluated at X^1 against a continuation value upon reaching X^2 of $J(X^2)$. But this implies that R would do better by using a threshold of $\underline{B}_2(X^2)$ at X^1 (which would again preserve all DP constraints by the single-peaked property of A 's preferences), a contradiction of the optimality of $\underline{B}_1(X^1)$. Therefore it cannot be that \underline{B} switches from \underline{B}_1 to \underline{B}_2 .

Finally, we argue that the optimal curve cannot be only $\underline{B}_2(x)$. Let us assume that the approval rule follows \underline{B}_2 from X_1 onwards. Then let us consider a relaxed problem in which R can choose an approval rule subject to the condition that R can approval only at levels less than B_A^* and R must reject at b_r (and we add only a participation constraint at X_1):

$$\sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r(\tau \wedge \tau(b_r))} \tilde{g}(X_{\tau \wedge \tau(b)}, d_\tau \mathbb{1}(\tau \leq \tau(b))) | X_1]$$

subject to

$$P : \mathbb{E}[e^{-r(\tau \wedge \tau(b))} f(X_{\tau \wedge \tau(b)}, d_\tau \mathbb{1}(\tau \leq \tau(b))) | X_1]$$

where $\tilde{g}(X_\tau, d_\tau) = g(X_\tau, d_\tau)$ if $X_\tau \leq B_A^*$ and is equal to $-\infty$ otherwise. Similar to our mechanism with two-sided commitment, it is straightforward to show that the optimal mechanism here is a threshold rule. Because we assume that R 's preferences are single-peaked and $B_R^*(b_r) > B_A^*(b_r)$, then it must be that the optimal threshold in this relaxed problem is $B_A^*(b_r)$. However, this threshold could be used in our original problem from X^1 onwards (since it clearly satisfies all DP constraints as it gives A his

best approval threshold subject to rejection at b_r). Therefore we conclude that \underline{B}_2 is never used in the optimal mechanism. \square

Lemma 26. *Let (τ^N, d_τ^N) be the solution to $GRSM_N$ and define $(\tau, d_\tau) = \lim_{N \rightarrow \infty} (\tau^N, d_\tau^N)$. Then (τ, d_τ) is admissible with respect to GSM .*

Proof. Our stopping rule $\tau = \inf\{t : X_t \geq B(M_t^X)\} \wedge \tau(b_r)$ is clearly a \mathcal{F}_t^X -measurable and thus is admissible if it satisfies all DP constraints. We need to verify that after any history, the continuation value for A is weakly positive. Since the mechanism (τ, d_τ) is (X_t, M_t^X) -Markov, we need only check that $\mathbb{E}[e^{-r\tau} f(X_\tau, d_\tau) | X_t, M_t^X] \geq f(X_t, 0)$.

When $M_t^X > b^*(B^1)$, then A 's continuation utility is given by $\widehat{V}(B^1, b^*(B^1), X_t)$, which is greater than $f(X_t, 0)$ by definition of $b^*(B^1)$ and the fact that $X_t > b^*(B^1)$. When $M_t^X \leq b^*(B^1)$, then A 's continuation utility is given by $\widehat{V}(\underline{B}(M_t^X), M_t^X, X_t)$, which is greater than $f(X_t, 0)$ since $X_t \geq M_t^X$. Thus, for all (X_t, M_t^X) , dynamic participation constraints are satisfied. \square

I.1 Proof of Theorem 4

Proof. Let $(\tau, d_\tau) := \lim_{N \rightarrow \infty} (\tau^N, d_\tau^N)$ where (τ^N, d_τ^N) solves $GRSM_N$. We know that the value of $GRSM_N$ (i.e., $J(\tau^N, d_\tau^N, X_0)$) is an upper bound on the value of GSM . Moreover, since the $J(\tau^N, d_\tau^N, X_0)$ is bounded above by $\sup_{(\tau, d_\tau)} \mathbb{E}[e^{-r\tau} f(X_\tau, d_\tau) | X_0]$, the dominated convergence theorem implies that $J(\tau, d_\tau, X_0) = \lim_{N \rightarrow \infty} J(\tau^N, d_\tau^N, X_0)$. Therefore, we can conclude that $J(\tau, d_\tau, X_0)$ is an upper bound on the value of GSM . Since (τ, d_τ) is admissible to GSM by Lemma 26, this implies that (τ, d_τ) is a solution to GSM .

Finally, we need to show that the threshold $B(M_t^X)$ is continuous. By Lemma 23, we know that the approval threshold (call it B^1) is constant until $\tau(X^1)$. After $\tau(X^1)$, Lemma 20 implies that $B(M_t^X) = \underline{B}(M_t^X)$. Thus, we will be done if we can show that $X^1 = b^*(B^1)$.

Clearly, we cannot have $X^1 < b^*(B^1)$, since this would violate the RDP constraints for $X_i \in (X^1, b^*(B^1))$. For the sake of contradiction, suppose that for large enough N we had $X^1 > b^*(B^1)$ where B^1 is the initial threshold for $GRSM_N$. Let $\widehat{J}(X_c)$ be the continuation value of the optimal mechanism (τ^N, d_τ^N) of $GRSM_N$ to R at some $X_c \in (X^1, X_0)$ when $M_t^X > X^1$ and let $\widetilde{J}(X_c)$ be the continuation value to R at X_c when $M_t^X \in (X^2, X^1)$. The utility to R at $t = 0$ is given by

$$\mathbb{E}[\mathbf{1}(\tau \leq \tau(X_c))e^{-r\tau}g(X_\tau, d_\tau) + \mathbf{1}(\tau > \tau(X_c))e^{-r\tau(X_c)}\widehat{J}(X_c)|X_0, M_t^X = X_0].$$

Switching to the mechanism which delivers $\widetilde{J}(X_c)$ at $\tau(X_c)$ would be admissible since $\widetilde{J}(X_c)$ satisfies all *RDP* constraints. Therefore, for this to not be optimal, it must be that $\widetilde{J}(X_c) < \widehat{J}(X_c)$. Similarly, consider the continuation value of the optimal mechanism at $\tau(X^1)$, which is given by

$$\mathbb{E}[\mathbf{1}(\tau(X^2) < \tau(X_c))e^{-r\tau(X^2)}\widetilde{J}(X^2) + \mathbf{1}(\tau(X^2) > \tau(X_c))e^{-r\tau(X_c)}\widetilde{J}(X_c)|X^1, M_t^X = X^1].$$

We can see that switching to the mechanism which delivers $\widehat{J}(X_c)$ if X_c is reached before $\tau(X^2)$ will satisfy all *RDP* at $\tau(X^1)$ (since the mechanism at $\widetilde{J}(X_c)$ satisfies all *RDP* constraints). Therefore, for this to not be optimal, we must have $\widetilde{J}(X_c) > \widehat{J}(X_c)$, a contradiction. Letting $N \rightarrow \infty$, we can conclude that $X^1 = b^*(B^1)$. □