Endogenous Experimentation in Organizations*

Germán Gieczewski† Svetlana Kosterina‡

October 2019

Abstract

We study a model of policy experimentation in organizations. Members have a common objective but differ in their prior beliefs about a risky policy. Current members collectively decide whether to experiment with the policy. Agents can enter and leave the organization freely and receive payoffs from the organization’s policy only if they are members. We show that, for a wide range of parameters, there is too much experimentation. This is due to self-selection into the organization: unsuccessful experiments drive out conservative members, leaving the organization with a radical median voter who supports continued experimentation. When successes are imperfectly informative, unlike in a setting where a single agent controls the policy, these forces can sustain equilibria in which the organization experiments more when the policy is bad than when it is good. The model is applied to decision-making in environmental and civil rights organizations, non-profits, and cooperatives.

Keywords: experimentation, dynamics, median voter, endogenous population

---

*We would like to thank Daron Acemoglu, Alessandro Bonatti, Glenn Ellison, Robert Gibbons, Faruk Gul, Matias Iaryczower, Navin Kartik, Wolfgang Pesendorfer, Caroline Thomas, Juuso Toikka, the audiences at the 2018 American Political Science Association Conference (Boston), the 2018 Columbia Political Economy Conference and the 2019 Canadian Economic Theory Conference (Montreal), as well as seminar participants at MIT and Princeton for helpful comments. All remaining errors are our own.

†Department of Politics, Princeton University.

‡Department of Politics, Princeton University.
1 Introduction

Organizations frequently face uncertainty about the quality of the policies they can pursue and must experiment with a policy to find out its quality. Organizations are diverse: the members of an organization oftentimes disagree about the merits of its policies. Moreover, the membership is fluid: some members leave, disillusioned with the policies pursued, while others join, lured by the promise of greater benefits. As the membership of an organization changes, so do the policies it pursues. This paper aims to understand the dynamics of experimentation in environments with these features— that is, in environments where the membership of an organization is in flux, its members’ beliefs are diverse, and the decision-making process within the organization is responsive to the composition of its membership.

Consider the following example to which our model applies. A citizen desiring to change the government’s policy on an environmental issue or to affect the behavior of multinational corporations that impact the environment can act independently by, for instance, writing to her elected representatives, or she can join an environmental organization, such as Greenpeace, that has access to strategies not available to a citizen acting alone, such as lobbying or demonstrations. While all members of the organization want the government to change the policy, their beliefs as to the best means of achieving this goal differ. Some support safer strategies, such as lobbying, while others prefer riskier ones, such as demonstrations. If some candidates for leadership positions in the organization advocate using risky strategies while others support safe ones, the members favoring a certain approach can influence the organization’s strategy by supporting the candidate advocating for that approach.

In our model, an organization chooses between a safe policy and a risky policy at each point in time. The safe policy yields a flow payoff known by everyone, while the risky policy yields an uncertain flow payoff. There is a continuum of agents with resources to invest. At each point in time each agent decides whether to invest her resources with the organization or to invest them outside. If she invests with the organization, she obtains a flow payoff depending on the policy of the organization. Investing the resources outside yields a guaranteed flow payoff.

All agents want to maximize their returns but hold heterogeneous prior beliefs about the quality of the risky policy. As long as agents invest with the organization,
they remain voting members of the organization and vote on the policy it pursues. We assume that the median voter of the organization (that is, the voter with the median belief) chooses the organization’s policy. Whenever the risky policy is used, the results are publicly observed by all agents.

We show that experimentation in organizations is inefficient in two ways that are novel to the literature. Our first and main result shows that, under mild conditions, there is over-experimentation relative to the social optimum in the unique equilibrium. Over-experimentation takes a particularly stark form: the organization experiments forever regardless of the outcome. Because disagreements over policy arise only from differences in prior beliefs, which should vanish as information accumulates, we might expect that in the long run the policy of the organization will converge to the one desired by most agents. However, this is not what happens: we show that learning can lead to a capture of the organization by extremists. Our second major result establishes that, for some parameters, there is an equilibrium in which the organization is more likely to experiment forever if the risky policy is bad than if it is good.

Substantively, our result of perpetual experimentation speaks to the empirical reality of the survival of radical organizations which use tactics that most of the population is pessimistic about. Failure makes such organizations more radical in that they continue trying methods that repeatedly fail to achieve success. Conversely, success makes organizations more conservative, either by rendering the methods they are using more reasonable from the point of view of the general population, or by inducing the organization to switch to more widely accepted tactics.

Going back to our example of organizations engaged in environmental activism, our result can explain the survival of radical environmental organizations such as Greenpeace that keep using controversial tactics like direct action and sabotage, occupying coal plants and confronting whaling vessels. The result can also explain the survival of the more radical environmental groups such as Animal Liberation Front and Earth Liberation Front that engage in ecoterrorism and economic sabotage in spite of the failure of these tactics to achieve the organizations’ goals.

Two forces affect the amount of experimentation in our model. On the one hand, the median member of an organization is reluctant to experiment today if she
anticipates losing control of the organization tomorrow as a result. On the other hand, if no successes are observed, as time passes only the most optimistic members remain in the organization, and these are precisely the members who most want to continue experimenting. The first force makes under-experimentation more likely, while the second force pushes the organization to over-experiment. The first major contribution of this paper is showing that the second force can dominate.

In this regard, we prove two results. The first result provides simple sufficient conditions on the parameters that guarantee perpetual experimentation for several families of densities of the agents’ prior beliefs. We provide the conditions for the uniform, power law and arbitrary densities. The second result states that if there is perpetual experimentation under a given density, the same is true under any density that MLRP-dominates it. In other words, greater optimism in the sense of the MLRP increases the likelihood of over-experimentation.

We then characterize the set of equilibria when the parameters are such that experimentation cannot go on forever. From the point of view of the initial median voter, both over-experimentation and under-experimentation are possible in these equilibria.

The rest of the paper extends the model to more general settings. The first two extensions consider alternative learning models. In the first one, agents observe perfectly informative bad news instead of good news. We characterize the unique equilibrium in this case, showing that the equilibrium is given by a finite set of stopping intervals. In this case there also exist parameters under which the organization experiments forever. The dynamics of the organization’s membership differ under bad news: the organization expands gradually instead of shrinking. A switch to the safe policy must happen after a failure is observed but can also happen even if there are no failures. There cannot be over-experimentation from the point of view of any pivotal agent but under-experimentation is possible.

In the second extension, we consider more general good news learning processes under which success is only imperfectly informative about the quality of the technology. The results from the baseline model extend to this case. Unlike in the baseline model, here for some parameters there is an equilibrium in which the organization stops experimenting with a strictly positive probability only if enough successes are
observed. A consequence of this is that, conditional on the technology being bad, the organization experiments forever, but conditional on the technology being good, the organization stops experimenting with a strictly positive probability.

Finally, we consider extensions to more general voting rules and settings in which the members’ flow payoffs depend on the size of the organization. We show that over-experimentation in equilibrium also obtains in these settings. Voting rules requiring greater supermajorities to make policy changes make over-experimentation more likely, while the effect of increasing or decreasing returns to scale is ambiguous.

The rest of the paper proceeds as follows. Section 2 discusses the applications of the model. Section 3 reviews the related literature. Section 4 introduces the baseline model. Section 5 analyzes the set of the equilibria in the baseline model. Section 6 considers other learning processes, dealing with the case of bad news and the case of imperfectly informative good news. Section 7 considers extensions of the model to general voting rules and settings where the members’ flow payoffs depend on the size of the organization.

2 Applications

Our model has a variety of applications. The applications include civil rights organizations, cooperatives, non-profits and publicly traded firms. One prominent application of our model, that of environmental organizations, has been discussed in the introduction. In a similar manner, our model applies to other organizations which aim to change government policies or public opinion, such as civil rights organizations.

Our next application is a cooperative. Here agents are individual producers who own factors of production. In the case of a dairy cooperative, for example, each member owns a cow. The agent can manufacture and sell his own dairy products independently or he can join the cooperative. If he joins, his milk will be processed at the cooperative’s plants, which benefit from economies of scale. The cooperative can choose from a range of dairy production policies, some of which are riskier than others. For instance, it can limit itself to selling mainstream products or it can instead develop a line of premium cheeses that may or may not become profitable. Dairy farmers have different beliefs about the market viability of the latter strategy.
Should this strategy be used, only the more optimistic farmers will choose to join or remain in the cooperative. The members of the cooperative decide whether to keep experimenting with the risky policy by voting.

Non-profit organizations aimed at eradicating poverty are another application of our model. Here an individual has a choice of contributing to a charity by herself or joining a non-profit organization. The non-profit organization can pursue a policy known to be effective in alleviating poverty, such as cash transfers, or a less-known policy, such as microfinance loans, that would yield greater benefits if successful. Assessing untried policies is hard, so individuals within and outside the organization disagree about the efficacy of microfinance loans.

In a publicly traded firm, agents are individuals who can invest in the firm.\(^1\) Having bought the firm’s shares, they gain voting rights which afford them a measure of control over the firm’s decisions. Shareholders influence the firm’s policy by voting in the elections of the board of directors and voting on major corporate decisions such as mergers. All shareholders have an interest in maximizing the profits of the firm. However, their beliefs as to the best means of achieving this may differ. For example, in the case of a technology company, some shareholders may believe that the firm should focus on selling desktop computers, while others may think it will do better by expanding into the mobile market.

The example of a cooperative involves private values, while the examples of non-profits and environmental organizations involve common values since desiring a change in public policy benefit from this change happening regardless of how much their actions contributed to it. We write our main model for the case of private values. In Section 7, we show that our main results survive in the common values setting.\(^2\)

---

\(^1\)Because, by making public offerings or buying back shares, firms typically also control how many agents can become shareholders, this example has features not captured by our model. We discuss it here because it is an example of considerable economic importance and the channels producing over-experimentation in our model should still apply.

\(^2\)Another way to accommodate common values in our model is to endow agents with expressive payoffs, whereby agents benefit not just from a policy change but also from having participated in the efforts that brought it about.
3 Related Literature


Keller, Rady and Cripps (2005) develop a model where multiple agents with common priors control two-armed bandits of the same type which may have breakthroughs at different times. In contrast, the present paper considers multiple agents with heterogeneous beliefs who can influence whether they are affected by the organization’s policy by entering and exiting the organization, with the members of the organization making a single collective decision in each period about whether to continue experimenting. In Keller, Rady and Cripps the amount of experimentation in equilibrium is too low due to free-riding, whereas in the present paper there are parameters under which there is over-experimentation in equilibrium.

In Strulovici (2010) a community of agents decides by voting whether to collectively experiment with a risky technology or to switch to a safe one. Agents’ payoffs from the risky technology are heterogeneous: under complete information some would prefer it to the safe technology, while others would not. As experimentation continues, agents learn about their own type as well as the types of other players.

Strulovici finds that there is too little experimentation in equilibrium because agents fear being trapped into using a technology that turns out to be bad for them. The same incentive to under-experiment is present in our model. Indeed, consider an agent who would prefer to experiment today but not tomorrow. If this agent anticipates that learning will bring an agent who will want to continue experimenting to power, then she may choose not to experiment today, lest she be forced to over-experiment or switch to her inefficient outside option.

Strulovici’s model is similar to ours in that agents collectively decide whether to experiment by voting. The model is different in that in our model agents with the same preferences start with heterogeneous beliefs which converge to a common belief, while in Strulovici’s model agents with different preferences start with common prior
beliefs and learn about payoffs over time. The ability of the agents to opt out of using the risky technology by leaving and switching to the outside option at the cost of forfeiting their voting rights is another feature that distinguishes our model.

The literature on decision-making in clubs studies dynamic policy-making in a setting where control of the club depends on policy choices but there is no uncertainty about the consequences of policies. Instead, different agents prefer different policies. The present paper shares with this strand of literature (Acemoglu, Egorov, and Sonin (2008, 2012, 2015), Roberts 2015, Bai and Lagunoff 2011) the feature that the policy chosen by the pivotal decision-maker today affects the identities of future decision-makers, leading agents to fear that myopically attractive policies may lead to a future loss of control. Most closely related is Gieczewski (2017), which, like this paper, studies a setting in which agents can choose to join an organization or stay out and are only able to influence the policy if they do join the organization. The present paper differs in considering agents with heterogeneous beliefs rather than preferences and in allowing for new information to arrive as long as the risky policy is in place.

4 The Baseline Model

Time $t \in [0, \infty)$ is continuous. There is an organization that has access to a risky policy and a safe policy. The risky policy is either good or bad. We use the notation $\theta = G, B$ for each respective scenario.

The world is populated by a continuum of agents, represented by a continuous density $f$ over $[0, 1]$. The position of an agent in the interval $[0, 1]$ indicates her beliefs: an agent $x \in [0, 1]$ has a prior belief that the risky policy is good with probability $x$. All agents discount the future at rate $\gamma$. Each agent has one unit of capital.

At every instant, each agent chooses whether to be a member of the organization. The agents can enter and leave the organization at no cost. We use $X_t \subseteq [0, 1]$ to denote the subset of the population that belongs to the organization at time $t$.\footnote{This notation rules out mixed membership strategies, but the restriction is without loss of generality.} If an agent is not a member at time $t$, she invests her capital independently and obtains a guaranteed flow payoff $s$. If she is a member, her capital is invested with the
organization and generates payoffs depending on the organization’s policy.

Whenever the organization uses the safe policy ($\pi_t = 0$), all members receive a guaranteed flow payoff $r$. When the risky policy is used ($\pi_t = 1$), its payoffs depend on the state of the world. If the risky policy is good, it succeeds according to a Poisson process with rate $b$. If the risky policy is bad, it never succeeds. Each time the risky policy succeeds, all members receive a lump-sum unit payoff. At all other times, the members receive zero while the risky policy is used.

We assume that $0 < s < r < b$. This implies that the organization’s safe policy is always preferable to investing independently. Moreover, the risky policy would be the best choice were it known to be good, but the bad risky policy is the worst of all the options.

When the risky policy is used, its successes are observed by everyone, and agents update their beliefs based on this information. By Bayes’ rule, the posterior belief of an agent with prior $x$ who has seen $k$ successes after experimenting for a length of time $\tau$ is

$$x = \frac{x}{x + (1 - x)L(k, \tau)}$$

where $L(k, \tau) = \prod_{k=0}^{k} e^{br}$. Since $L(k, \tau)$ serves as a sufficient statistic for the information observed so far, suppressing the dependence of $L(k, \tau)$ on $k$ and $\tau$, we take $L = L(k, \tau)$ to be a state variable in our model and hereafter define $p(L, x)$ as the posterior of an agent with prior $x$ given that the state variable is $L$.

Recall that, at each time $t$, a subset of the population $X_t$ belongs to the organization. We assume that the organization is using the risky policy in the beginning of the game ($\pi_0 = 1$), and at every instant $t > 0$ the median member of the organization, denoted by $m_t$, chooses whether the organization should continue to experiment at that instant.\footnote{For the median to be well-defined, we require $X_t$ to be Lebesgue measurable. It can be shown that in any equilibrium $X_t$ is an interval.} Since there is a continuum of agents, an agent obtains no value from her ability to vote and behaves as a policy-taker with respect to her membership decision. That is, she joins the organization when she prefers the expected flow payoff it offers to that of investing independently.

Because we are working in continuous time, membership and policy decisions are
made simultaneously. This necessitates imposing a restriction on the set of equilibria we consider. We are interested in equilibria that are limits of the equilibria of a game in which membership and policy decisions are made at times \( t \in \{0, \epsilon, 2\epsilon, \ldots \} \) with \( \epsilon > 0 \) small. In this discrete-time game, at each time \( t \) in \( \{0, \epsilon, 2\epsilon, \ldots \} \), first the incumbent median chooses a policy \( \pi_t \) for time \([t, t+\epsilon)\), and then all agents choose whether to be members. The agents who choose to be members at time \( t \) — and hence accrue the flow payoffs generated by policy \( \pi_t \) — are the incumbent members at time \( t + \epsilon \). The median of this set of members then chooses \( \pi_{t+\epsilon} \). The small delay between joining the organization and voting on the policy rules out equilibria involving self-fulfilling prophecies. These are equilibria in which agents join the organization despite disliking its policy, because they expect other like-minded members to join at the same time and immediately change the policy.

In our continuous time setting, we incorporate this distinction between incumbent and current policies in the following way. We let \( \pi_t^- \) and \( \pi_t^+ \) denote the left and right limits of the policy path at time \( t \) respectively, whenever the limits are well-defined. Then \( \pi_t \), the current policy at time \( t \), should be optimal from the point of view of the decision-maker who is pivotal given the incumbent policy \( \pi_t^- \). Similarly, \( \pi_t^+ \) should be optimal from the point of view of the decision-maker who is pivotal given \( \pi_t \). These conditions require that, for the policy to change from \( \pi \) to \( \pi' \) along the path of play, the decision-maker induced by \( \pi \) must be in favor of the change.

We define a membership function \( \beta \) so that \( \beta(x, L, \pi) = 1 \) if agent \( x \) chooses to be a member of the organization given information \( L \) and policy \( \pi \), and \( \beta(x, L, \pi) = 0 \) otherwise. We define a policy function \( \alpha \) so that \( \alpha(L, \pi) \) is the set of policies the median voter, \( m(L, \pi) \), considers optimal.\(^5\)

Our notion of strategy profile summarizes the above requirements:

**Definition 1.** A strategy profile is given by a membership function \( \beta : [0, 1] \times \mathbb{R}_+ \times \{0, 1\} \to \{0, 1\} \), a policy function \( \alpha : \mathbb{R}_+ \times \{0, 1\} \to \{\{0\}, \{1\}, \{0, 1\}\} \), and a stochastic path of play consisting of information and policy paths \((L_t, \pi_t)_t\) satisfying the following:

(a) Conditional on the policy type \( \theta \), \((L_t, \pi_t)_{t \geq 0}\) is a progressively measurable Markov

\(^5\)Note that \( \alpha(L, \pi) \) can take the values \( \{0\}, \{1\} \) and \( \{0, 1\} \). Defining \( \alpha(L, \pi) \) in this way is convenient because some paths of play cannot be easily described in terms of the instantaneous switching probabilities of individual agents.
process with paths that have left and right limits at every $t \geq 0$ satisfying $(L_0, \pi_0) = (1, 1)$.

(b) Letting $(\tilde{k}_\tau)_\tau$ denote a Poisson process with rate $b$ or 0 if $\theta = G$ or $B$ respectively, letting $(\tilde{L}_\tau)_\tau$ be given by $\tilde{L}_\tau = L (\tilde{k}_\tau, \tau)$, and letting $n(t) = \int_0^t \pi_s ds$ denote the amount of experimentation up to time $t$, we have $L_t = \tilde{L}_{n(t)}$.

(c) $\pi_t \in \alpha(L_t, \pi_{t-})$ for all $t \geq 0$.

(d) $\pi_{t+} \in \alpha(L_t, \pi_t)$ for all $t \geq 0$.

Before we provide a definition of equilibrium, a short digression on continuation utilities after deviations is required. We define $V_x(L, \pi)$ as the continuation utility of an agent with prior belief $x$ given information $L$ and incumbent policy $\pi$. In other words, $V_x(L, \pi)$ is the utility agent $x$ expects to get starting at time $t_0$ when the state follows the process $(L_t, \pi_t)_{t \geq t_0}$ given that $(L_{t_0}, \pi_{t_0}) = (L, \pi)$. In state $(L, \pi)$, the median can choose between the continuations starting in states $(L, 1)$ and $(L, 0)$. In the well-behaved case in which these continuations are different, it is natural to define the set of optimal policies $\alpha(L, \pi)$ as the set of policies $\pi'$ that maximize the median’s continuation payoff $V_m(L, \pi')(L, \pi)$.

However, if the continuations are identical, applying this definition would imply that $\alpha(L, \pi) = \{0, 1\}$ because the choice of the median $m(L, \pi)$ has no impact on the continuation. This allows for unattractive equilibria in which weakly dominated policies may be chosen: even under common knowledge that the risky policy is good, there is equilibrium in which all decision-makers choose the safe policy because any deviation to the risky policy would be reversed immediately.

To eliminate these equilibria, our definition considers short-lived deviations optimal if they would be profitable when extended for a short amount of time. To formalize this, we define $\bar{V}_x(L, \pi, \epsilon)$ as $x$’s continuation utility under the following assumptions: the state is initially $(L, \pi)$ at time $t_0$, the policy $\pi$ is locked in for a length of time $\epsilon > 0$ irrespective of the equilibrium path of play, and the equilibrium path of play continues after time $t_0 + \epsilon$.

**Definition 2.** An equilibrium $\sigma$ is a strategy profile such that:

---

6This would happen, for example, if future decision-makers coming immediately after $m(L, \pi)$ are expected to choose the same policy $\pi'$ independently of the choice of $m(L, \pi)$. 

11
(i) \( \beta(x, L, 1) = 1 \) if \( p(L, x)b > s \), \( \beta(x, L, 1) = 0 \) if \( p(L, x)b < s \) and \( \beta(x, L, 0) = 1 \) if \( r > s \).

(ii) If \( V_{\pi(L, \pi)}(L, 1) > V_{\pi(L, \pi)}(L, 0) \), then \( \alpha(L, \pi) = \{\pi'\} \).

(iii) If \( V_{\pi(L, \pi)}(L, 1) = V_{\pi(L, \pi)}(L, 0) \) but \( \nabla V_{\pi(L, \pi)}(L, \pi', \epsilon) - \nabla V_{\pi(L, \pi)}(L, 1 - \pi', \epsilon) > 0 \) for all \( \epsilon > 0 \) small enough, then \( \alpha(L, \pi) = \{\pi'\} \).

Part (i) of the definition of equilibrium says that agents are policy-takers with respect to their membership decisions. Part (ii) says that the pivotal agent chooses her preferred policy based on her expected utility, assuming that the equilibrium strategies are played in the continuation. Part (iii) is a tie-breaking rule which enforces optimal behavior even when the agent’s policy choice only affects the path of play for an infinitesimal amount of time.

Note that our definition is a special case of Markov Perfect Equilibrium, as we only allow the strategies to condition on the information about the risky policy revealed so far and on the existing policy (which determines the identity of the current median voter).

5 Equilibria in the Baseline Model

In this section we characterize the equilibria of the model described above. The presentation of the results is structured as follows. We first explain who the members of the organization are depending on what has happened in the game so far. We then make several observations which allow us to reduce the problem of equilibrium characterization to finding the optimal stopping time. Next, we state our first main result, which shows that the organization may experiment forever and provides simple sufficient conditions for this to happen (Propositions 1 and 2). Finally, in Proposition 3 we characterize the equilibria in cases when the sufficient conditions for obtaining experimentation forever fail.

We start with three useful observations. First note that, because the bad risky policy never succeeds, the posterior belief of every agent with a positive prior jumps to 1 if a success is observed. Because \( b > r, s \), if a success is ever observed, the risky policy is always used thereafter, and all agents enter the organization and remain
Second, recall that, whenever the risky policy is being used, the set of members is the set of agents for whom \( p(L, x)b \geq s \). It is clear that, for any \( L > 0 \) (that is, if no successes have been observed), \( p(L, x) \) is increasing in \( x \). That is, agents who are more optimistic at the outset remain more optimistic after observing additional information. Hence the set of members \( X_t \) is an interval of the form \([y_t, 1]\).

Third, since \( r > s \), whenever the safe policy is used, all agents choose to join the organization, and the population median becomes the pivotal decision-maker. Observe that the population median is more pessimistic than the median of any interval of the form \([y, 1]\) with \( y > 0 \). In particular, she is more pessimistic than the median voter of the organization before a switch to the safe policy. Thus if the median of the organization prefers to switch to the safe policy, so does the population median. Because no further learning happens when the safe policy is used, a switch to the safe policy is permanent.

The above observations imply that an equilibrium path must have the following structure. The risky policy is used until some time \( t^* \in [0, \infty] \). If it succeeds by then, it is used forever. Otherwise, the organization switches to the safe policy at time \( t^* \).\(^7\) While no successes are observed, agents become more pessimistic over time and the organization becomes smaller. As soon as a success occurs or the organization switches to the safe policy, all agents join and remain members of the organization forever, and no further learning occurs.

More generally, a pure strategy equilibrium can be described by a set \( t_0 < t_1 < t_2 < \ldots \) of stopping times, as follows. For any \( t \in (t_{n-1}, t_n] \), if the risky policy was used in the period \([0, t] \) and no successes were observed, the organization continues using it until time \( t_n \). If the risky policy has not succeeded by \( t_n \), the organization switches to the safe policy at \( t_n \).\(^8\)\(^9\)

Proposition 1 states our first main result. The result provides a simple condition that is sufficient for over-experimentation to arise in equilibrium. More specifically, if this condition is satisfied, then the organization uses the risky policy forever regardless

\(^7\)If \( t^* = \infty \), the risky policy is used forever.
\(^8\)\( t_0 \) is the only stopping time on the equilibrium path.
\(^9\)In principle, stopping times may be random, but for clarity it is convenient to focus on pure strategy equilibria first.
of its results.

To state Proposition 1, we will need the following definitions. We let \( V(x) \) denote the continuation utility of an agent with posterior belief \( x \) at time \( t \), provided that she expects experimentation to continue for all \( s \geq t \). We let \( m_t \) denote the median voter at time \( t \) provided that the organization has experimented unsuccessfully up to time \( t \), and \( p_t(m_t) \) is \( m_t \)'s posterior in this case.

**Proposition 1.** If \( V(p_t(m_t)) > \frac{\xi}{\gamma} \) for all \( t \), then there is an essentially unique\(^{10}\) equilibrium. In this equilibrium, if the risky policy is used at \( t = 0 \), the organization experiments forever. If \( \inf_{t \geq 0} V(p_t(m_t)) < \frac{\xi}{\gamma} \), then there is no equilibrium in which the organization experiments forever.

The intuition behind Proposition 1 is illustrated in Figure 1. As the organization experiments unsuccessfully on the equilibrium path, all agents become more pessimistic. That is, \( p_t(x) \) is decreasing in \( t \) for fixed \( x \). Letting \( x_t \) denote the agent indifferent about continuing experimentation at time \( t \), so that \( V(p_t(x_t)) = \frac{\xi}{\gamma} \), this implies that \( x_t \) must be increasing in \( t \). Thus there is a shrinking mass of agents in favor of the risky policy (the agents shaded in blue in Figure 1) and a growing mass of agents against it (the agents shaded in red and green). For high \( t \), almost all agents agree that experimentation should be stopped.

However, at the same time, growing pessimism induces members to leave. Hence the marginal member becomes more extreme, and so does the median member. If

\(^{10}\)If \( V(p_t(m_0)) = \frac{\xi}{\gamma} \) for some \( t \), there are multiple equilibria that differ in histories off the equilibrium path where the safe policy is being used and \( m_0 \) decides whether to switch to the risky policy, but all such equilibria feature the same equilibrium path.
$m_t \geq x_t$ for all $t$, that is, if the prior of the median is always higher than the prior of the indifferent agent, then the agents in favor of the risky policy always retain a majority within the organization, due to most of their opposition forfeiting their voting rights.

Figure 2 shows the same result in the space of posterior beliefs. The accumulation of negative information puts downward pressure on $p_t(m_t)$ as $t$ grows, but selection forces prevent it from converging to zero. Instead, $p_t(m_t)$ converges to a belief strictly between 0 and 1, which is above the critical value $p_t(x_t)$ in this example. Hence the median voter always remains optimistic enough to continue experimenting.

To establish whether this equilibrium entails over-experimentation, we need a definition of over-experimentation in a setting with heterogeneous priors. We will use the following notion. Consider an alternative model in which an agent with initial belief $x$ controls the policy at all times. It is well-known that whenever $0 < x < 1$, the agent would experiment until some finite time $t^*(x)$ such that her posterior belief at time $t^*(x)$ equals $\frac{\xi - 1 + b - r}{\xi + b - r}$. We say that an equilibrium of our model features over-experimentation from $x$’s point of view if experimentation continues up to some time $T > t^*(x)$. By this definition, when the condition in Proposition 1 is satisfied, there is over-experimentation from the point of view of all agents except those with prior belief exactly equal to 1.

The level of experimentation in equilibrium is determined by the interaction of two opposing forces, in addition to the usual incentives present in the canonical single-
agent bandit problem. When the pivotal agent decides whether to stop experimenting at time \( t \), she takes into account the difference in the expected flow payoffs generated by the safe policy and the risky one, as well as the option value of experimenting further. However, because the identity of the median voter changes over time, the pivotal agent knows that if she chooses to continue experimenting, the organization will stop at a time chosen by some other agent, which she likely considers suboptimal. This force encourages her to stop experimentation while the decision is still in her hands, leading to under-experimentation. It is similar to the force behind the under-experimentation result in Strulovici (2010) in that, in both cases, agents prefer a sub-optimal amount of experimentation because they expect a loss of control over future decisions if they allow experimentation to continue. It is also closely related to the concerns about slippery slopes faced by agents in the clubs literature (see, for example, Bai and Lagunoff (2011) and Acemoglu et. al. (2015)).

The second force stems from the endogeneity of the median voter’s position in the distribution. As discussed above, the more pessimistic a fixed observer becomes about the risky policy, the more extreme the median voter is. This effect is so strong that, as time passes, the posterior belief of the median after observing no successes does not converge to zero, and the median voter may choose to continue experimenting when no successes have been observed for an arbitrarily long time.

The following Proposition provides specific parameter conditions under which the organization experiments forever:

**Proposition 2.** The value function \( V \) in Proposition 1 satisfies the following:

(i) If \( f \) is non-decreasing, then

\[
\gamma \inf_{t \geq 0} V(p_t(m_t)) = \gamma V\left(\frac{2s}{b + s}\right) = \frac{2bs}{b + s} + \left(\frac{1}{2}\right)^{\gamma} \frac{s(b - s)}{b + s} \frac{b}{\gamma + b}
\]

(ii) Given \( \alpha > 0 \), let \( f_\alpha(x) \) denote a density with support \([0, 1]\) such that \( f_\alpha(x) = (\alpha+1)(1-x)^\alpha \) for \( x \in [0, 1] \). Let \( f \) be a density with support \([0, 1]\) that dominates \( f_\alpha \) in the MLRP sense, that is, \( \frac{f(x)}{f_\alpha(x)} \) is non-decreasing for \( x \in [0, 1] \). Let \( \lambda = \frac{1}{2^{\frac{1}{\alpha+1}}} \). Then

\[
\gamma \inf_{t \geq 0} V(p_t(m_t)) \geq V\left(\frac{s}{\lambda b + (1 - \lambda)s}\right) = \frac{bs}{\lambda b + (1 - \lambda)s} + \lambda^{\frac{\gamma + b}{b}} \frac{s(b - s)}{\lambda b + (1 - \lambda)s} \frac{b}{\gamma + b}
\]
(iii) Let \( f \) be any density with support \([0, 1]\). Then

\[
\gamma \inf_{t \geq 0} V(p_t(m_t)) \geq \gamma V \left( \frac{s}{b} \right) = s + \frac{s(b-s)}{\gamma + b}
\]

In all cases, the value function of the pivotal agent has a simple interpretation. The first term represents the agent’s expected flow payoff from experimentation, while the second term is the option value derived from her ability to leave the organization when she becomes pessimistic enough, and to return if there is a success.

Proposition 2 shows that the sufficient condition for obtaining experimentation forever provided in Proposition 1 is not difficult to satisfy. It is more likely to hold when \( b \) is high relative to \( r \) and \( s \), that is, when the returns from good risky technology are high, when \( \gamma \) is low, that is, when the agents are sufficiently patient, and when \( f \) does not decrease too quickly, that is, when there are enough optimists in the distribution. For example, if \( f \) is uniform, \( r = 3 \) and \( s = 2 \), then when \( b \geq 6 \), the condition holds regardless of \( \gamma \); when \( \frac{10}{3} < b < 6 \), it holds for low enough \( \gamma \); and when \( b \leq \frac{10}{3} \), it cannot hold.

The logic behind the bounds provided in Proposition 2 can be explained as follows. When \( f \) is uniform or follows a power law distribution, \( p_t(m_t) \) is decreasing and converges to some value between 0 and 1 as \( t \to \infty \). Hence, the condition in Proposition 1 reduces to checking that \( V(\lim_{t \to \infty} p_t(m_t)) \geq \frac{r}{\gamma} \). Since the marginal member \( y_t \) always has posterior belief \( \frac{s}{b} \), and the most optimistic member has posterior belief 1, \( \lim_{t \to \infty} p_t(m_t) \) must be between these values, but its position depends on the position of \( m_t \) in the interval \([y_t, 1]\). If \( f \) is uniform, \( m_t \) is the midpoint between \( y_t \) and 1, while if \( f \) decreases steeply near \( x = 1 \), then \( m_t \) is closer to \( y_t \) than 1, resulting in a value of \( \lim_{t \to \infty} p_t(m_t) \) closer to \( \frac{s}{b} \).

If there does not exist an equilibrium in which experimentation continues forever, the equilibrium analysis is more complicated. In this case there are multiple equilibria featuring different levels of experimentation on the equilibrium path, which are supported by different behavior off the path.

To characterize the set of equilibria, it is useful to define a stopping function \( \tau : [0, \infty) \to [0, \infty] \) as follows. For each \( t \geq 0 \), \( \tau(t) \geq t \) is such that \( m_t \) is indifferent about switching to the safe policy at time \( t \) if she expects a continuation where
experimentation will stop at time $\tau(t)$ should she fail to stop at $t$. If the agent never wants to continue experimenting regardless of the expected continuation, then $\tau(t) = t$, while if she always does, then $\tau(t) = \infty$.\footnote{$\tau(t)$ is unique.} Proposition 3 characterizes the set of pure strategy equilibria in this setting.

**Proposition 3.** Any pure strategy equilibrium $\sigma$ in which the organization does not experiment forever is given by a sequence of stopping times $t_0(\sigma) \leq t_1(\sigma) \leq t_2(\sigma) \leq \ldots$ such that $t_n(\sigma) = \tau(t_{n-1}(\sigma))$ for all $n > 0$ and $t_0(\sigma) \leq \tau(0)$.

There always exists $t \in [0, \tau(0)]$ for which $(t, \tau(t), \tau(\tau(t)), \ldots)$ constitutes an equilibrium. Moreover, if $\tau$ is weakly increasing, then $(t, \tau(t), \tau(\tau(t)), \ldots)$ constitutes an equilibrium for all $t \in [0, \tau(0)]$.

Proposition 3 says that, if experimenting forever is not compatible with equilibrium, then, provided that the stopping function $\tau$ is increasing, experimentation can continue on the equilibrium path for any length of time $t$ between 0 and $\tau(0)$. For each possible stopping time $t$, there is a unique sequence of off-path future stopping times that makes stopping at $t$ optimal for $m_t$. In particular, the time $t_{n+1}(\sigma)$ is chosen to leave $m_{t_{n}}(\sigma)$ indifferent about whether to continue to experiment at $t = t_{n}(\sigma)$.

The condition that $\tau$ be increasing rules out situations in which, despite $m_{t_{n}}$ being indifferent between experimenting until $t_{n+1}$ and stopping at $t_{n}$ for all $n$, the given sequence of stopping times is incompatible with equilibrium because there is some $t \in (t_{n}, t_{n+1})$ for which $m_t$ would rather stop at $t$ than experiment until $t_{n+1}$. If $\tau$ is nonmonotonic, then all equilibria must still be of the form specified in Proposition 3 but it may be that only some values of $t \in [0, \tau(0)]$ can be supported as equilibrium stopping times. Proposition 3 shows that we can always find at least one $t \in [0, \tau(0)]$ that can be supported as an equilibrium stopping time.

Finally, it can be shown that the initial median voter’s optimal stopping time in the hypothetical single-agent bandit problem where she controls the policy at all times falls between 0 and $\tau(0)$. Consequently, from the point of view of the initial median voter, both over and under-experimentation are possible depending on which equilibrium is played.
6 Other Learning Processes

The baseline model presented above has two salient features. First, when an organization pursues the risky policy for a short period of time, there is a low probability of observing a success, which increases agents’ posterior beliefs substantially, and a high probability of observing no success, which lowers their posteriors slightly. In other words, the baseline model is a model of good news. Second, because the risky policy can only succeed when it is good, good news are perfectly informative. These assumptions afford us enough tractability that closed-form solutions and a detailed equilibrium characterization can be given. When these assumptions are relaxed, more limited results can be proven. In this section, we present these results, generalizing the model to allow for bad news and imperfectly informative news.

6.1 A Model of Bad News

In this section we consider the same model as in Section 4, except that now the risky policy generates different flow payoffs. In particular, if the risky policy is good, then it generates a guaranteed flow payoff $b$. If it is bad, then it generates a guaranteed flow payoff $b$ at all times except when it experiences a failure. When using the bad risky policy, the organization experiences failures following a Poisson process with rate $b$. A failure discontinuously lowers the payoffs of all members by 1. Thus, as in the baseline model, the expected flow payoff of using the risky policy is $b$ when the policy is good and 0 when it is bad. The learning process, however, is different from the one in the baseline model.

Before characterizing the equilibrium in a model of bad news, we make a genericity assumption on the parameters. As before, $m_t$ is the median member at time $t$ provided that the risky policy has been used up to time $t$ and no failures have been observed, and $p_t(m_t)$ is her posterior belief at time $t$. We let $W_{\tau-t}(x)$ denote the value function starting at time $t$ of an agent with belief $x$ at time $t$ given a continuation equilibrium path on which the organization experiments until $T$ and then switches to the safe technology.

**Assumption 1.** The parameters are such that for all $\tau > t$, $\frac{\partial}{\partial x} W_{\tau-t}(p_t(m_t)) \neq 0$ whenever $W_{\tau-t}(p_t(m_t)) = \frac{r}{\gamma}$. 

19
Assumption 1 states that the parameters of the problem – \( b, r, \gamma \) and \( f \) – are such that the agents’ value functions are well-behaved: that is, for each \( \tau \), the derivative of the function \( t \mapsto W_{\tau-t}(p_t(m_t)) \) is not zero at any point where the function crosses the threshold \( \frac{\gamma}{\gamma} \).

Proposition 4 characterizes the equilibrium in our model that features both bad news and endogenous membership.

**Proposition 4.** Under Assumption 1, there is a unique equilibrium. The equilibrium can be described by a finite, possibly empty set of stopping intervals \( I_0 = [t_0, t_1] \), \( I_1 = [t_2, t_3], \ldots, I_n \) such that \( t_0 < t_1 < t_2 < \ldots \) as follows: conditional on the risky policy having been used during \([0, t]\) with no failures, the median \( m_t \) switches to the safe policy at time \( t \) if and only if \( t \in I_k \) for some \( k \).

Moreover, if \( f \) is non-decreasing and \( V \left( \frac{2s}{b+s} \right) \geq \frac{\gamma}{\gamma} \), the organization experiments forever unless a failure is observed.

Proposition 4 shows that the dynamics of organizations under bad news differ substantially from the dynamics observed under good news. As usual in models of bad news, as long as no failures are observed, all agents become more optimistic about the risky technology, so the organization expands over time instead of shrinking, as it did in the case of good news. This gradual expansion continues either forever or until some time \( T \) unless a failure occurs, in which case the organization switches to the safe technology and all agents previously outside the organization become members. Interestingly, the switch to the safe technology must happen upon observing a failure of the risky technology but may happen even if no failures are observed.

To understand the intuition for the results we obtain in the model of bad news, it is instructive to consider the associated single-agent bandit problem. In a bandit problem with good news, the agent uses the risky policy forever after observing a success, and becomes more and more pessimistic over time while experimenting should no successes be observed. This implies that the optimal strategy is to experiment up to some time \( t^* \) and quit if no successes have been observed by \( t^* \). In contrast, in a model of bad news, the agent switches to the safe policy forever upon observing a failure and becomes more optimistic over time if she pursues the risky policy and observes no failures. Hence, if she decides to use the risky policy at all, she uses it forever unless it fails.

20
In our model of bad news, when an agent $m_t$ is pivotal, she is more likely to choose to experiment if she expects experimentation to continue in the future. Indeed, if she prefers not to experiment at all in the single-agent bandit model, she would also switch to the safe policy here. Conversely, if she prefers to experiment in a world where she has full control over the policy, she would prefer to experiment forever. Any expected limitations to future experimentation discourage her from experimenting now, because they reduce the option value of learning about the policy.

This idea underlies the structure of the equilibrium described in Proposition 4. For $t \geq T$ and $T$ large enough, if the risky policy has been used in $[0, t]$ and no failures have been observed, most agents—including the median member, $m_t$, who will approach the population median—will be very optimistic and hence will continue to experiment forever. We can then proceed backwards and ask if there is any time $t < T$ for which $m_t$ would prefer to quit under the expectation that, if she instead experiments, experimentation will continue forever until a failure occurs. If there is some such $t$, call the highest such time $t_{2n+1}$. Now the medians $m_t$ for $t < t_{2n+1}$ face a very different problem: they know that even if they choose to experiment, $m_{t_{2n+1}}$ will switch to the safe policy at time $t_{2n+1}$. Hence, the option value of experimenting is discontinuously lower for $m_{t_{2n+1}-\epsilon}$ than it is for $m_{t_{2n+1}}$. As a result, these medians choose not to experiment for $t$ close to $t_{2n+1}$: indeed, due to their proximity to $m_{t_{2n+1}}$, they would only be slightly willing to experiment even with the maximum option value available. In turn, each $m_t$ that chooses not to experiment eliminates the option value of experimentation for $m_{\tau}$, $\tau < t$. The highest $t < t_{2n+1}$ for which $m_t$ chooses to experiment, if there is any such $t$, will be such that $m_t$ is willing to experiment in the complete absence of option value, that is, if $p_t(m_t)b \geq r$. If there is some such $t$, denote it $t_{2n}$. We can proceed in the same manner to characterize all the intervals $I_k$.

Conversely, recalling that $V(x)$ is the value function of an agent with prior $x$ provided that the organization experiments forever, if $V(p_t(m_t)) > \frac{r}{\gamma}$ for all $t$, then the organization must experiment forever. The last part of Proposition 4 follows from the fact that, if $f$ is non-decreasing, then $p_t(m_t) \geq \frac{2\epsilon}{b+\delta}$ for all $t$, as was the case in Proposition 1.

Several important conclusions follow from the analysis above. First, as in the previous model, experimentation can continue forever (although in this case it is not...
as surprising because this result can arise even in the associated single-agent bandit problem). Second, over-experimentation is never possible from the point of view of any pivotal agent. Indeed, if $m_t$ did not want to experiment in a single-agent bandit problem, then she could stop at time $t$. If she did want to experiment, she would want to experiment forever. Therefore, whatever level of experimentation the organization allows would at most be equal to her bliss point.

Third, under-experimentation (from the point of view of pivotal agents) is possible, and often obtains when experimentation does not continue forever. Indeed, if the equilibrium described in Proposition 4 has two intervals, $I_0 = [t_0, t_1]$ and $I_1 = [t_2, t_3]$, then all agents $m_t$ for $t$ between $t_1$ and $t_2$ would rather experiment forever than experiment until time $t_2$. The same logic applies whenever the equilibrium features three or more intervals.

Fourth, although under-experimentation was also possible in the previous model, the mechanism is different in this case. Here agents do not stop experimenting lest experimentation continue for too long – they stop experimenting because experimentation will not continue for long enough.

### 6.2 A Model of Imperfectly Informative (Good) News

In the previous models, agents’ posterior beliefs only move in one direction, except for when a perfectly informative event occurs, after which no more interesting decisions are made. The reader might wonder whether the results are sensitive to this feature of our assumptions. To address this, we consider a model with imperfectly informative news, which allows for rich dynamics even after observing successes or failures. For brevity, we consider the case of good news, but similar results can be obtained for imperfectly informative bad news.

Again, the model is the same as in Section 4 except for the payoffs generated by the risky policy. If the risky policy is good, it generates successes according to a Poisson process with rate $b$. If it is bad, it generates successes according to a Poisson process with rate $c$. We now assume that $b > r > s > c > 0$.

As before, the effect of past information on the agents’ beliefs can be aggregated into a one-dimensional sufficient statistic. Suppose that the organization has used the
risky policy for a length of time $t$ and that $k$ successes have occurred during that time. Define

$$L(k, t) = \left(\frac{c}{b}\right)^k e^{(b-c)t}$$

Then the posterior of an agent with prior $x$ at time $t$ after observing the organization use the risky policy for a length of time $t$ and achieve $k$ successes is

$$\frac{x}{x + (1-x)L(k, t)}$$

As before, we suppress the dependence of $L(k, t)$ on $k$ and $t$ and use $L$ to denote our sufficient statistic. Note that high $L$ indicates bad news about the risky policy.

We use $V_x(L)$ to denote the value function of an agent with prior $x$ given that the state is $L$ provided that on the equilibrium path experimentation continues forever. We let $V(x) = V_x(1)$ denote the ex-ante value function of an agent with prior $x$ in the same model and under the same equilibrium. The next Proposition shows that, as in the previous variants of the model, for certain parameter values experimentation can continue forever regardless of how badly the risky policy performs.

**Proposition 5.** If $f$ is non-decreasing and $V\left(\frac{2(s-c)}{(b-c)+(s-c)}\right) \geq \frac{r}{\gamma}$, then there is a unique equilibrium. In this equilibrium, if the risky policy is used at $t = 0$, the organization experiments forever. If $V\left(\frac{2(s-c)}{(b-c)+(s-c)}\right) < \frac{r}{\gamma}$, then there is no equilibrium in which the organization experiments forever.

Moreover, $V\left(\frac{2(s-c)}{(b-c)+(s-c)}\right) \geq \frac{1}{\gamma} \frac{(b-c)s+(s-c)b}{(b-c)+(s-c)}$, so there exist parameter values such that $V\left(\frac{2(s-c)}{(b-c)+(s-c)}\right) \geq \frac{r}{\gamma}$.

It is more difficult to give an exact expression for the value function $V$ in this case owing to the complicated behavior of $L$ over time. For the same reason, it is not feasible to fully characterize the set of equilibria. Nevertheless, the following result illustrates the novel outcomes that can arise in this case.

**Proposition 6.** There exist $b$, $r$, $s$, $c$ and $f$, and $\epsilon \in (0, 1]$ and $L^* > 0$, such that an equilibrium of the following form exists: whenever $L = L^*$, the organization stops experimenting with probability $\epsilon$, and if $L \neq L^*$, then the organization continues

\[\text{[12]}\text{The agent's posterior only depends on } k \text{ and } t, \text{ and not on the timing of the successes.}\]
experimenting with probability one.

The intuition behind the equilibrium is as follows. Suppose that the density of prior beliefs $f$ is such that $\mu(L)$ increases in $L$ quickly to the right of a certain value $L^*$, but slowly to its left— for instance, because $f(x)$ is low for $x > \mu(L^*)$ and high for $x < \mu(L^*)$ — and that, as a result, $L \mapsto p(L, \mu(L))$ is decreasing for $L < L^*$ but increasing for $L > L^*$. It may then be that $p(L, \mu(L))$ has a global minimum at $L = L^*$, that is, the median voter is most pessimistic when $L = L^*$. If the parameters are chosen appropriately, this median voter will be indifferent about stopping experimentation, and hence willing to do so with some probability $\epsilon > 0$, while other agents prefer to continue experimenting when they are pivotal decision-makers.

The striking feature of this equilibrium is that stopping only happens for an intermediate value of $L$. In particular, if $L^* < L(0,0) = 1$, the only way experimentation will stop is if it succeeds enough times for $L$ to decrease all the way to $L^*$. As a result, we obtain the counterintuitive result that experimentation may be more likely to continue forever precisely when the risky policy is bad:

**Corollary 1.** The parameters in Proposition 6 can be chosen so that, in addition to the equilibrium being as described there, the probability that the organization never stops experimenting is higher when the risky policy is bad than when it is good.

In the equilibrium we just described persistent failure makes organizations more radical, as reflected in their willingness to experiment with a policy disliked by most agents, while success may make organizations more conservative and more prone to stop experimentation.

The result in Proposition 6 can be obtained only for certain densities of prior beliefs. In particular, it can be shown that if $f$ is uniform, or follows any power law, then the posterior belief of the median $p(L, \mu(L))$ must be decreasing in $L$ (Lemma 17) and hence such an equilibrium cannot be constructed.
7 Other Extensions

7.1 General Voting Rules

We assume throughout the paper that the median voter—that is, the voter at the 50th percentile of the belief distribution among members—is pivotal. This assumption is not essential to our analysis: our results can be extended to other voting rules under which the agent at the $q$-th percentile is pivotal.

It is instructive to consider how the results change as we vary $q$, assuming that $f$ is uniform. In this case, if the organization has experimented continuously for time $t$ and has not experienced a success, the posterior belief of the decision-maker $q_t$ in the organization at time $t$ is given by $p_t(q_t) = \frac{s+q(b-s)e^{-bt}}{qs+(1-q)b+q(1-s)e^{-bt}}$. As $t \to \infty$, the posterior of the decision-maker converges to $\frac{s}{qs+(1-q)b}$. This observation has two important consequences. First, Proposition 2 can be extended in straightforward fashion to this more general case. In particular, for any voting rule $q \in (0,1)$, there exist parameters such that in the unique equilibrium the organization experiments forever. Second, as $q \mapsto \frac{s}{qs+(1-q)b}$ is increasing, more stringent supermajority requirements are functionally equivalent to more optimistic leadership of the organization, and make it easier to sustain an equilibrium with over-experimentation.

7.2 Size-Dependent Payoffs

In some settings the payoffs that an organization generates may depend on its size. In this section we discuss how different operationalizations of this assumption affect our results. We find that our main result is robust to this extension, and that different kinds of size-dependent payoffs may make over-experimentation easier or more difficult to obtain.

We consider three types of size-dependent payoffs. For the first two, we suppose that when the set of members of the organization has measure $\mu$, the safe policy gives members a flow payoff $rg(\mu)$, the good risky policy yields instantaneous payoffs of size $g(\mu)$ generated at rate $b$, and the bad risky policy yields zero payoffs. We assume that $g(1) = 1$, so that $b, r$ and 0 are the expected flow payoffs from the good risky policy, the safe policy and the bad risky policy respectively when all agents are in
the organization. For the first type of payoffs we consider, \( g(\mu) \) is increasing in \( \mu \), so there are economies of scale. For the second type, \( g(\mu) \) is decreasing in \( \mu \), so there is a congestion effect.

In general, the effect of size-dependent payoffs on the level of experimentation is ambiguous because of two countervailing effects. On the one hand, when there is a congestion effect, as the organization shrinks, higher flow payoffs increase the benefits from experimentation, which makes experimentation more attractive.\(^{13}\) We call this the \textit{payoff effect}. On the other hand, because increasing flow payoffs provide incentives for agents to stay in the organization, the organization shrinks at a lower speed, which causes the median voter in control of the organization to be more pessimistic about the risky policy. We call this the \textit{control effect}. When there are economies of scale, these effects are reversed.

When there are economies of scale, \( X_t \) may not be uniquely determined as a function of the state at time \( t \). This is because the more members there are, the higher payoffs are, so the membership stage may have multiple equilibria. We will assume, for simplicity, that \( X_t \) is uniquely determined.\(^{14}\) It is sufficient to assume that \( g \) does not increase too fast.

The following Proposition presents our first result:

\textbf{Proposition 7.} Suppose that \( f = f_\alpha \).\(^{15}\) Let \( \bar{g} = \lim_{\mu \to 0} g(\mu) \), and let \( V_{g,t}(p_t(m_t)) \) denote the utility of the pivotal agent at time \( t \) if she expects experimentation to continue forever. If

\[
\frac{\bar{g}}{\lambda} s \frac{b}{\gamma + b} + \frac{s}{\lambda} \frac{\gamma}{\gamma + b} > b \frac{b}{\gamma + b} + s \frac{\gamma}{\gamma + b}
\]

then \( \lim_{t \to \infty} V_{g,t}(p_t(m_t)) \) is strictly increasing in \( \bar{g} \) for all \( \bar{g} \in [s, \infty) \). In this case, perpetual experimentation obtains for a greater set of parameter values with the congestion effect and for a smaller set of parameter values with economies of scale, relative to the baseline model.

\(^{13}\) Note that, while the safe policy could also yield high payoffs when the organization is small, all agents will enter as soon as the safe policy is implemented, so these high payoffs can never be captured.

\(^{14}\) Formally, we require that the equation \( \frac{\mu_t}{y_t + (1-y_t)e^{bt}} = \frac{s}{g(1-y_t)b} \) has a unique fixed point for all \( t \geq 0 \).

\(^{15}\) Recall that \( f_\alpha(x) \) is a density with support \([0, 1]\) such that \( f_\alpha(x) = (\alpha + 1)(1-x)^{\alpha} \) for \( x \in [0, 1] \).
Conversely, if the reverse inequality holds strictly, then \( \lim_{t \to \infty} V_{g,t}(p_t(m_t)) \) is strictly decreasing in \( \bar{g} \) for all \( \bar{g} \in [s, \infty) \).

The intuition for the Proposition is as follows. By the same argument as in the baseline model, Proposition 1 holds: a sufficient condition to obtain experimentation forever is that \( V_{g,t}(p_t(m_t)) \geq r \gamma \) for all \( t \). While it is difficult to calculate \( V_{g,t}(p_t(m_t)) \) explicitly for all \( t \), calculating its limit as \( t \to \infty \) is tractable and often allows us to determine whether the needed condition holds for all \( t \). We show that the limit depends only on \( \bar{g} \) rather than the entire function \( g \). Moreover, it is a hyperbola in \( \bar{g} \), so it is either increasing or decreasing in \( \bar{g} \) everywhere. In the first case, size-dependent payoffs affect the equilibrium mainly through the payoff effect, so experimentation is more attractive with a congestion effect and less so with economies of scale. In the second case, the control effect dominates, and the comparative statics are reversed. These statements are precise as \( t \to \infty \) (that is, conditional on the risky policy having been used for a long time). We can show that when congestion effects make experimentation more likely in the limit, they do so for all \( t \).

The inequality in the Proposition determines which case we are in. Because \( b > \lambda^2 s \) and \( \frac{\xi}{\gamma} > s \), if \( b \) is large enough relative to \( \gamma \), then over-experimentation is easier to obtain with economies of scale than in the baseline model, and easier to obtain in the baseline model than with a congestion effect. The opposite happens if \( \gamma \) is large relative to \( b \). The reason is that, under economies of scale, the pivotal decision-maker is very optimistic about the risky policy but expects to receive a low payoff from the first success. If \( \frac{b}{\gamma} \) is large, so that successes arrive at a high rate or the agent is very patient, the first success is expected to be one of many, while if \( \frac{b}{\gamma} \) is small, further successes are expected to be heavily discounted. Conversely, with a congestion effect, for large \( t \) the pivotal decision-maker is almost certain that the risky policy is bad but believes that, with a low probability, it will net a very large payoff before she leaves.

The third way to operationalize size-dependent payoffs that we consider deals with changes of the learning rate rather than flow payoffs. Here we suppose that when the organization is of size \( \mu \), the good risky policy generates successes at a rate \( b \mu \). Each success pays a total of 1, which is split evenly among members, so that

\[16\text{It can be shown that when congestion effects make experimentation less likely in the limit, they may not do so for all } t.\]
each member gets $\frac{1}{\mu}$. All other payoffs are the same as in the baseline model. An example that fits this setting is a group of researchers trying to find a breakthrough. If there are fewer researchers, breakthroughs are just as valuable but happen less often. Letting $f = f_\alpha$ and using $\hat{V}$ to denote the continuation utility under perpetual experimentation, we have

$$
\gamma \inf \hat{V}_t(p_t(m_t)) = \gamma \lim_{t \to \infty} \hat{V}_t \left( \frac{s}{\lambda b + (1 - \lambda)s} \right) = \frac{bs}{\lambda b + (1 - \lambda)s}
$$

Comparing this with the condition in Proposition 2, we find that the condition for obtaining experimentation forever is more difficult to satisfy but can still hold as long as the flow payoff that the pivotal agent obtains from experimentation is higher than that from the safe policy. This is a consequence of Proposition 2, where we showed that the pivotal agent’s expected utility from experimentation equals her flow payoff from the risky policy plus the option value derived from her ability to leave and re-enter. Here, in the limit, the option value vanishes as learning becomes very slow, so only the flow payoff from experimentation is left.

### 7.3 Common Values

Although our model features agents with private values, our results can be extended to a model with common values, which is more appropriate for some of our applications, such as environmental organizations or civil rights activism.

We discuss this extension in the context of our example of civil rights activism. The flow payoff of an agent with belief $x$ is now a flow contribution to a rate of change in the relevant laws which can be attributed to the activism of agent $x$. The mapping from membership and policy decisions to the outcomes is the same as in Section 4, but now agents care only about the overall rate of changes in the law, and not about their own contribution.

Formally, we let $U_x(\sigma_y, \sigma)$ denote the private utility of agent $x$ when she plays the equilibrium strategy of agent $y$ and the equilibrium path is dictated by the strategy profile $\sigma$. Then in the private values case $x$’s equilibrium utility is $U_x(\sigma_x, \sigma)$, while in the common values case it is $\int_0^1 U_x(\sigma_y, \sigma) f(y) dy$. Note that, even though all agents
share the objective of maximizing the aggregate rate of legal change, their utility functions still differ due to differences in the prior beliefs.

Agents are now indifferent about their membership decisions: the membership status of a set of agents of measure zero has no impact on anyone’s payoffs. However, it is natural to assume that each agent should choose to join when doing so would be optimal if her behavior had a positive weight in her own utility function.\footnote{For instance, this is the case in a model with a finite number of agents.} Under this assumption, agents still make membership decisions that maximize their flow contributions at each point in time, just as in Section 4.

Let us conjecture a strategy profile in which the organization experiments forever, and let $\tilde{V}_t(x)$ denote the continuation utility at time $t$ of an agent who has posterior belief $x$ at time $t$ under this strategy profile.\footnote{This matters in this case, in contrast to the model in Section 4, because the membership strategies of other agents, which depend on $t$ rather than $x$, enter the agent’s utility function.} Then Proposition 1 holds with the same proof, replacing $V(p_t(m_t))$ in the original statement with $\tilde{V}_t(p_t(m_t))$. Moreover, the following lower bound for the value function holds:\footnote{Unlike the result in Proposition 2, the bound for $\tilde{V}_t(x)$ provided here is not tight, because a tight bound that can be expressed in closed form does not exist for general densities $f$.}

**Proposition 8.** For any $x \in [0, 1]$ and any $t \geq 0$,

$$V(x) \geq \tilde{V}_t(x) \geq \min \left\{ \frac{b}{\gamma} x - \frac{b}{\gamma} \frac{b s}{\gamma} - x \frac{b - s}{\gamma + b} \right\}$$

Proposition 8 can be used to obtain an analog of Proposition 2 for this setting. For instance, when the density of prior beliefs $f$ is uniform, experimentation continues forever as long as

$$\min \left\{ \frac{2bs}{b + s} - \frac{2bs}{b + s} + \frac{(b - s)s}{b + s} \left( 1 - \frac{2\gamma}{\gamma + b} \right) \right\} \geq r$$

In other words, over-experimentation can still occur in equilibrium for reasonable parameter values.

The above result differs from that in Proposition 2 in two important ways. First, the fact that $\tilde{V}_t(x) \leq V(x)$ means that agents’ payoffs from experimentation are always weakly lower in the common values case than in the private values case,
while the payoffs from the safe policy are identical. As a result, over-experimentation occurs for a strictly smaller set of parameter values in the common values case. The reason is that, under common values, an agent considers the entry and exit decisions of other agents suboptimal, and her payoff is affected by these decisions as long as experimentation continues. In contrast, in the private values case, agents’ payoffs depend only on their own entry and exit decisions, which are chosen optimally given their beliefs.

The second important difference is that, as shown in Proposition 2, in the private values case, experimentation can continue forever even if agents are impatient, as long as the density $f$ does not decrease too quickly near 1 and other parameters are chosen appropriately (for example, $s$ is close to $r$ and $b$ is high enough). This occurs because the pivotal agent is optimistic enough that the expected flow payoff from experimentation is higher than $r$, even without taking the option value into account. In contrast, in the common values case, the expected flow payoff from experimentation goes to $s$ as $t \to \infty$ if there are no successes, no matter how optimistic the agent is. Indeed, here agents care about the contributions of all players, and they understand that for large $t$ most players will become outsiders and generate $s$, regardless of the quality of the policy. Hence perpetual experimentation is only possible if agents are patient enough.
References


Lemma 1. Suppose that the initial distribution of priors is $f_\alpha$ for some $\alpha \geq 0$, as in Proposition 2. The posterior belief of the median member of the organization at time $t$, provided that experimentation has continued from time 0 to time $t$ and no successes have been observed, is

$$p_t(m_t) = \frac{s + (1 - \lambda)(b - s)e^{-bt}}{\lambda(b - s) + s + (1 - \lambda)(b - s)e^{-bt}}$$

Proof of Lemma 1. The posterior belief of agent $x$ at time $t$ is given by

$$p_t(x) = \frac{xe^{-bt}}{xe^{-bt} + 1 - x}.$$

Using the fact that $p_t(y_t) = \frac{s}{b}$ for the marginal member $y_t$, we set $\frac{y_te^{-bt}}{y_te^{-bt} + 1 - y_t} = \frac{s}{b}$. Solving for $y_t$, we obtain

$$y_t = \frac{s}{b} + \frac{(1 - \frac{s}{b})e^{-bt}}{s + (b - s)e^{-bt}} = \frac{s}{s + (b - s)e^{-bt}}$$

The median $m_t$ must satisfy the condition $2 \int_{m_t}^{1} f_\alpha(x)dx = \int_{y_t}^{1} f_\alpha(x)dx$, so that $2(1 - m_t)^{\alpha+1} = (1 - y_t)^{\alpha+1}$. Hence $1 - m_t = \lambda(1 - y_t)$, which implies that

$$m_t = 1 - \lambda + \lambda y_t = 1 - \lambda + \lambda \frac{s}{s + (b - s)e^{-bt}} = \frac{s + (1 - \lambda)(b - s)e^{-bt}}{s + (b - s)e^{-bt}}$$

Substituting the above expression into the formula for $p_t(x)$, we obtain

$$p_t(m_t) = \frac{m_t e^{-bt}}{1 - m_t + m_t e^{-bt}} = \frac{m_t}{(1 - m_t)e^{-bt} + m_t e^{-bt}} = \frac{m_t}{s + (1 - \lambda)(b - s)e^{-bt}}$$

$$= \frac{s + (1 - \lambda)(b - s)e^{-bt}}{\lambda(b - s) + s + (1 - \lambda)(b - s)e^{-bt}}$$

In particular, if $\alpha = 0$, then $f$ is uniform and we have

$$p_t(m_t) = \frac{2s + (b - s)e^{-bt}}{b + s + (b - s)e^{-bt}}$$

\[\blacksquare\]
We now provide a formula for $V(x)$.

**Lemma 2.** Let $t(x)$ denote the time it will take for an agent’s posterior belief to go from $x$ to $\frac{x}{b}$ (provided that no successes are observed during this time), at which time she would leave the organization. Then

$$V(x) = xb \cdot \frac{1}{\gamma} + (1-x)e^{-\gamma t(x)} \cdot \frac{s}{\gamma} - x(b-s) \frac{e^{-(\gamma+b)t(x)}}{\gamma+b}$$

**Proof of Lemma 2.**

Let $P_t = x(1-e^{-bt})$ denote the probability that an agent with prior belief $x$ assigns to having a success by time $t$. Then

$$V(x) = x \int_0^{t(x)} be^{-\gamma \tau} d\tau + \int_{t(x)}^\infty (P_\tau b + (1-P_\tau)s) e^{-\gamma \tau} d\tau$$

The first term is the payoff from time 0 to time $t(x)$, when the agent stays in the organization. The second term is the payoff after time $t(x)$, when the agent leaves the organization and obtains the flow payoff $s$ thereafter, unless the risky technology has had a success (in which case the agent returns to the organization and receives a guaranteed expected flow payoff $b$).

We have

$$V(x) = x \int_0^{t(x)} be^{-\gamma \tau} d\tau + \int_{t(x)}^\infty P_\tau (b-s) e^{-\gamma \tau} d\tau$$

$$= xb \cdot \frac{1-e^{-\gamma t(x)}}{\gamma} + e^{-\gamma t(x)} \cdot \frac{s}{\gamma} + x(b-s) \left( \int_{t(x)}^\infty P_\tau e^{-\gamma \tau} d\tau - \int_{t(x)}^\infty (1-P_\tau) e^{-\gamma \tau} d\tau \right)$$

$$= xb \cdot \frac{1-e^{-\gamma t(x)}}{\gamma} + e^{-\gamma t(x)} \cdot \frac{s}{\gamma} + x(b-s) \left( \frac{e^{-(\gamma+b)t(x)}}{\gamma+b} - \frac{e^{-(\gamma+b)t(x)}}{\gamma+b} \right)$$

$$= xb \cdot \frac{1}{\gamma} + (1-x)e^{-\gamma t(x)} \cdot \frac{s}{\gamma} - x(b-s) \frac{e^{-(\gamma+b)t(x)}}{\gamma+b}$$

where the second equality follows from the fact that $\int_0^t e^{-\gamma \tau} d\tau = \frac{1-e^{-\gamma t}}{\gamma}$, $\int_t^\infty e^{-\gamma \tau} d\tau = \frac{e^{-\gamma t}}{\gamma}$, and $P_t = x(1-e^{-bt})$, and the third equality follows from the fact that $\int_t^\infty e^{-\gamma \tau} d\tau = \frac{e^{-(\gamma+b)t}}{\gamma+b}$. 

**Lemma 3.** Let $t^y(x)$ denote the time it takes for an agent’s posterior belief to go from
x to y. Then
\[ t^y(x) = -\frac{\ln\left(\frac{y}{1-y} - \frac{x}{b}\right)}{x} \quad t(x) = -\frac{\ln\left(\frac{s(1-x)}{(b-s)x}\right)}{b} \]

If \( x = \frac{2s}{b+s} \), then \( e^{-bt(x)} = \frac{1}{2} \). If \( x = \frac{s}{\lambda b+(1-\lambda)s} \), then \( e^{-bt(x)} = \lambda \).

**Proof of Lemma 3.**

We solve \( p_t(x) = \frac{x e^{-bt}}{x e^{-bt} + (1-x)} = y \) for \( t \). Then we obtain \( e^{-bt^y(x)} = \frac{y}{1-y} - \frac{x}{b} \) or, equivalently, \( t^y(x) = -\frac{\ln\left(\frac{y}{1-y} - \frac{x}{b}\right)}{x} \).

In particular, \( t(x) = t^x(x) = -\frac{\ln\left(\frac{s(1-x)}{(b-s)x}\right)}{b} \). Substituting \( x = \frac{2s}{b+s} \) into \( e^{-bt(x)} = \frac{s(1-x)}{(b-s)x} \) and simplifying, we obtain \( e^{-bt} = \frac{1}{2} \). Substituting \( x = \frac{s}{\lambda b+(1-\lambda)s} \) into \( e^{-bt(x)} = \frac{s(1-x)}{(b-s)x} \) and simplifying, we obtain \( e^{-bt(x)} = \lambda \). ■

**Lemma 4.** Let \( V_x(L, \pi) \) denote the value function of the agent with prior \( x \) when the initial state is \((L, \pi)\). Then for all \((L, \pi)\), \( x \mapsto V_x(L, \pi) \) is strictly increasing for all agents with prior \( x \) that are in the experimenting organization at a set of times of positive measure with a positive probability (on the equilibrium path).

**Proof of Lemma 4.**

Consider agents with priors \( x' > x \). Let \( V^y_{x'}(L, \pi) \) denote the payoff to the agent with prior \( x' \) from copying the equilibrium strategy of the agent with prior \( x \). When \( x \) and \( x' \) are outside the organization, their flow payoffs are equal to \( s \) and do not depend on their priors.

When \( x' \) is in the organization, if the organization is using the risky policy, at a continuation where the state variable is \( \tilde{L} \), \( x' \)'s expected flow payoff is \( p(\tilde{L}, x') b \). Because \( x' > x \), we have \( p(\tilde{L}, x') > p(\tilde{L}, x) \) and thus \( p(\tilde{L}, x') b > p(\tilde{L}, x) b \), so the flow payoff of \( x' \) is higher than the flow payoff of \( x \) when \( x \) and \( x' \) are inside the organization.

Observe that we have \( V^y_{x'}(L, \pi) > V_x(L, \pi) \) for all agents with prior \( x \) that are in the experimenting organization at a set of times of positive measure with a positive probability. Because \( V_{x'}(L, \pi) \geq V^y_{x'}(L, \pi) \), we then have \( V_{x'}(L, \pi) \geq V^y_{x'}(L, \pi) > V_x(L, \pi) \), which implies that \( V_{x'}(L, \pi) > V_x(L, \pi) \) for all agents with prior \( x \) that are in the experimenting organization at a set of times of positive measure with a positive probability, as required. ■
Corollary 2. If the organization experiments forever on the equilibrium path, then
\[ V_x(L, \pi) = V_{p(L,x)}(1, \pi). \] Moreover, \( V_{p(L,x)}(1, \pi) \) is increasing in \( p(L, x) \).

Lemma 5. For any policy path \((\pi_t)_t\) with left and right-limits everywhere, there is
another policy path \((\hat{\pi}_t)_t\) such that \(\hat{\pi}_0 = \pi_0\), \((\hat{\pi}_t)_t\) is càdlàg for all \(t > 0\), and \((\hat{\pi}_t)_t\) is
equal to \((\pi_t)_t\) almost everywhere.\(^{20}\)

Proof of Lemma 5.

Define \(\hat{\pi}_0 = \pi_0\) and \(\hat{\pi}_t = \pi_t \) for all \(t > 0\). Let
\(T = \mathbb{R}_{>0} \setminus \{t \geq 0 : \pi_{t^-} = \pi_t = \pi_{t^+}\}\). Because \((\pi_t)_t\) has left and right-limits everywhere, \(T\) must be countable—
otherwise \(T\) would have an accumulation point \(t_0\), and either the left-limit or right-
limit of \((\pi_t)_t\) at \(t_0\) would not be well-defined. Then, since \(\hat{\pi}_t = \pi_t\) for all \(t \notin T\), \((\hat{\pi}_t)_t\)
and \((\pi_t)_t\) only differ on a countable set. Moreover, it is straightforward to show that,
for all \(t > 0\), \(\hat{\pi}_{t^-} = \pi_{t^-}\) and \(\hat{\pi}_{t^+} = \pi_{t^+}\), so \((\hat{\pi}_t)_t\) is càdlàg. \(\blacksquare\)

Corollary 3. For any strategy profile \((\beta, \alpha, (L_t, \pi_t)| (L, \pi, \theta))\) the stochastic process
\((L_t, \hat{\pi}_t)| (L, \pi, \theta)\) (where \((\hat{\pi}_t)_t\) is as in Lemma 5) has càdlàg paths, satisfies Conditions
(a) and (b), and induces a path of play that yields the same payoffs as the strategy.

Lemma 6 (Recursive Decomposition). Let \(\Theta \subseteq \mathbb{R}^n\) be a closed set, let \((\theta_t)_t\)
be a right-continuous progressively measurable Markov process with support contained in
\(\Theta\), let \(f\) be a bounded function, and let
\[ U(\theta_0) = \int_0^\infty e^{-\gamma t} E_{\theta_0}[f(\theta_t)] dt. \]

Let \(A\) be a closed subset of \(\Theta\) and define a stochastic process \((s_t)_t\) with a co-
domain \((A \cup \{\emptyset\})\) as follows: \(s_t = \emptyset\) if there exists \(t' \leq t\) such that \(\theta_{t'} = \emptyset\) and
\(\theta_{t''} \notin A\) for all \(t'' < t'\). If this is not true for any \(\theta \in A\), then \(s_t = \emptyset\).\(^{21}\) Then
\[ U(\theta_0) = \int_0^\infty e^{-\gamma t} E_{\theta_0} [f(\theta_t) \mathbb{1}_{\{s_t = \emptyset\}}] dt + \int_A U(\theta) dP_s \]
where \(P_s\) is defined as follows: \(P_{s_t}\) is the probability measure on \(A \cup \emptyset\) induced by \(s_t\),
and \(P_s = \gamma \int_0^\infty e^{-\gamma t} P_{s_t} dt.\)

\(^{20}\)Hence it is payoff-equivalent to \(\pi_t\) and generates the same learning path \((L_t)_t\).

\(^{21}\)In other words, \(s_t\) takes the value of the first \(\theta \in A\) that \((\theta_t)_t\) hits.
Proof of Lemma 6.

We have

\[
U(\theta_0) = \int_0^\infty e^{-\gamma t} E_{\theta_0} [f(\theta_t)] dt = \\
\int_0^\infty e^{-\gamma t} E_{\theta_0} [f(\theta_t) \mathbb{1}_{\{s_t = \emptyset\}}] dt + \int_0^\infty e^{-\gamma t} E_{\theta_0} [f(\theta_t) \mathbb{1}_{\{s_t \in A\}}] dt
\]

So it remains to show that

\[
\int_A U(\theta) dP_s = \int_0^\infty e^{-\gamma t} E_{\theta_0} [f(\theta_t) \mathbb{1}_{\{s_t \in A\}}] dt
\]

Define a random variable \(z\) with a co-domain \((A \times [0, \infty)) \cup \{\emptyset\}\) as follows:

\[z = (\theta, t)\] if \(\theta_t = \theta \in A\) and \(\theta_{t'} \notin A\) for all \(t' < t\). If this is not true for any \(\theta \in A\) and \(t \geq 0\), then \(z = \emptyset\).\(^{22}\) Let \(P_z\) be the probability measure on \((A \times [0, \infty)) \cup \{\emptyset\}\) induced by \(z\). Let \(\theta(z)\) and \(t(z)\) be the random variables equal to the first and second coordinates of \(z\), conditional on \(z \neq \emptyset\). Note that \(s_t = \theta\) if and only if \(z = (\theta, t')\) for some \(t' \leq t\).

Then we can write

\[
\int_0^\infty e^{-\gamma t} E_{\theta_0} [f(\theta_t) \mathbb{1}_{\{s_t \in A\}}] dt = \int_0^\infty e^{-\gamma t} E_{\theta_0} [f(\theta_t) \mathbb{1}_{\{z \in A \times [0, \infty)\}} \mathbb{1}_{\{t \geq t(z)\}}] dt = \\
= \int_{A \times [0, \infty)} \left( \int_{t(z)}^\infty e^{-\gamma t} E_{\theta_0} [f(\theta_t) \mathbb{1}_{\{z \in A \times [0, \infty)\}}] dt \right) dP_z = \\
= \int_{A \times [0, \infty)} e^{-\gamma t(z)} U(\theta(z)) dP_z = \\
= \int_{A \times [0, \infty)} \left( \int_{t(z)}^\infty \gamma e^{-\gamma t} dt \right) U(\theta(z)) dP_z = \\
= \int_0^\infty \int_{A \times [0, \infty)} \gamma e^{-\gamma t} \mathbb{1}_{\{t \geq t(z)\}} U(\theta(z)) dP_z dt = \\
= \int_0^\infty \gamma e^{-\gamma t} \left( \int_{A \times [0, \infty)} \mathbb{1}_{\{t \geq t(z)\}} U(\theta(z)) dP_z \right) dt = \\
= \int_0^\infty \gamma e^{-\gamma t} \left( \int_A U(\theta) dP_s \right) dt = \int_A U(\theta) dP_s
\]

\(^{22}\)In other words, \(z\) takes the value of the first \(\theta \in A\) that \((\theta_t)_t\) hits, and the time when it hits.
as desired.

Let $W_T(x)$ denote the continuation value of an agent with current belief $x$ in an equilibrium in which the organization stops experimenting after a length of time $T$. Let $T^* = \arg\max_T W_T(x)$ denote the optimal amount of time that an agent with prior $x$ would want to experiment for if she was always in control of the organization.

Lemma 7. (i) $T \mapsto W_T(x)$ is differentiable for all $T \in (0, \infty)$ and right-differentiable at $T = 0$.

(ii) $W_0(x) = \frac{r}{\gamma}$ and $\left. \frac{\partial}{\partial T} W_T(x) \right|_{T=0} = \max\{xb, s\} - r + \frac{xb(b-r)}{\gamma}$.

(iii) $T \mapsto W_T(x)$ is strictly increasing for $T \in [0, T^*]$ and strictly decreasing for $T > T^*$.

(iv) If $V(x) > \frac{r}{\gamma}$, then $W_T(x) > \frac{r}{\gamma}$ for all $T > 0$.

Proof of Lemma 7.

Fix $T_0 \geq 0$ and $\epsilon > 0$. By Lemma 6, we have

$$W_{T_0+\epsilon}(x) - W_{T_0}(x) = e^{-rT_0}Q_{T_0}(x) \left( W_\epsilon(p_{T_0}(x)) - W_0(p_{T_0}(x)) \right)$$

$$= e^{-rT_0}Q_{T_0}(x) \left( W_\epsilon(p_{T_0}(x)) - \frac{r}{\gamma} \right)$$

where $Q_T(x)$ is the probability that there is no success up to time $T$, based on the prior $x$, and $p_T(x)$ is the posterior belief of an agent with prior $x$ in this case. Then

$$\left. \frac{\partial W_T(x)}{\partial T} \right|_{T=T_0} = \lim_{\epsilon \searrow 0} \frac{W_{T_0+\epsilon}(x) - W_{T_0}(x)}{\epsilon} =$$

$$= \lim_{\epsilon \searrow 0} e^{-rT_0}Q_{T_0}(x) \frac{W_\epsilon(p_{T_0}(x)) - \frac{r}{\gamma}}{\epsilon} = e^{-rT_0}Q_{T_0}(x) \left. \frac{\partial W_T(p_{T_0}(x))}{\partial T} \right|_{T=0}$$

The case with $\epsilon < 0$ is analogous. Then it is enough to prove that $T \mapsto W_T(x)$ is right-differentiable at $T = 0$. This can be done by calculating $W_T(x)$ explicitly for small $T > 0$. We do this for $x > \frac{b}{b-r}$. In this case, for $T$ sufficiently small, $x$ is in the organization because the experiment with the risky policy is going to stop before she

---

23 This is defined for $T \in [0, \infty]$, where $W_\infty(x) = V(x)$. 

37
wants to leave. Then

\[ W_T(x) = \int_0^T x e^{-\gamma t} dt + \int_T^\infty e^{-\gamma t} (Q_T(x)r + (1 - Q_T(x))b) dt = \]

\[ = xb \frac{1 - e^{-\gamma T}}{\gamma} + \frac{e^{-\gamma T}}{\gamma} (r + (1 - Q_T(x))(b - r)) = \]

\[ = xb \frac{1 - e^{-\gamma T}}{\gamma} + \frac{e^{-\gamma T}}{\gamma} (r + x (1 - e^{-bT})(b - r)) \]

This implies that \( \frac{\partial W_T(x)}{\partial T} \bigg|_{T=0} = xb - r + \frac{xb(b-r)}{\gamma} \).

Similarly, it can be shown that if \( x \leq \frac{s}{b} \), then \( \frac{\partial W_T(x)}{\partial T} \bigg|_{T=0} = s - r + \frac{xb(b-r)}{\gamma} \). This proves (i) and (ii).

For (iii), note that, by our previous result, \( \frac{\partial W_T(x)}{\partial T} \bigg|_{T=T^*} \) is positive (negative) whenever \( \frac{\partial W_T(p_{T_0}(x))}{\partial T} \bigg|_{T=0} \) is positive (negative). In addition, it follows from our calculations that \( y \mapsto \frac{\partial W_T(y)}{\partial T} \bigg|_{T=0} \) is increasing and \( T_0 \mapsto p_{T_0}(x) \) is decreasing. Moreover, for large \( T_0 \), \( p_{T_0}(x) \) is close to zero, so \( \frac{\partial W_T(p_{T_0}(x))}{\partial T} \bigg|_{T=0} \) is negative. It follows that \( T \mapsto W_T(x) \) is single-peaked. If \( \frac{\partial W_T(x)}{\partial T} \bigg|_{T=T^*} > 0 \), then the peak is the unique \( T^* \) satisfying \( \frac{\partial W_T(x)}{\partial T} \bigg|_{T=T^*} = 0 \). If \( \frac{\partial W_T(x)}{\partial T} \bigg|_{T=0} \leq 0 \), then \( T^* = 0 \).

Hence if \( 0 < T \leq T^* \), then \( W_T(x) > W_0(x) = \frac{x}{\gamma} \) because in this case \( T \mapsto W_T(x) \) is increasing by (iii), and if \( T > T^* \), then \( W_T(x) \geq \lim_{T \to \infty} W_T(x) = V(x) \) because in this case \( T \mapsto W_T(x) \) is decreasing by (iii). This proves (iv).

Lemma 8. Let \( m(L) \) and \( \tilde{m}(L) \) denote the median voter when the state variable is \( (L, 1) \) and the density is \( f \) and \( \tilde{f} \) respectively. Suppose that \( \tilde{f} \) MLRP-dominates \( f \). Then \( \tilde{m}(L) \geq m(L) \) for all \( L \).

Proof of lemma 8.

Let \( y(L) \) denote the indifferent agent given information \( L \) under either density (note that \( y(L) \) is given by the condition \( p(L, y(L)) = \frac{x}{b} \), which is independent of the density). By definition, we have

\[ \int_{y(L)}^{m(L)} f(x) dx = \int_0^1 f(x) dx \]

38
Suppose that \( \tilde{m}(L) < m(L) \). This is equivalent to
\[
\int_{y(L)}^{m(L)} s(x)f(x)dx = \int_{y(L)}^{m(L)} \tilde{f}(x)dx > \int_{m(L)}^{1} \tilde{f}(x)dx = \int_{m(L)}^{1} s(x)f(x)dx
\]
where \( s(x) = \frac{\tilde{f}(x)}{f(x)} \).

Since \( \tilde{f} \) MLRP-dominates \( f \), \( s(x) \) is weakly increasing. Thus
\[
\int_{y(L)}^{m(L)} s(m(L))f(x)dx \geq \int_{y(L)}^{m(L)} s(x)f(x)dx > \int_{m(L)}^{1} s(x)f(x)dx \geq \int_{m(L)}^{1} s(m(L))f(x)dx
\]
which is a contradiction. \( \blacksquare \)

**Lemma 9.** Let \( m(L) \) and \( \tilde{m}(L) \) denote the median voters when the state variable is \( L \) under the uniform density and a non-decreasing density \( f \) respectively. Suppose that \( y(L) \to 1 \) as \( L \to \infty \). Then \( \frac{1-\tilde{m}(L)}{1-m(L)} \to 1 \) as \( L \to \infty \).

**Proof of Lemma 9.**

Given a state variable \( L \) and the marginal member \( y(L) \) corresponding to it, let \( f_{0L} = f(y(L)) \) and \( f_1 = f(1) \). By Lemma 8, we have \( m(L) \leq \tilde{m}(L) \leq \hat{m}(L) \) where \( \hat{m}(L) \) is the median corresponding to a density \( \hat{f} \) such that \( \hat{f}(x) = f_{0L} \) for \( x \in [y(L), \hat{m}(L)] \) and \( \hat{f}(x) = f_1 \) for \( x \in [\hat{m}(L), 1] \).

By construction, because \( \hat{m}(L) \) is the median, we have \( f_{0L}(\hat{m}(L) - y(L)) = f_1(1 - \hat{m}(L)) \), so \( \hat{m}(L) = \frac{f_{0L}y(L) + f_1}{f_{0L} + f_1} \). Thus \( 1 - \hat{m}(L) = \frac{f_{0L}(1-y(L))}{2f_{0L} + f_1} \) and, because \( m(L) = \frac{y(L) + 1}{2} \) so that \( 1 - m(L) = \frac{1-y(L)}{2} \), we have \( \frac{1-\hat{m}(L)}{1-m(L)} = \frac{2f_{0L}}{f_{0L} + f_1} \).

Since \( f \) is increasing, using the fact that \( f(x) \to \sup_{y \in [0,1]} f(y) \) as \( x \to 1 \), we find that \( f(x) \to f(1) \) as \( x \to 1 \). Then, as \( t \to \infty \), we have \( y(L) \to 1 \), \( f_{0L} = f(y(L)) \to f_1 \) and \( \frac{1-\hat{m}(L)}{1-m(L)} \to 1 \). \( \blacksquare \)

**Lemma 10.** Let \( x_t, \tilde{x}_t \) be two time-indexed sequences of agents such that \( x_t \leq \tilde{x}_t \) for all \( t \) and \( x_t \to 1 \) as \( t \to \infty \). If \( \frac{1-x_t}{1-\tilde{x}_t} \to 1 \), then \( \frac{p_t(\tilde{x}_t)}{p_t(x_t)} \to 1 \).

**Proof of Lemma 10.**
Using the formula for the posterior beliefs, we have

\[
\frac{p_t(\tilde{x}_t)}{p_t(x_t)} = \frac{\tilde{x}_t x_t + (1 - x_t)L_t}{\tilde{x}_t x_t + (1 - \tilde{x}_t)L_t}.
\]

Since \(x_t \to 1\) and \(\tilde{x}_t \geq x_t\) for all \(t\), \(\tilde{x}_t \to 1\), whence \(\tilde{x}_t x_t \to 1\). In addition, since \(1 - x_t \to 1\), \(1 - \tilde{x}_t \to 1\), whence \(\tilde{x}_t x_t \to 1\). As a result, for all \(t\),

\[
\min \left\{ \frac{x_t}{\tilde{x}_t}, \frac{(1 - x_t)L_t}{(1 - \tilde{x}_t)L_t} \right\} \leq \frac{x_t + (1 - x_t)L_t}{\tilde{x}_t + (1 - \tilde{x}_t)L_t} \leq \max \left\{ \frac{x_t}{\tilde{x}_t}, \frac{(1 - x_t)L_t}{(1 - \tilde{x}_t)L_t} \right\}
\]

so \(\frac{x_t + (1 - x_t)L_t}{\tilde{x}_t + (1 - \tilde{x}_t)L_t} \to 1\), which concludes the proof. 

Lemma 11 shows that agents strictly prefer the risky policy after a success.

**Lemma 11.** In any equilibrium, \(\alpha(0, 1) = \alpha(0, 0) = 1\).

**Proof of Lemma 11.**

Observe that \(L_{t_0} = 0\) implies \(L_t = 0\) for all \(t \geq t_0\) no matter what policy path is followed, and hence \(p(L_t, x) = 1\) for all \(t\) and \(x\). For the rest of the argument, we can then write \(V(0, \pi)\) instead of \(V_x(0, \pi)\).

By Lemma 6, there is \(\rho \in [0, 1]\) such that

\[
V(0, 0) = \rho \frac{r}{\gamma} + (1 - \rho)V(0, 1) \tag{1}
\]

It follows that there exist \(\eta \in [0, 1]\) and \(\eta' \in [0, 1]\) such that \(\eta \geq \eta'\)

and

\[
V(0, 0) = \eta \frac{r}{\gamma} + (1 - \eta) \frac{b}{\gamma} \quad V(0, 1) = \eta' \frac{r}{\gamma} + (1 - \eta') \frac{b}{\gamma}
\]

\(\eta\) and \(\eta'\) are the discounted fractions of the expected time that the organization spends on the safe policy, when starting in states \((0, 0)\) and \((0, 1)\), respectively.

Observe that if \(\eta > \eta'\), then \(V(0, 0) < V(0, 1)\). In particular, \(V_{m(0, \pi)}(0, 1) > \)

\[\text{We have } \eta \geq \eta' \text{ for the following reason. } V(0, 1) = \eta' \frac{r}{\gamma} + (1 - \eta' \frac{b}{\gamma} \text{ and (1) imply that } \eta \frac{r}{\gamma} + (1 - \eta) \frac{b}{\gamma} = V(0, 0) = \rho \frac{r}{\gamma} + (1 - \rho)V(0, 1) = (\rho + (1 - \rho)\eta') \frac{r}{\gamma} + (1 - \rho)(1 - \eta') \frac{b}{\gamma}. \text{ Then } \eta = \rho + (1 - \rho)\eta', \text{ which implies that } \eta \geq \eta', \text{ as required.}\]
$V_{m(0,\pi)}(0,0)$ for all $\pi$, which implies that $\alpha(0,\pi) = 1$ for all $\pi$ by Condition (ii). If $\eta = \eta'$, then $V(0,0) = V(0,1)$. Because it can be shown that $V(0,0,\epsilon) < V(0,1,\epsilon)$ for any $\epsilon > 0$, by Condition (iii), in this case we must also have $\alpha(0,\pi) = 1$ for all $\pi$. ■

Lemma 12. For any state $(L,\pi)$, there is a CDF $G$ with support\textsuperscript{25} contained in $[0,\infty]$ such that

$$V_x(L,\pi) = \int_0^\infty W_T(p(L,x))dG(T)$$

for all $x \in [0,1]$, where $W_T(y)$ is as defined in Lemma 7.

Similarly, for any state $(L,\pi)$ and any $\epsilon > 0$, there is a distribution $G_\epsilon$ with support contained in $[0,\infty]$ such that

$$\nabla_x(L,\pi,\epsilon) = \int_0^\infty W_T(p(L,x))dG_\epsilon(T)$$

for any $x \in [0,1]$.

Proof of Lemma 12.

We prove the first statement. The proof of the second statement is analogous.

Note that we can, without loss of generality, assume that the distribution over future states $(L,\pi)$ induced by the continuation starting in state $(L,\pi)$ satisfies the following: the policy is equal to 1 in the beginning and, if it ever changes from 1 to 0, it never changes back to 1. Indeed, suppose that $\pi$ switches from 1 to 0 at time $t$ and switches back at a random time $t + \nu$, where $\nu$ is distributed according to some CDF $H$. Let $p = \int_0^\infty e^{-\gamma \nu}dH(\nu)$. Then a continuation path on which the policy only switches to 0 at time $t$ with probability $1 - p$ and never returns to 1 after switching induces the same discounted distribution over future states.

Under the above assumption and given that the policy always remains at 1 after a success by Lemma 11, the path of play can be described as follows: experimentation continues uninterrupted until a success or a permanent stop. Then we can let $G$ be the CDF of the stopping time, conditional on no success being observed. ■

\textsuperscript{25}G is a degenerate CDF that can take the value $\infty$ with positive probability. Equivalently, $G$ satisfies all the standard conditions for the definition of a CDF, except that $\lim_{T \to \infty} G(T) \leq 1$ instead of $\lim_{T \to \infty} G(T) = 1$. This is needed to allow for the case where experimentation continues forever.
Lemma 13 shows that switches to the safe policy are permanent.

**Lemma 13.** In any equilibrium, for any \( L \), if \( 0 \in \alpha(L, 1) \), then \( \alpha(L, 0) = 0 \).

**Proof of Lemma 13.**

If \( L = 0 \), then \( \alpha(0, \pi) = 1 \) for all \( \pi \) by Lemma 11, so the statement is vacuously true. Suppose then that \( L > 0 \). Suppose for the sake of contradiction that the statement is false.

Observe that for all \( L \) there is \( \rho_L \in [0, 1] \) independent of \( x \) such that

\[
V_x(L, 0) = \rho_L \frac{r}{\gamma} + (1 - \rho_L)V_x(L, 1)
\]

for all \( x \). This follows from Lemma 6, with the added observation that \( \rho_L \) (equivalently, \( P_s \) in the notation of Lemma 6) is independent of \( x \) in this case because the stochastic process governing \((L, \pi)\) is independent of \( x \) if \( \pi = 0 \).

**Case 1:** Suppose that \( \rho_L > 0 \), and that the expected amount of experimentation under the continuation starting in state \((L, 1)\) is positive. By Lemma 4, \( V_x(L, 1) \) is strictly increasing in \( x \). Then equation (2) implies that \( V_x(L, 1) - V_x(L, 0) \) is strictly increasing in \( x \). Since \( m(L, 1) > m(L, 0) \), we have \( V_{m(L,1)}(L, 1) - V_{m(L,1)}(L, 0) > V_{m(L,0)}(L, 1) - V_{m(L,0)}(L, 0) \). Since \( 1 \in \alpha(L, 0) \) implies that \( V_{m(L,0)}(L, 1) - V_{m(L,0)}(L, 0) \geq 0 \), we have \( V_{m(L,1)}(L, 1) - V_{m(L,1)}(L, 0) > 0 \), and thus \( \alpha(L, 1) = 1 \), a contradiction.

**Case 2:** Suppose that \( \rho_L = 0 \). We make two observations. First, \( V_x(L, 0) = V_x(L, 1) \) for all \( x \). Second, the expected amount of experimentation under the continuation starting in state \((L, 1)\) is positive. Indeed, \( \rho_L = 0 \) implies that, conditional on the state at \( t \) being \((L_t, \pi_t) = (L, 0)\), we have \( \text{inf}\{t' > t : \pi_{t'} = 1\} = t \) a.s. Since Condition (d) requires that \( \pi_{t+} \) exists and the result that \( \text{inf}\{t' > t : \pi_{t'} = 1\} = t \) a.s. rules out that \( \pi_{t+} = 0 \) with a positive probability, it must be that \( \pi_{t+} = 1 \) a.s. In turn, this implies that \( \text{inf}\{t' > t : \pi_{t'} = 0\} > t \) a.s. Then \( E[\text{inf}\{t' > t : \pi_{t'} = 0\}] - t > 0 \).

\(^{26}\)If \((L_t, \pi_t)\) has càdlàg paths, this follows from Lemma 6. If not, then Lemma 6 cannot be applied because the stochastic process in question is not necessarily right-continuous. However, we can use Corollary 3 of Lemma 5 to obtain a payoff-equivalent path of play with càdlàg paths and then apply Lemma 6 to it.
By definition, we have

\[
\bar{V}_x(L, 0, \epsilon) = \rho_\epsilon \frac{r}{\gamma} + (1 - \rho_\epsilon)V_x(L, 1)
\]

(3)

for \( \rho_\epsilon = 1 - e^{-\gamma \epsilon} \).

In the following argument, for convenience, we subtract \( \frac{r}{\gamma} \) from every value function.\(^{27}\) Note that, because we lock policy 1 in for time \( \epsilon \), \( G_\epsilon \) satisfies \( 1 - G_\epsilon(T) = \min \left\{ \frac{1 - G(T)}{1 - G(\epsilon)}, 1 \right\} \). Then for \( T \in [0, \epsilon] \), \( 1 - G_\epsilon(T) = 1 \) and for \( T > \epsilon \), \( 1 - G_\epsilon(T) = \frac{1 - G(T)}{1 - G(\epsilon)} \). Hence for \( \epsilon > 0 \) sufficiently small we have

\[
\bar{V}_x(L, 1, \epsilon) = \int_0^\infty W_T(p(L, x)) dG_\epsilon(T) = \\
= \int_0^\epsilon W_T(p(L, x)) dG_\epsilon(T) + \int_\epsilon^\infty W_T(p(L, x)) dG_\epsilon(T) = \\
= 0 + \frac{1}{1 - G(\epsilon)} \int_\epsilon^\infty W_T(p(L, x)) dG(T) = \\
= \frac{V_x(L, 1)}{1 - G(\epsilon)} - \frac{1}{1 - G(\epsilon)} \int_0^\epsilon W_T(p(L, x)) dG(T) = \frac{V_x(L, 1)}{1 - G(\epsilon)} + G(\epsilon)\mathcal{O}(\epsilon)
\]

The first part of the third equality follows from the fact that \( G_\epsilon(T) = 0 \) for all \( T \in [0, \epsilon] \) and the second part of the third equality follows from the fact that \( G_\epsilon(T) = \frac{G(T) - G(\epsilon)}{1 - G(\epsilon)} \) for \( T > \epsilon \). The last equality follows from the fact that \( \lim_{\epsilon \to 0} G(\epsilon) = 0 \) since \( \inf \{ t' > t : \pi_{t'} = 1 \} = t \) a.s. and from the fact that \( \frac{\partial_x W_T(x)}{\partial T} \bigg|_{T=0} = \max \{ x, s \} - r + \frac{x b (b - r)}{\gamma} \) by part (ii) of lemma 7.\(^{28}\)

Suppose for the sake of contradiction that \( V_{m(L, 0)}(L, 1) < 0 \). Note that (3) then implies that \( V_{m(L, 0)}(L, 1) < \bar{V}_{m(L, 0)}(L, 0, \epsilon) \). Therefore, we have \( \bar{V}_{m(L, 0)}(L, 1) \leq V_{m(L, 0)}(L, 1) < \bar{V}_{m(L, 0)}(L, 0, \epsilon) \) for all \( \epsilon > 0 \) sufficiently small and hence \( \alpha(L, 0) = 0 \), a contradiction. Hence \( V_{m(L, 0)}(L, 1) \geq 0 \). It follows that, because \( m(L, 1) > m(L, 0) \) and, by Lemma 4, \( x \mapsto V_x(L, 1) \) is strictly increasing, we have \( V_{m(L, 1)}(L, 1) > 0 \).

---

\(^{27}\)That is, we let \( \hat{V}_x(L, \pi) = V_x(L, \pi) - \frac{r}{\gamma}, \hat{W}_T(x) = W_T(x) - \frac{r}{\gamma}, \hat{V}_x(L, \pi, \epsilon) = V_x(L, \pi, \epsilon) - \frac{r}{\gamma} \). For the rest of this proof, we work with the normalized functions \( \hat{V}, \hat{W}, \hat{V} \), but drop the operator \( \hat{\cdot} \) to simplify notation.

\(^{28}\)In greater detail, \( \int_0^T W_T(p(L, x)) dG(T) \approx \int_0^T (\alpha_0 T + \alpha_1) dG(T) \leq \int_0^T (\alpha_0 \epsilon + \alpha_1) dG(T) = (\alpha_0 \epsilon + \alpha_1) \int_0^T dG(T) = (\alpha_0 \epsilon + \alpha_1)(G(\epsilon) - G(0)) = (\alpha_0 \epsilon + \alpha_1)G(\epsilon) = \alpha_0 G(\epsilon) = G(\epsilon)\mathcal{O}(\epsilon) \) where we have used the fact that we subtracted \( \frac{r}{\gamma} \) from every value function to get rid of the constant \( \alpha_1 \).
Then $\nabla x(L, 1, \epsilon) = \frac{V_x(L, 1)}{1 - G(\epsilon)} + G(\epsilon)O(\epsilon)$ implies that $\nabla m(L, 1)(L, 1, \epsilon) \geq V_m(L, 1)(L, 1)$.\footnote{\(\nabla m(L, 1)(L, 1, \epsilon) \geq V_m(L, 1)(L, 1)\) is then equivalent to \(G(\epsilon)(1 - G(\epsilon))O(\epsilon) \geq -G(\epsilon)V_m(L, 1)(L, 1)\), which is satisfied for \(V_m(L, 1)(L, 1) > 0\).}

Moreover, because \(V_m(L, 1)(L, 1) > 0\), we have $V_m(L, 1)(L, 1) > \nabla m(L, 1)(L, 0, \epsilon)$, as $\nabla m(L, 1)(L, 0, \epsilon)$ is a convex combination of $V_m(L, 1)(L, 1)$ and $0$.\footnote{Recall that we have subtracted \(\frac{L}{\gamma}\) from every value function.}

Then $\nabla m(L, 1)(L, 1, \epsilon) \geq V_m(L, 1)(L, 1) > \nabla m(L, 1)(L, 0, \epsilon)$ for all $\epsilon > 0$ sufficiently small. By Condition (iii), this implies that $\alpha(L, 1) = 1$, a contradiction.

Case 3: Suppose that the expected amount of experimentation starting in state $(L, 1)$ is zero. In this case $V_x(L, 0) = V_x(L, 1) = \frac{L}{\gamma}$ for all $x$, and $\nabla x(L, 0, \epsilon) = \frac{L}{\gamma}$ for all $x$ and $\epsilon > 0$.

Again, we subtract $\frac{L}{\gamma}$ from every value function for simplicity.

By definition, for all $\epsilon > 0$ the path starting in state $(L, 1, \epsilon)$ has a positive expected amount of experimentation. Moreover, $G_\epsilon$ defined in Lemma 12 is FOSD-decreasing in $\epsilon$ (that is, if $\epsilon' < \epsilon$, then $G_{\epsilon'} \geq G_\epsilon$) and hence, taken as a function of $\epsilon$, has a pointwise limit $G$ (that is, $G_\epsilon(T) \to G(T)$ for all $T \geq 0$). Then

$$\nabla x(L, 1, \epsilon) \to \int_0^\infty W_T(p(L, x))dG(T)$$

Since $1 \in \alpha(L, 0)$, there exists a sequence $\epsilon_n \searrow 0$ such that $\nabla m(L, 0)(L, 1, \epsilon_n) \geq 0$ for all $n$,\footnote{Suppose for the sake of contradiction that $1 \in \alpha(L, 0)$ and such a sequence does not exist. Then for all $\epsilon > 0$ sufficiently small we have $\nabla m(L, 0)(L, 1, \epsilon) < 0$ (note that we have used the fact that we subtract $\frac{L}{\gamma}$ from every value function here). Then $\nabla m(L, 0)(L, 1, \epsilon) < \nabla m(L, 0)(L, 0, \epsilon) = 0$, which contradicts $1 \in \alpha(L, 0)$ by Condition (iii).} whence $\lim_{\epsilon \to 0} \nabla m(L, 0)(L, 1, \epsilon) \geq 0$.

There are now two cases. First, if $E_G[T] > 0$, we can use the following argument. $\lim_{\epsilon \to 0} \nabla m(L, 0)(L, 1, \epsilon) \geq 0$ implies that $\lim_{\epsilon \to 0} \nabla m(L, 1)(L, 1, \epsilon) > 0$ because $m(L, 1) > m(L, 0)$ and $x \mapsto V_x(L, 1)$ is strictly increasing by Lemma 4 (note that we have used the fact that $E_G[T] > 0$ to apply Lemma 4 here). Because $\epsilon \mapsto \nabla m(L, 1)(L, 1, \epsilon)$ is continuous, it follows that $\nabla m(L, 1)(L, 1, \epsilon) > 0$ for all $\epsilon > 0$ sufficiently small. But then $\alpha(L, 1) = 1$ by Condition (iii), a contradiction.

Second, if $E_G[T] = 0$, then we can employ a similar argument using the fact
that, by part (ii) of lemma 7, $\frac{\partial W_\epsilon(p(L,x))}{\partial \epsilon}\bigg|_{\epsilon=0}$ is strictly increasing in $x$ and that, by Lemma 7, we have

$$\lim_{\epsilon \to 0} \frac{V_x(L, 1, \epsilon)}{E_{G_\epsilon}[T]} = \lim_{\epsilon \to 0} \frac{W_\epsilon(p(L, x))}{\epsilon} = \frac{\partial W_\epsilon(p(L, x))}{\partial \epsilon}\bigg|_{\epsilon=0}$$

Proof of Proposition 1.

We first argue that if $\inf_{t \geq 0} V(p_t(m_t)) \geq \frac{\xi}{\gamma}$, then experimenting forever is an equilibrium if the organization is experimenting at time $t = 0$.

Consider the following strategy profile: $\alpha(L, 1) = 1$ for all $L \in \{0\} \cup [1, \infty)$, $\alpha(L, 0) = 1$ if $V(p(L, m(L, 0))) > \frac{\xi}{\gamma}$, and $\beta(x, L, \pi)$ is given by (i) in the definition of the equilibrium. The path of play is as follows. If the organization is in state $(L, 1)$ at time $t_0$, then $\pi_t = 1$ for all $t > t_0$. If the organization is in state $(L, 0)$ at time $t_0$, then $t_1 = \inf \left\{ t \geq t_0 : V(p(L_t, m(L_t, 0))) > \frac{\xi}{\gamma} \right\}$. It follows that $\pi_t = 1$ for all $t \geq t_1$ and $\pi_t = 0$ for $t < t_1$. We can check that Conditions (a)-(d) and (i)-(iii) hold, so this is an equilibrium.

Next, we argue that if $V(p_t(m_t)) > \frac{\xi}{\gamma}$ for all $t \geq 0$, then any equilibrium must be of this form.

Let $\sigma$ be an equilibrium. By Lemma 11, the risky policy must always be used after a success. By Lemma 13, after a switch from the risky policy to the safe policy, the safe policy will be used forever.

Now suppose for the sake of contradiction that $\sigma$ is not such that, starting with policy 1, policy 1 is used forever with probability one. In other words, suppose that there is $L$ for which $0 \in \alpha(L, 1)$. By Condition (ii) in the definition of the equilibrium, this requires that $\frac{\xi}{\gamma} \geq V_{m(L, 1)}(L, 1)$.

Let $T$ be the (possibly random) time until the policy first switches to 0, starting in state $(L, 1)$. Then $V_{m(L, 1)}(L, 1) = E[W_T(p(L, m(L, 1)))], \text{ where } W_T(x)$ is as in Lemma 7 and the expectation is taken over $T$.

Recall that $V(p(L, m(L, 1))) > \frac{\xi}{\gamma}$ by assumption. By Lemma 7, this im-
plies that $W_T(p(L, m(L, 1)))) > \frac{r}{\gamma}$ for all $T > 0$. If $E[T] > 0$, it follows that $E[W_T(p(L, m(L, 1))))] > \frac{r}{\gamma}$, implying that $V_{m(L, 1)}(L, 1) > \frac{r}{\gamma}$. This implies that $\alpha(L, 1) = 1$ by Condition (ii), which is a contradiction. If $E[T] = 0$, then $V_{m(L, 1)}(L, 1) = \frac{r}{\gamma}$ but, by the same argument, $V_{m(L, 1)}(L, \epsilon) = \frac{r}{\gamma}$ for all $\epsilon > 0$ and thus $\alpha(L, 1) = 1$ by Condition (iii), which is a contradiction.

Finally, suppose that $\inf_{t \geq 0} V(p_t(m_t)) < \frac{r}{\gamma}$, so that $V(p_{t_0}(m_{t_0})) < \frac{r}{\gamma}$ for some $t_0$, and suppose that there is an equilibrium in which $\pi_t = 1$ for all $t$ starting in state $(1, 1)$. This requires that $1 \in \alpha(L_{t_0}, 1)$, which implies that $V(p_{t_0}(m_{t_0})) \geq \frac{r}{\gamma}$ by Condition (ii), a contradiction.

\[ \blacksquare \]

**Proof of Proposition 2.**

We prove each inequality in three steps.

First, we show that the median posterior belief is uniformly bounded below for all $t$, with different bounds depending on the density. When $f$ is uniform, we have $p_t(m_t) \geq \frac{2s}{b+s}$. For any $\alpha > 0$, when $f = f_{\alpha}$, we have $p_t(m_t) \geq \frac{s}{\lambda b + (1-\lambda)s}$ for $\lambda = \frac{1}{2\pi \tau}$. Finally, if $f$ has full support, we have $p_t(m_t) \geq \frac{s}{b}$. The first two claims follow from Lemma 1, which shows that $p_t(m_t) \searrow \frac{2s}{b+s}$ as $t \to \infty$ when $f$ is uniform, and $p_t(m_t) \searrow \frac{s}{\lambda b + (1-\lambda)s}$ when $f = f_{\alpha}$. The last claim is implied by the fact that $m_t \geq y_t$ and $p_t(y_t) = \frac{s}{b}$.

Second, we argue that these bounds hold not just for the aforementioned densities but also for any that dominate them in the MLRP sense. This follows from Lemma 8 and the fact that the function $x \mapsto p_t(x)$ is strictly increasing.

Third, we observe that, since $V(x)$ is strictly increasing and continuous in $x$ (by Lemmas 2 and 4), we have $\inf_{t \geq 0} V(p_t(m_t)) = V(\inf_{t \geq 0} p_t(m_t))$. Hence, to arrive at the bounds in the Proposition, it is enough to evaluate $V$ at the appropriate beliefs.

To calculate $V\left(\frac{s}{\lambda b + (1-\lambda)s}\right)$, we use Lemmas 2 and 3. The time it takes for an agent with belief $\frac{s}{\lambda b + (1-\lambda)s}$ to reach posterior $\frac{s}{b}$ is

\[
t \left( \frac{s}{\lambda b + (1-\lambda)s} \right) = -\ln \left( \frac{s}{b} \right) \frac{\lambda(b-s)}{b} = -\ln \lambda \frac{b}{b} \]
Thus, taking \( x = \frac{s}{\lambda b + (1 - \lambda)s} \), we have \( e^{-bl(x)} = \lambda \) and \( e^{-\gamma t(x)} = \lambda^{\frac{t}{b}} \). Substituting this value of \( x \) into the formula for \( V(x) \) from Lemma 2, we obtain

\[
\gamma V \left( \frac{s}{\lambda b + (1 - \lambda)s} \right) = \frac{bs}{\lambda b + (1 - \lambda)s} + \frac{\lambda(b - s)s}{\lambda b + (1 - \lambda)s} \lambda^{\frac{t}{b}} - \frac{s}{\lambda b + (1 - \lambda)s} \frac{(b - s)\gamma \lambda^{\frac{t}{b}}}{\gamma + b} \\
= \frac{bs}{\lambda b + (1 - \lambda)s} + \frac{(b - s)s}{\lambda b + (1 - \lambda)s} \frac{b}{\gamma + b} \lambda^{\frac{t}{b}}
\]

In particular, for \( \alpha = 0 \), this becomes

\[
\gamma V \left( \frac{2s}{b + s} \right) = \frac{2bs}{b + s} + \left( 1 + \frac{1}{2} \right) \frac{b}{s} \frac{(b - s)s}{b + s} \frac{\gamma}{\gamma + b}
\]

On the other hand, for \( x = \frac{s}{b} \), we have \( t(x) = 0 \). Substituting this in, we obtain

\[
\gamma V \left( \frac{s}{b} \right) = s + \frac{b - s}{b} s - \frac{s}{b}(b - s) \frac{\gamma}{\gamma + b} = s + \frac{(b - s)s}{\gamma + b}
\]

An additional argument is required to show that the bound is tight in part (i).

Take \( f \) to be any non-decreasing density. Let \( \tilde{m}_t \) denote the median at time \( t \) under \( f \), and let \( m_t \) denote the median at time \( t \) under the uniform density. It is sufficient to show that the asymptotic posterior of the median is \( \frac{2s}{b + s} \) under \( f \), that is, that \( \lim_{t \to \infty} p_t(\tilde{m}_t) = \lim_{t \to \infty} p_t(m_t) = \frac{2s}{b + s} \).

Lemma 9 shows that \( \frac{1 - \tilde{m}_t}{1 - m_t} \to 1 \). Note that we have \( \tilde{m}_t \geq m_t \) for all \( t \) by Lemma 8 and \( m_t \to 1 \) as \( t \to \infty \). Then Lemma 10 applies. Lemma 10 applied to the sequences \( \tilde{m}_t \) and \( m_t \) guarantees that the ratio \( \frac{p_t(\tilde{m}_t)}{p_t(m_t)} \) converges to 1.\[\blacksquare\]

**Lemma 14.** There exist parameters such that \( V \left( \frac{2s}{b + s} \right) \geq \frac{r}{\gamma} \).

**Proof of Lemma 14.**

It is easy to show that there exist parameters such that \( \frac{2bs}{b + s} + \left( \frac{1}{2} \right) \frac{b}{s} \frac{(b - s)s}{b + s} \frac{\gamma}{\gamma + b} \geq r \) is satisfied. For example, suppose that \( \gamma \approx \infty \) (agents are infinitely impatient). Then we need that \( \frac{2bs}{b + s} \geq r \). For this, it is sufficient to have \( s > \frac{r}{2} \) and \( b \geq \frac{s \gamma}{2s - r} \).\[\blacksquare\]

**Proof of Proposition 3.**

47
We first argue that the stopping function \( \tau \) is well-defined.

Let \( t \) be the current time and let \( t^* \) be the time at which \( m_t \) would choose to stop experimenting if she had complete control over the policy. Recall the definition of \( W_{T-t}(x) \) from Lemma 7: \( W_{T-t}(x) \) is the value function starting at time \( t \) of an agent with belief \( x \) at time \( t \) given a continuation equilibrium path on which the organization experiments until \( T \) and then switches to the safe technology. Then, equivalently, \( t^* = \arg \max_T W_{T-t}(x). \)

There are three cases. If \( t^* = t \), then \( \tau(t) = t \). If \( t^* > t \), that is, if \( x \) wants to experiment for a positive amount of time, and \( V(p_t(m_t)) < \frac{\gamma}{\gamma} \), then \( W_{T-t}(p_t(m_t)) \) is strictly increasing in \( T \) for \( T \in [t, t^*) \) and strictly decreasing in \( T \) for \( T > t^* \), as shown in Lemma 7, and there is a unique \( \tau(t) > t^* \) for which \( W_{\tau(t)-t}(p_t(m_t)) = \frac{\gamma}{\gamma} \). Finally, if \( t^* > t \) and \( V(p_t(m_t)) \geq \frac{\gamma}{\gamma} \), then \( \tau(t) = \infty \).

Next, note that \( \tau \) is continuous. If \( \tau(t_0) \in (t_0, \infty) \), then for \( t \) in a neighborhood of \( t_0 \), \( \tau(t) \) is defined by the condition \( W_{\tau(t)-t}(p_t(m_t)) = \frac{\gamma}{\gamma} \), where \( p_t(m_t) \) is continuous in \( t \), and \( W_T(x) \) is differentiable in \( (T, x) \) at \( (T, x) = (\tau(t), p_t(m_t)) \) and strictly decreasing in \( T \), so the continuity of \( \tau \) follows from the Implicit Function Theorem. The proofs for the cases when \( \tau(t_0) = 0 \) or \( \tau(t_0) = \infty \) are similar.

Consider a pure strategy equilibrium \( \sigma \) in which the organization does not experiment forever on the equilibrium path. Let \( t_0(\sigma) \) be the time at which experimentation stops on the equilibrium path. Clearly, we have \( t_0(\sigma) \leq \tau(0) \), as otherwise \( m_0 \) would switch to the safe policy at time 0. As before, if a success occurs or if the organization switches to the safe policy, everyone joins the organization permanently.

Consider what happens at time \( t_0(\sigma) \) if \( m_{t_0(\sigma)} \) deviates and continues experimenting. Suppose first that \( \tau(t_0(\sigma)) \in (t_0(\sigma), \infty) \). In a pure strategy equilibrium, there must be a time \( t_1(\sigma) \geq t_0(\sigma) \) for which experimentation stops in this continuation, and it must satisfy \( t_1(\sigma) = \tau(t_0(\sigma)) \). To see why, suppose that \( t_1(\sigma) > \tau(t_0(\sigma)) \). In this case, for \( \epsilon > 0 \) sufficiently small, \( m_{t_0(\sigma)+\epsilon} \) would strictly prefer to stop experimenting, a contradiction. On the other hand, if \( t_1(\sigma) < \tau(t_0(\sigma)) \), then \( m_{\tau_0(\sigma)} \) would strictly prefer to deviate from the equilibrium path and not stop.

Next, suppose that \( \tau(t_0(\sigma)) = \infty \), that is, \( m_{t_0(\sigma)} \) weakly prefers to continue.

\footnote{The fact that \( W_T(x) \) is differentiable in \( T \) at \( (T, x) = \tau(t) \) and strictly decreasing in \( T \) is implied by Lemma 7.}
experimenting regardless of the continuation. Then it must be that \( t_1(\sigma) = \infty \) and 
\[ V\left(p_{t_0(\sigma)}(m_{t_0(\sigma)})\right) = \xi, \] and in this case we must still have \( t_1(\sigma) = \tau(t_0(\sigma)). \)

Now suppose that \( \tau(t_0(\sigma)) = t_0(\sigma), \) that is, \( m_{t_0(\sigma)} \) weakly prefers to stop regardless of the continuation. In this case, the implied sequence of stopping points is 
\( (t_0(\sigma), t_0(\sigma), \ldots). \) This does not fully describe the equilibrium, as it does not specify what happens conditional on not stopping experimentation by \( t_0(\sigma), \) but still provides enough information to characterize the equilibrium path fully, as in any equilibrium experimentation must stop at \( t_0(\sigma). \)

Next, we show that if \( \tau \) is increasing and \( t \in [0, \tau(0)] \), then \( (t, \tau(t), \tau(\tau(t)), \ldots) \) constitutes an equilibrium. Our construction already shows that \( m_{t_0(\sigma)} \) is indifferent between switching to the safe policy at time \( t_n(\sigma) \) and continuing to experiment. To finish the proof, we have to show that for \( t \) not in the sequence of the stopping times, \( m_t \) weakly prefers to continue experimenting. Fix \( t \in (t_n(\sigma), t_{n+1}(\sigma)). \) Since \( t > t_n(\sigma) \) and \( \tau \) is increasing, we have \( \tau(t) \geq \tau(t_n(\sigma)) = t_{n+1}(\sigma). \) Hence the definition of \( \tau(t) \) and the fact that \( T \mapsto W_T(x) \) is single-peaked by Lemma 7 imply that \( W_{t_{n+1}(\sigma)} - \tau(p_t(m_t)) \geq \xi, \) and Conditions (ii) and (iii) imply that \( m_t \) weakly prefers to continue experimenting.

Finally, we show that even if \( \tau \) is not increasing, this construction yields an equilibrium for at least one value of \( t \in [0, \tau(0)]. \) Note that our construction fails if and only if there is \( t \in (t_k(\sigma), t_{k+1}(\sigma)) \) for which \( \tau(t) < t_{k+1}(\sigma). \) Motivated by this, we say \( t \) is valid if \( \tau(t) = \inf_{\nu \geq \tau} \tau(t'), \) and say \( t \) is \( n \)-valid if \( t, \tau(t), \ldots, \tau^{(n-1)}(t) \) are all valid. Let \( A_0 = [0, \tau(0)] \) and, for \( n \geq 1, \) let \( A_n = \{ t \in [0, \tau(0)] : t \) is \( n \)-valid.\}

Suppose that \( \tau(t) > t \) and \( \tau(t) < \infty \) for all \( t. \) Clearly, \( A_n \supseteq A_{n+1} \) for all \( n, \) and the continuity of \( \tau \) implies that \( A_n \) is closed for all \( n. \) In addition, \( A_n \) must be non-empty for all \( n \) by the following argument. Take \( t_0 = t \) and define a sequence \( \{t_0, t_{-1}, t_{-2}, \ldots, t_{-k}\} \) by \( t_{-i} = \max \{\tau^{-1}(t_{-i+1})\} \) for \( i \leq -1, \) and \( t_{-k} \in [0, \tau(0)]. \) By construction, \( t_{-k} \in A_0 \) is \( k \)-valid, and, because \( \tau(t) < \infty \) for all \( t, \) if we choose \( t \) large enough, we can make \( k \) arbitrarily large.\footnote{Under the assumption that \( \tau(t) < \infty \) for all \( t, \) since \( \tau \) is continuous, the image of \( \tau^l \) restricted to the set \([0, \tau(0)]\) is compact and hence bounded for all \( l. \) Thus for any \( t \) larger than the supremum of this image, \( k \) must be larger than \( l. \)} Then \( A = \cap_0^{\infty} A_n \neq \emptyset \) by Cantor’s intersection theorem, and any sequence \((t, \tau(t), \ldots)\) with \( t \in A \) yields an equilibrium. The same argument goes through if \( \tau(t) = \infty \) for some values of \( t \) but there are
arbitrarily large $t$ for which $\tau(t) < \infty$.

If $\tau(t) = t$ for some $t$, let $\bar{t} = \min\{t \geq 0 : \tau(t) = t\}$. If there is $\epsilon > 0$ such that $\tau(t) \geq \tau(\bar{t})$ for all $t \in (\bar{t} - \epsilon, \bar{t})$, then we can find a finite equilibrium sequence of stopping times by setting $t_0 = \bar{t}$ and using the construction in the previous paragraph. If there is no such $\epsilon$, then the previous argument works.\(^{34}\) The only difference is that, to show the non-emptiness of $A_n$, we take $t \to \bar{t}$ instead of making $t$ arbitrarily large.

If $\tau > t$ for all $t$ and there is $\tilde{t}$ for which $\tau(t) = \infty$ for all $t \geq \tilde{t}$, without loss of generality, take $\tilde{t}$ to be minimal (that is, let $\tilde{t} = \min\{t \geq 0 : \tau(t) = \infty\}$). Then we can find a finite sequence of stopping times compatible with equilibrium by taking $t_0 = \tilde{t}$, assuming that $m_{t_0}$ stops at $t_0$ and using the above construction. \(\blacksquare\)

\section{A Model of Bad News}

\textbf{Lemma 15.} In a model of bad news, the value function of an agent with prior $x$ who is in the organization and expects the organization to continue forever unless a failure is observed is

\[
V(x) = (xb + (1 - x)r)\frac{1}{\gamma} - (1 - x)r\frac{1}{\gamma + b}
\]

\textbf{Proof of lemma 15.}

Note that an agent receives an expected flow payoff of $b$ only if the technology is good and the organization has not switched to the safe technology upon observing a failure. Because a good technology cannot experience a failure, as long as experimentation continues, an agent with posterior belief $x$ receives an expected flow payoff of $b$ with probability $x$.

Let $P_t = x + (1 - x)e^{-bt}$ denote the probability that an agent with prior belief $x$ assigns to not having a failure by time $t$. Note that at each time $t$, the probability that the organization has switched to the safe technology by this time is $1 - P_t = (1 - x)(1 - e^{-bt})$.

\(^{34}\)If there is $\epsilon > 0$ with the aforementioned property, then $\tau^{-1}(\bar{t})$ is strictly lower than $\bar{t}$ and reaching $[0, \tau(0)]$ takes finitely many steps. If there is no such $\epsilon$, then $\tau^{-1}(\bar{t}) = \bar{t}$ and there exists a sequence converging to $\bar{t}$. 

50
Then
\[
V(x) = \int_0^\infty (xb + (1 - P_x)r) e^{-\gamma \tau} d\tau \\
= \int_0^\infty \left( xb + (1 - x)(1 - e^{-b\tau})r \right) e^{-\gamma \tau} d\tau \\
= (xb + (1 - x)r) \frac{1}{\gamma} - (1 - x)r \frac{1}{\gamma + b}
\]
where the last equality follows from the fact that \( \int_0^\infty e^{-\gamma \tau} d\tau = \frac{1}{\gamma} \) and \( \int_0^\infty e^{-(\gamma+b)\tau} d\tau = \frac{1}{\gamma+b} \).

Proof of Proposition 4.

Claim 1. In a model of bad news, if the initial distribution of priors is uniform, then \( p_0(m_0) = \frac{b+s}{2b} \) and \( t \mapsto p_t(m_t) \) is strictly increasing.

Proof of claim 1.

Observe that in a model of bad news, we have \( p_t(y_t) = \frac{y_t}{y_t + e^{-bt}(1 - y_t)} \). Because \( p_t(y_t) = \frac{s}{b} \) must be satisfied, using the formula for \( p_t(y_t) \) and solving for \( y_t \), we obtain
\[
y_t = \frac{s}{s + (b-s)e^{bt}}.
\]
If the density is uniform, the median is given by \( m_t = \frac{1 + y_t}{2} \). Substituting the above formula for \( y_t \) into \( m_t = \frac{1 + y_t}{2} \), we obtain \( m_t = \frac{1}{2} \frac{2s + (b-s)e^{bt}}{s + (b-s)e^{bt}} \).

Substituting the above formula for \( m_t \) into \( p_t(m_t) = \frac{m_t}{m_t + e^{-bt}(1 - m_t)} \), we obtain
\[
p_t(m_t) = \frac{\frac{1}{2} \frac{2s + (b-s)e^{bt}}{s + (b-s)e^{bt}}}{\frac{1}{2} \frac{2s + (b-s)e^{bt}}{s + (b-s)e^{bt}} + e^{-bt} \left( 1 - \frac{\frac{1}{2} \frac{2s + (b-s)e^{bt}}{s + (b-s)e^{bt}}}{2s + (b-s)e^{bt}} \right)}
\]
\[
= \frac{2s + (b-s)e^{bt} + e^{-bt}(2(s + (b-s)e^{bt}) - (2s + (b-s)e^{bt}))}{2s + (b-s)(1 + e^{bt})}
\]
Thus we have \( p_t(m_t) = \frac{2s + (b-s)e^{bt}}{2s + (b-s)(1 + e^{bt})} \). Then \( p_0(m_0) = \frac{b+s}{2b} \).

Moreover, it can be verified that \( t \mapsto p_t(m_t) \) is strictly increasing. In particular, let \( p(e) = \frac{2s + (b-s)e}{2s + (b-s)(1+e)} \). We have \( p'(e) \propto (2s + (b-s)(1 + e))(b-s) - (2s + (b- b}
Suppose first that \( f \) is non-decreasing and \( V(\frac{2s}{b+s}) \geq \frac{r}{\gamma} \). Because \( \frac{2s}{b+s} = \frac{2s}{s+b} > 0 \), the fact that \( V(\frac{2s}{b+s}) \geq \frac{r}{\gamma} \) implies that \( V(\frac{b+s}{2b}) > \frac{r}{\gamma} \). Because, by claim 1, \( p_0(m_0) = \frac{b+s}{2b} \) and \( t \mapsto p_t(m_t) \) is strictly increasing and lemma 15 implies that \( V \) is increasing, this implies that \( V(p_t(m_t)) > \frac{r}{\gamma} \) for all \( t \).

Then, using the fact that that \( f \) is non-decreasing, we can use an argument similar to the one used in the proofs of Propositions 1 and 2 to show that there exists an equilibrium where the organization experiments forever unless a failure is observed. Moreover, an argument similar to the one used in the proofs of Propositions 1 and 2 can be used to show that this equilibrium is unique. In particular, letting \( \tilde{m}_t \) denote the median under a non-decreasing \( f \) and letting \( m_t \) denote the median under the uniform distribution, because \( m_t \mapsto p_t(m) \) is strictly increasing, it is sufficient to show that \( \tilde{m}_t \geq m_t \) for all \( t \), and \( f \) being non-decreasing ensures that \( \tilde{m}_t \geq m_t \) for all \( t \) by the same argument as in the proof of Proposition 2.

We now show that there exists \( T \) such that for all \( t \geq T \), if no failures have been observed during \([0,t]\), then \( V(p_t(m_t)) \geq \frac{r}{\gamma} \). First note that, because in a model of bad news agents do not leave the organization, we have \( \lim_{t \to \infty} m_t > 0 \). Moreover, \( \lim_{t \to \infty} e^{-bt} = 0 \). This implies that \( \lim_{t \to \infty} p_t(m_t) = \lim_{t \to \infty} \frac{m_t}{m_t + e^{-\gamma t}(1-m_t)} = 1 \). Then, provided that no failures have been observed during \([0,t]\), we have \( \lim_{t \to \infty} V(p_t(m_t)) = V(1) \) because \( V \) is continuous, and \( V(1) > \frac{r}{\gamma} \).

Next, observe that two cases are possible: either \( V(p_t(m_t)) \geq \frac{r}{\gamma} \) for all \( t \leq T \), or there exists \( t \leq T \) such that \( V(p_t(m_t)) < \frac{r}{\gamma} \). If \( V(p_t(m_t)) \geq \frac{r}{\gamma} \) for all \( t \leq T \), then the organization experiments forever, so suppose that there exists \( t \leq T \) such that \( V(p_t(m_t)) < \frac{r}{\gamma} \).

Claim 2. Suppose that on the equilibrium path, the organization continues experimenting for time \( t_\perp \) unless a failure occurs and then switches to the safe policy. Then the value function of an agent with prior \( x \) in this equilibrium is given by

\[
(xb + (1-x)r) \frac{1 - e^{-\gamma t_*}}{\gamma} - (1-x)r \frac{1 - e^{-(\gamma+b)t_*}}{\gamma + b} + e^{-\gamma t_*} \frac{r}{\gamma}
\]

Proof of claim 2.
Because the median switches to the safe policy after a time period of length $t_+$, the value function of an agent with prior $x$ in this equilibrium is given by

$$
\int_0^{t_+} (xb + (1 - P_r)r) e^{-\gamma \tau} d\tau + \int_{t_+}^\infty re^{-\gamma \tau} d\tau
$$

$$
= \int_0^{t_+} (xb + (1 - x)(1 - e^{-\epsilon t})r) e^{-\gamma \tau} d\tau + \int_{t_+}^\infty re^{-\gamma \tau} d\tau
$$

$$
= (xb + (1 - x)r) \int_0^{t_+} e^{-\gamma \tau} d\tau - (1 - x)r \int_0^\infty e^{-(\gamma + b)\tau} d\tau + \int_{t_+}^\infty re^{-\gamma \tau} d\tau
$$

$$
= (xb + (1 - x)r) \frac{1 - e^{-\gamma t_+}}{\gamma} - (1 - x)r \frac{1 - e^{-(\gamma + b)t_+}}{\gamma + b} + e^{-\gamma t_+}r
$$

where the last equality follows from the fact that $\int_0^{t_+} e^{-\gamma \tau} d\tau = \frac{1 - e^{-\gamma t_+}}{\gamma}$, $\int_{t_+}^\infty e^{-(\gamma + b)\tau} d\tau = \frac{e^{-\gamma t_+}r}{\gamma}$.

\textbf{Claim 3.} Suppose that in some equilibrium the median $m_t$ stops experimenting. If for all $t \in [\bar{t}, t_0)$ we have $p_t(m_t)b < r$, then for all $t \in [\bar{t}, t_0)$, $m_t$ stops experimenting.

\textbf{Proof of claim 3.}

Suppose for the sake of contradiction that this is not the case. Then there exists a non-empty subset $B \subseteq [\bar{t}, t_0)$ such that for all $t \in B$, $m_t$ continues experimenting.

There are two cases. In the first case, $B$ has a non-empty interior. In this case, for all $\epsilon > 0$ small, there must exist $\tau \in [\bar{t}, t_0)$ such that, starting at time $\tau$, experimentation continues up to time $\tau + \epsilon$ and then stops.\textsuperscript{35}

By claim 2, the payoff to $m_\tau$ from continuing experimentation is $W_{t_2-\tau}(p_\tau(m_\tau)) = (p_\tau(m_\tau)b + (1 - p_\tau(m_\tau)r)\frac{1 - e^{-\gamma \tau}}{\gamma} - (1 - p_\tau(m_\tau)r)\frac{1 - e^{-(\gamma + b)\tau}}{\gamma + b} + e^{-\gamma \tau}r$.

The payoff to stopping experimentation is $\frac{\bar{t}}{\gamma}$. Then if $W_{t_2-\tau}(p_\tau(m_\tau)) < \frac{\bar{t}}{\gamma}$, the median strictly prefers to stop experimenting. This is equivalent to

$$
\left( p_\tau(m_\tau)b + (1 - p_\tau(m_\tau)r) \right) \frac{1 - e^{-\gamma \epsilon}}{\gamma} - (1 - p_\tau(m_\tau)r) \frac{1 - e^{-(\gamma + b)\epsilon}}{\gamma + b} < \frac{1 - e^{-\gamma \epsilon}}{\gamma}r
$$

Equivalently, $p_\tau(m_\tau)(b - r)\frac{1 - e^{-\gamma \epsilon}}{\gamma} < (1 - p_\tau(m_\tau)r)\frac{1 - e^{-(\gamma + b)\epsilon}}{\gamma + b}$. Rearranging, we

\textsuperscript{35}To find such $\tau$, let $\tilde{t}$ be in the interior of $B$, and let $\tilde{t} = \inf\{t \geq \tilde{t} : t \notin B\}$. Then $\tau = \tilde{t} - \epsilon$ works for all $\epsilon > 0$ small enough.
obtain
\[ \frac{p_\tau(m_\tau)}{1 - p_\tau(m_\tau)} \cdot \frac{b - r \gamma + b}{r} < \frac{1 - e^{-(\gamma + b)\epsilon}}{1 - e^{-\gamma \epsilon}} \]

By L'Hospital's rule, when we take the limit as \( \epsilon \to 0 \), we obtain \( \frac{p_\tau(m_\tau)(b-r)\gamma+b}{(1-p_\tau(m_\tau))r} < \frac{\gamma+b}{\gamma} \). Equivalently, \( p_\tau(m_\tau)b < r \). By the hypothesis, we have \( p_t(m_t)b < r \) for all \( t \in [t, t_0) \). Then, because \( \tau \in [t, t_0) \), the inequality \( p_\tau(m_\tau)b < r \) is satisfied. Then \( m_\tau \) strictly prefers to stop experimenting, which is a contradiction.

In the second case, the interior of \( B \) is empty. In this case, the proof follows a similar argument leveraging Condition (iii).

**Claim 4.** Suppose that in some equilibrium the median \( m_{t_0} \) stops experimenting. If for all \( t \in [\frac{t}{t}, t_0) \) we have \( p_t(m_t)b > r \), then in any equilibrium, for all \( t \in [t, t_0) \), \( m_t \) continues experimenting.

**Proof of claim 4.**

Suppose for the sake of contradiction that this is not the case. Then there exists a non-empty subset \( T' \subseteq [\frac{t}{t}, t_0) \) such that for all \( t \in T' \), \( m_t \) stops experimenting. Fix \( t \in T' \). Let \( t_+ \) denote the length of the time period after which the equilibrium prescribes a switch to the safe policy.\(^{36}\)

Note that, because the median switches to the safe policy after a time period of length \( t_+ \), the payoff to \( m_t \) from continuing experimentation is

\[
W_{t_+}(p_t(m_t)) = \int_0^{t_+} (p_t(m_t)b + (1 - P_\tau)r)e^{-\gamma \tau}d\tau + \int_{t_+}^{\infty} re^{-\gamma \tau}d\tau \\
\geq \int_0^{t_+} p_t(m_t)b e^{-\gamma \tau}d\tau + \int_{t_+}^{\infty} re^{-\gamma \tau}d\tau = p_t(m_t)b \frac{1 - e^{-\gamma t_+}}{\gamma} + r \frac{e^{-\gamma t_+}}{\gamma}
\]

The payoff to stopping experimentation is \( \frac{\tau}{\gamma} \). Then if \( \frac{1 - e^{-\gamma t_+}}{\gamma} p_t(m_t)b + \frac{e^{-\gamma t_+}}{\gamma} r > \frac{\tau}{\gamma} \), the median \( m_t \) strictly prefers to continue experimenting. The above inequality is equivalent to \( p_t(m_t)b > r \), which is satisfied. Then \( m_t \) strictly prefers to continue experimenting, which is a contradiction. ■

Let \( t_{2n+1} = \sup \{ t : V(p_t(m_t)) \leq \frac{\tau}{\gamma} \} \) denote the largest time for which the me-

\(^{36}\)We write the argument assuming that \( t_+ > 0 \). If \( t_+ = 0 \), the proof follows a similar argument leveraging Condition (iii).
that, for all \( t < t_{2n+1} \). In this case, by claim 3, for all \( t \leq t_{2n+1} \), \( m_t \) stops experimentation. Then we set \( n = 0 \), \( t_0 = 0 \) and \( I_0 = [t_0, t_1] \).

Suppose next that there exists \( t < t_{2n+1} \) such that \( p_t(m_t)b > r \). Set \( t_{2n} = \sup\{t < t_{2n+1} : p_t(m_t)b > r\} \). Note that, because \( F \) admits a continuous density, \( t \mapsto p_t(m_t) \) is continuous, which implies that we must have \( p_{t_{2n}}(m_{t_{2n}})b - r = 0 \). Then claim 3 implies that for all \( t \in [t_{2n}, t_{2n+1}] \), \( m_t \) stops experimentation.

Let us conjecture a continuation equilibrium path on which, starting at \( t \), the organization experiments until \( t_{2n} \). Recall that \( W_{t_{2n} - t}(x) \) denotes the value function of an agent with belief \( x \) (at time \( t \)) given this continuation equilibrium path. We then let

\[
 t_{2n-1} = \sup \left\{ t < t_{2n} : W_{t_{2n} - t}(p_t(m_t)) \leq \frac{\epsilon}{\gamma} \right\}.
\]

Note that, because, by construction, for \( t \in (t_{2n-1}, t_{2n}) \) we have \( W_{t_{2n} - t}(p_t(m_t)) > \frac{\epsilon}{\gamma} \), the median \( m_t \) continues experimentation for all \( t \in (t_{2n-1}, t_{2n}) \).

Since \( F \) admits a continuous density, \( t \mapsto W_{t_{2n} - t}(p_t(m_t)) \) is continuous, which implies that we must have \( t_{2n-1} = \max \left\{ t < t_{2n} : W_{t_{2n} - t}(p_t(m_t)) \leq \frac{\epsilon}{\gamma} \right\} \). Note that it is then consistent with equilibrium for the median \( m_{t_{2n}} \) to stop experimenting.

Now note that if \( W_{t_{2n} - t_{2n-1}}(p_{t_{2n-1}}(m_{t_{2n-1}})) = \frac{\epsilon}{\gamma} \), then \( p_{t_{2n-1}}(m_{t_{2n-1}})b < r \). By continuity, there exists an interval \([t_i, t_i]\) in \( T_1 \) such that \( t_{2n-1} \in [t_i, t_i] \) (and \( t_i \) satisfies \( t_i = \min\{ t < t_{2n-1} : p_t(m_t)b \leq r \} \)).

Set \( t_{2n-2} = t_i \). Because \( p_t(m_t)b \leq r \) for all \( t \in [t_{2n-2}, t_{2n-1}] \), claim 4 implies that, for all \( t \in [t_{2n-2}, t_{2n-1}] \), \( m_t \) stops experimenting.

We then proceed inductively in the above manner, finding the largest \( t \) strictly less than \( t_{2n-2} \) such that \( W_{t_{2n-2} - t}(p_t(m_t)) \leq \frac{\epsilon}{\gamma} \). Because \( T_1 \) is finite collection of intervals, the induction terminates in a finite number of steps.

The equilibrium is generically unique for the following reason. Under Assumption 1, each \( t_{2k+1} \) satisfies not only \( W_{t_{2k+2}}(p_{t_{2k+1}}(m_{t_{2k+1}})) = \frac{\epsilon}{\gamma} \) but also \( \frac{\partial}{\partial t} W_{t_{2k+2} - t}(p_t(m_t))|_{t=t_{2k+1}} > 0 \), that is, \( W_{t_{2k+2} - t}(p_t(m_t)) < \frac{\epsilon}{\gamma} \) for all \( t \leq t_{2k+1} \) close enough to \( t_{2k+1} \). Thus, even if
we allow \( m_{t2k+1} \) to continue experimenting, all agents in \((t_{2k+1} - \epsilon, t_{2k+1})\) must stop as they strictly prefer to do so. Likewise, each \( t_{2k} \) satisfies not only \( p_{t2k}(m_{t2k})b - r = 0 \) but also \( \frac{\partial}{\partial r} p_{t}(m_{t})|_{t=t_{2k}} < 0 \), that is, \( p_{t}(m_{t})b - r > 0 \) for all \( t < t_{2k} \) close enough to \( t_{2k} \). Thus, even if we allow \( m_{t2k} \) to stop experimenting, all agents in \((t_{2k} - \epsilon, t_{2k})\) must stop as they strictly prefer to do so.

\[ L \]

## B Imperfectly Informative Experimentation Technology

**Lemma 16.** If the organization is experimenting at time \( t \), then an agent with belief \( x \) at time 0 is in the organization at time \( t \) if and only if \( L(k, t) \leq \frac{x(b-s)}{(1-\alpha)(s-c)} \).

**Proof of lemma 16.**

Let \( x_t \) denote the belief at time \( t \) of an agent with belief \( x \) at time 0. Because agents make their membership decisions based on the expected flow payoffs, this agent is in the organization at time \( t \) if and only if \( x_t b + (1-x_t)c \geq s \), that is, if \( x_t \geq \frac{s-c}{b-c} \). Since \( x_t = \frac{x}{x+(1-x)L(k,t)} \), this is equivalent to \( L(k, t) \leq \frac{x(b-s)}{(1-\alpha)(s-c)} \).

**Lemma 17.** If the distribution of priors is power law, then \( L \mapsto p(L, m(L)) \) is decreasing. Moreover, if \( L_0 m'(L_0) < m(L_0)(1-m(L_0)) \), then \( L \mapsto p(L, m(L)) \) is strictly decreasing at \( L = L_0 \), and if \( L_0 m'(L_0) > m(L_0)(1-m(L_0)) \), then \( L \mapsto p(L, m(L)) \) is strictly increasing at \( L = L_0 \).

**Proof of lemma 17.**

The density of the power law distribution is given by \( f(x) = (1-x)^{\alpha}c \) where \( c \) is a constant ensuring that the density integrates to 1. In particular, if the support of the distribution is \([0, 1]\), then we have \( F(z) = \int_0^z (1-x)^{\alpha}c dx = \frac{c}{\alpha+1} (1 - (1 - z)^{\alpha+1}) \). Because \( F(1) = 1 \), we must have \( c = \alpha + 1 \). Then \( F(z) = 1 - (1 - z)^{\alpha+1} \) and the CDF of the distribution with support on \([y, 1]\) is given by \( \frac{(1-y)^{\alpha+1} - (1-z)^{\alpha+1}}{(1-y)^{\alpha+1}} \).

Recall that \( m(L) \) and \( y(L) \) denote the median and the marginal members of the organization respectively when the state variable is \( L \). The above argument implies that the median must satisfy \( \frac{(1-y(L))^{\alpha+1} - (1-m(L))^{\alpha+1}}{(1-y(L))^{\alpha+1}} = \frac{1}{2} \). Equivalently, we must have \( (1-m(L))^{\alpha+1} = \frac{1}{2} (1-y(L))^{\alpha+1} \). Then the median must satisfy \( 1-m(L) = \)
\[
(1 - y(L))2^{-\frac{1}{\alpha+1}} \text{, or } m(L) = 1 - \kappa + \kappa y(L) \text{ for } \kappa = 2^{-\frac{1}{\alpha+1}}.
\]

Note that \( p(L, m(L)) = \frac{1}{1+(m(L)-1)L} \). Then \( \frac{\partial}{\partial L} p(L, m(L)) \propto -\frac{\partial}{\partial L} \left( 1 + \left( \frac{1}{m(L)} - 1 \right) L \right) \) and \( \frac{\partial}{\partial L} \left( 1 + \left( \frac{1}{m(L)} - 1 \right) L \right) = \frac{\partial}{\partial L} \left( \frac{1}{m(L)} - 1 \right) L = \frac{1}{m(L)} - 1 - \frac{L}{(m(L))^2} m'(L). \)

This implies that if \( L_0 m'(L_0) < m(L_0)(1-m(L_0)) \), then \( L \mapsto p(L, m(L)) \) is strictly decreasing at \( L = L_0 \), and if \( L_0 m'(L_0) > m(L_0)(1-m(L_0)) \), then \( L \mapsto p(L, m(L)) \) is strictly increasing at \( L = L_0 \).

After some algebra, using the fact that \( y(L) = \frac{s-c}{s-c+(b-s)L} \), we get that if the distribution of priors is power law, then \( L m'(L) < m(L)(1-m(L)) \) is equivalent to \( 0 < (1-\kappa)(1-\zeta) \), where \( \zeta = \frac{s-c}{b-c} \). Since \( \kappa \) and \( \zeta \) are between 0 and 1, this always holds. \( \blacksquare \)

**Lemma 18.** There exist distributions for which there exist states \( L_1 < L_2 \) such that \( L_1 \) is a unique minimizer of \( p(L, m(L)) \) and \( L \mapsto p(L, m(L)) \) is strictly increasing on \((L_1, L_2)\).

**Proof of lemma 18.**

Consider a distribution with a density \( f(x) = a_1 \) for \( x \in [0, b_1] \) and \( f(x) = a_2 \) for \( x \in [b_1, 1] \). Note that we must have \( a_1 b_1 + a_2(1-b_1) = 1 \) to ensure that \( f \) integrates to 1. Define \( \bar{b} = b - c, \bar{s} = s - c, y = y(L), m = m(L), z = p(L, m(L)) \). Let \( L_1 \) be such that \( m(L_1) = b_1 \) and let \( L_2 \) be such that \( y(L_2) = b_1 \). Clearly, \( 0 < L_1 < L_2 \).

For \( L > L_2, m(L) \) and \( p(L, m(L)) \) are the same as in the uniform case. In particular, \( p(L, m(L)) = \frac{2L \bar{s} + L \bar{b} - L \overline{s+b}}{L \bar{s} + b - s} \), which is decreasing in \( L \). Moreover, with the notation we have defined, the formula for \( y(L) \) can be written as \( y = \frac{L \bar{s}}{L \bar{s} + b - s} \).

For \( L \in (L_1, L_2) \), we have \( a_1(b_1 - y) + a_2(m - b_1) = a_2(1 - m) \), that is, \( m = \frac{1+b_1}{2} - \frac{a_1 b_1}{2a_2} + \frac{a_1}{2a_2} y \). Equivalently, \( m = \left( 1 - \frac{1}{2a_2} \right) + \frac{a_1}{2a_2} y = \left( 1 - \frac{1}{2a_2} \right) + \frac{a_1}{2a_2} \frac{L \bar{s}}{L \bar{s} + b - s} \). Then

\[
1 - \frac{1}{z} = \frac{L(1-m)}{m} = L \frac{L \frac{1-a_1 \bar{s}}{2a_2} + \frac{1}{2a_2} \left( \bar{b} - \bar{s} \right)}{\left( 1 - \frac{1}{2a_2} + \frac{a_1}{2a_2} \right) \frac{L \bar{s}}{L \bar{s} + b - s} + \left( 1 - \frac{1}{2a_2} \right) \left( \bar{b} - \bar{s} \right)}
\]

For \( L < L_1 \), we have \( a_1(m - y) = a_1(b_1 - m) + a_2(1 - b_1) \), that is, \( 2a_1 m = \ldots \)
\[ a_1 b_1 + a_2 (1 - b_1) + a_1 y = 1 + a_1 y, \text{ so } m = \frac{1}{2a_1} + \frac{1}{2} y, \text{ and} \]

\[
\frac{1}{z} - 1 = \frac{L(1 - m)}{m} = L \left( \frac{1}{z} - \frac{1}{2a_1} \right) \bar{s} + \left( 1 - \frac{1}{2a_1} \right) (\bar{b} - \bar{s}) \left( \frac{1}{2a_1} + \frac{1}{2} \right) L \bar{s} + \frac{1}{2a_1} (\bar{b} - \bar{s}) \right).
\]

Now take \( a_2 = \frac{1}{2} \) and any \( a_1 > 1 \) (note that choosing both pins down \( b_1 = \frac{1}{2a_1 - 1} \)). Then we can verify that \( L \mapsto \frac{1}{p(L, m(L))} - 1 \) is increasing on \((0, L_1)\) and decreasing on \((L_1, L_2)\). In other words, \( L \mapsto p(L, m(L)) \) is decreasing on \((0, L_1)\) and \((L_2, \infty)\) but increasing on \((L_1, L_2)\), so \( L_1 \) is a local minimizer for \( p(L, m(L)) \).

Moreover, we can verify that under some extra conditions \( L_1 \) is a global minimizer: note that \( \lim_{L \to \infty} \frac{1}{p(L, m(L))} - 1 = \frac{\bar{b} - \bar{s}}{2\bar{s}} \), while \( \frac{1}{p(L_1, m(L_1))} - 1 = \frac{L_1 (1 - a_1) \bar{s} + \bar{b} - \bar{s}}{a_1 \bar{s}} \).

Since \( m(L_1) = b_1 \), we have

\[
\frac{1}{p(L_1, m(L_1))} - 1 = \frac{L_1}{m(L_1)} - L_1 = \frac{L_1}{b_1} - L_1 = \frac{L_1 (1 - a_1) \bar{s} + \bar{b} - \bar{s}}{a_1 \bar{s}}
\]

\[
L_1 = \frac{\bar{b} - \bar{s}}{\bar{s} \left( \frac{a_1}{b_1} - 1 \right)}
\]

\[
\frac{1}{p(L_1, m(L_1))} - 1 = \frac{L_1}{b_1} - L_1 = \frac{\bar{b} - \bar{s}}{\bar{s} \left( \frac{a_1}{b_1} - 1 \right)} = \frac{\bar{b} - \bar{s}}{\bar{s} \left( a_1 - b_1 \right)}
\]

\[
= \frac{\bar{b} - \bar{s}}{\bar{s}} \frac{2a_1 - 2}{a_1^2 - a_1 - 1} = \frac{\bar{b} - \bar{s}}{\bar{s}} \frac{1}{a_1 + \frac{1}{2}}
\]

so \( L_1 \) is a global minimizer if we take \( a_1 \in \left( 1, \frac{3}{2} \right) \).

**Proof of Proposition 5.**

**Claim 1.** For a uniform distribution of priors, \( \lim_{L \to \infty} p(L, m(L)) = \frac{2(s-c)}{(b-c)+(s-c)} \) and \( L \mapsto p(L, m(L)) \) is decreasing.

**Proof of claim 1.**

Observe that \( p(L, y(L)) \) is the posterior belief of the marginal member of the organization when the state variable is \( L \). Because marginal agents make their membership decisions based on the flow payoffs, \( y(L) \) satisfies \( p(L, y(L)) b + (1 - p(L, y(L))) c = s \). Equivalently, \( p(L, y(L)) = \frac{s-c}{b-c} \). Because \( p(L, y(L)) = \frac{\gamma(L)}{\gamma(L) + (1 - y(L)) L} \), this is equiva-
lent to \( \frac{y(L)}{y(L)+(1-y(L))L} = \frac{s-c}{b-c} \). Solving for \( y(L) \), we obtain \( y(L) = \frac{s-c}{s-c+(b-s)\frac{L}{x}} \).

Because \( y(L) = \frac{s-c}{s-c+(b-s)\frac{L}{x}} \) and, for a uniform distribution, we have \( m(L) = \frac{1+\gamma(L)}{2} \), substituting the formula for \( y(L) \) into the formula for \( m(L) \) for a uniform distribution, we obtain \( m(L) = \frac{1}{2} \frac{2L(s-c)+b-s}{(s-c)+b-s} \).

Because \( p(L,m(L)) = \frac{1}{1+(\frac{1}{m(L)}-1)L} \), substituting the above formula for \( m(L) \) into the formula for \( p(L,m(L)) \), we obtain \( p(L,m(L)) = \frac{2L(s-c)+b-s}{L(2(s-c)+b-s)+b-s} \). Then \( \lim_{L \to \infty} p(L,m(L)) = \lim_{L \to \infty} \frac{2L(s-c)+b-s}{L(2(s-c)+b-s)+b-s} = \frac{(b-c)+(s-c)}{(b-c)+(s-c)} \).

By Lemma 17, if the distribution of priors is power law, that is, \( f(x) = (1-x)^\alpha c \), then \( L \mapsto p(L,m(L)) \) is decreasing. In particular, this applies to the uniform distribution if we take \( \alpha = 0 \) and \( c = 1 \).

The rest of the proof is then similar to the proof for the baseline model (Propositions 1 and 2). In particular, because, by Lemma 4, \( x \mapsto V(x) \) is strictly increasing and, by claim 1, \( L \mapsto p(L,m(L)) \) is decreasing for a uniform distribution of priors, \( L \mapsto V(p(L,m(L))) \) is decreasing for a uniform distribution. Thus to ensure that \( V(p(L,m(L))) \geq \frac{\gamma}{\gamma} \) for all \( L \), it is enough to ensure that \( \lim_{L \to \infty} V(p(L,m(L))) \geq \frac{\gamma}{\gamma} \). Because, by claim 1, \( \lim_{L \to \infty} p(L,m(L)) = \frac{2(s-c)}{(b-c)+(s-c)} \) and \( L \mapsto V(p(L,m(L))) \) is continuous (because \( x \mapsto V(x) \) is continuous and \( L \mapsto p(L,m(L)) \) is continuous for a uniform distribution), it is enough to ensure that \( V \left( \frac{2(s-c)}{(b-c)+(s-c)} \right) \geq \frac{\gamma}{\gamma} \).

Next, given that we have shown that \( V \left( \frac{2(s-c)}{(b-c)+(s-c)} \right) \geq \frac{\gamma}{\gamma} \) implies that \( V(p(L,m(L))) \geq \frac{\gamma}{\gamma} \) for all \( L \) for a uniform distribution of priors, by an argument similar to the one in the proof of Proposition 2 the hypothesis that \( f \) is non-decreasing ensures that that we have \( V(p(L,m(L))) \geq \frac{\gamma}{\gamma} \) for all \( L \) under \( f \).

To show that if \( V \left( \frac{2(s-c)}{(b-c)+(s-c)} \right) < \frac{\gamma}{\gamma} \), there is no equilibrium in which the organization experiments forever, we can use the result in Lemma 10 that \( \lim_{L \to \infty} p(L,m(L)) = \frac{2(s-c)}{(b-c)+(s-c)} \) and an argument similar to the one used in the proof of Proposition 2.

Finally, we show that there exist parameter values such that \( V \left( \frac{2(s-c)}{(b-c)+(s-c)} \right) \geq \frac{\gamma}{\gamma} \) is satisfied. Note that, because an agent can always leave the organization, her payoff in an equilibrium in which the organization experiments forever on the equilibrium path is bounded below by her payoff from staying in the organization forever. If she stays in the organization forever, she gets a payoff of \( b \) forever if the risky technology
is good and a payoff of \( c \) forever if the risky technology is good. Then \( V(x) \geq x^b + (1 - x)^c \), which implies that \( V\left(\frac{2(s-c)}{(b-c)+(s-c)}\right) \geq \frac{1}{\gamma} \frac{(b-c)s+(s-c)b}{(b-c)+(s-c)} \), as required.

Then to show that there exist parameter values such that \( V\left(2(s-c)b,(s-c)+s\right) \geq \frac{r}{\gamma} \) it is sufficient to check that there exist parameter values such that \( \frac{(b-c)s+(s-c)b}{(b-c)+(s-c)} \geq r \). In general, for any values of \( b, s \) and \( c \) satisfying \( b > s > c > 0 \), there is \( r^*(b, s, c) \) such that the condition holds if \( r \leq r^*(b, s, c) \), and, moreover, \( r^*(b, s, c) \in (s, b) \).

**Proof of Proposition 6.**

For convenience, we multiply all the value functions in this proof by \( \gamma \). Let \( V_{x,\theta}(L) \) denote the value function of an agent with prior belief \( x \) given that the state is \( L(k, t) = L \), the technology is \( \theta \in \{G, B\} \) and the behavior on the equilibrium path is as described in the Proposition. Let \( V_{x}^e(L) \) denote the value function of an agent with prior belief \( x \) given that the state is \( L \) and the behavior on the equilibrium path is as described in the Proposition. The value function of the median is then given by

\[
V_{m(L)}^e(L) = p(L, m(L))V_{m(L),G}^e(L) + (1 - p(L, m(L)))V_{m(L),B}^e(L)
\]

By Proposition 5, there exist parameters such that there is an equilibrium in which the organization experiments forever. Note that \( V_{x}^0(L) \) is the value function of an agent with prior \( x \) when the state is \( L \) in the equilibrium in which the organization experiments forever.

We claim that we can choose the density \( f \) such that there is a unique global minimum of \( L \mapsto V_{m(L)}^0(L) \), which we will call \( L^* \). Because, by Corollary 2, \( V_{m(L)}^0(L) = V_{p(L,m(L))}^0(0) \) and \( x \mapsto V_x^0(0) \) is strictly increasing, it is enough to show that there exists a density such that the minimum of the posterior of the median \( p(L, m(L)) \) over \( L \) is a singleton. This follows from Lemma 18.\(^{37}\)

Thus we fix \( f \) constructed in the proof of Lemma 18. Observe that the construction is such that \( L \mapsto p(L, m(L)) \) has a kink at \( L^* \).

Note that \( V_{m(L)}^0(L) \) does not depend on \( r \) and the density \( f \) constructed in the proof of Lemma 18 does not depend on \( r \), which implies that \( L^* \) does not depend on\(^{37}\)

\(^{37}\)Technically, we also need the condition that \( V_{m(L^*)}^0(L^*) < \lim_{L \to \infty} V_{m(L)}^0(L) \), but this is also satisfied by the construction in Lemma 18.
Then we can choose \( r \) such that

\[
V^0_{m(L^*)}(L^*) = r
\]

(4)

Note that then, because \( L^* \) is the unique minimizer of \( L \mapsto V^0_{m(L)}(L) \), we have \( V^0_{m(L)}(L) > r \) for all \( L \neq L^* \).

We aim to show that if we change the equilibrium to require that experimentation stops at \( L = L^* \) with an appropriately chosen probability \( \epsilon > 0 \), the constraints \( V^\epsilon_{m(L^*)}(L^*) = r \) and \( V^\epsilon_{m(L)}(L) \geq r \) for all \( L \neq L^* \) still hold.

It is useful to note at this point that the value function can be written recursively. Towards this end, we introduce the following notation. For any strategy profile and any \( L, L' \in \mathbb{R} \), define

\[
T_{x,L'}(L) = \int_0^\infty \gamma e^{-\gamma t} Pr[\exists s \in [0, t]: L_s = L' | L_0 = L] dt
\]

\[
\hat{V}_{x,L'}(L) = \int_0^\infty \gamma e^{-\gamma t} E[u_x(h^t) | \exists s \in [0, t]: L_s = L' | L_0 = L] dt
\]

\[
\tilde{V}_{x,L'}(L) = \frac{\hat{V}_{x,L'}(L)}{1 - T_{x,L'}(L)}
\]

where \( u_x(h^t) \) is agent \( x \)'s flow payoff at time \( t \) and history \( h^t \) and the expectation is taken with respect to the stochastic process induced by the equilibrium strategy and the stochastic process \( (\tilde{L}_\tau)_{\tau} \).

Intuitively, \( T_{x,L'}(L) \) is the weighted discounted probability that the stochastic process \( (L_s)_s \) hits the value \( L' \) at least once, \( \hat{V}_{x,L'}(L) \) is the expected utility of agent \( x \) starting with \( L_0 = L \) but setting the continuation value to zero when \( (L_s)_s \) hits \( L' \), and \( \tilde{V}_{x,L'}(L) \) is a normalization.

Then the value function can be written recursively as

\[
V_x(L) = (1 - T_{x,L'}(L))\tilde{V}_{x,L'}(L) + T_{x,L'}(L)V_x(L')
\]

Taking \( L' = L^* \), this implies that for any \( \epsilon \in [0, 1] \), we have

\[
V^\epsilon_x(L) = (1 - T_{x,L^*}(L))\tilde{V}_{x,L^*}(L) + T_{x,L^*}(L)V^\epsilon_x(L^*)
\]

(5)
where $T_{x,L^*}(L)$ is independent of $\epsilon$, since changing $\epsilon$ has no impact on the policy path except when $L = L^*$.

Let

$$T_{x,L^*}(L^+) = \int_0^\infty \gamma e^{-\gamma t} \Pr[\exists s \in (0,t] : L_s = L' | L_0 = L] dt$$

$$\hat{V}_{x,L^*}(L^+) = \int_0^\infty \gamma e^{-\gamma t} E[u_x(h^t)|\exists s \in (0,t] : L_s = L' | L_0 = l] dt$$

and

$$\hat{V}_{x,L^*}(L^+) = \frac{\hat{V}_{x,L^*}(L^+)}{1 - T_{x,L^*}(L^+)}$$

Observe that $\tilde{W}_x(L^*) = \lim_{L \searrow L^*} \hat{V}_{x,L^*}(L)$ and $T_{x,L^*}(L^*) = \lim_{L \searrow L^*} T_{x,L^*}(L)$. Let $\tilde{W}_x = \hat{V}_{x,L^*}(L^*)$.

Let $W_x = \lim_{L \searrow L^*} V_x(L)$. $W_x$ is the expected continuation value of agent $x$ when $L = L^*$ and the median member, $m(L^*)$, has just decided not to stop experimenting. This is closely related to $V_x(L^*)$, which is the expected continuation value where the expectation is taken before $m(L^*)$ has decided whether to stop experimenting or not. Specifically, we have

$$V_x(L^*) = \epsilon r + (1 - \epsilon) W_x = \epsilon r + (1 - \epsilon) \left( (1 - T_{x,L^*}(L^*)) \tilde{W}_x + T_{x,L^*}(L^*) V_x(L^*) \right)$$

Solving this for $V_x(L^*)$, we obtain

$$V_x(L^*) = \frac{\epsilon r + (1 - \epsilon)(1 - T_{x,L^*}(L^*)) \tilde{W}_x}{1 - (1 - \epsilon) T_{x,L^*}(L^*)} = V_x^0(L^*) + \epsilon \frac{r - V_x^0(L^*)}{1 - (1 - \epsilon) T_{x,L^*}(L^*)}$$

(6)

where the second equality follows from the fact that $\tilde{W}_x = V_x^0(L^*)$ because $\tilde{W}_x$ is the continuation value of the agent conditional on the event that $(L_s)_s$ never hits $L^*$ again, which means that in this case experimentation continues forever.
Hence, substituting (6) into (5), we obtain

\[
V^\epsilon_x(L) = (1 - T_{x,L^*}(L)) \tilde{V}_{x,L^*}(L) + T_{x,L^*}(l) \left( V^0_x(L^*) + \epsilon \frac{r - V^0_x(L^*)}{1 - (1 - \epsilon)T_{x,L^*}(L^*)} \right) \\
= T_{x,L^*}(l)\epsilon \frac{r - V^0_x(L^*)}{1 - (1 - \epsilon)T_{x,L^*}(L^*)} + V^0_x(L) 
\]

(7)

At the same time, there exists \( \delta > 0 \) sufficiently small such that for all \( L \in (L^* - \delta, L^* + \delta) \) there exists \( K > 0 \) satisfying

\[
V^0_{m(L)}(L) = V^0_{p(L,m(L))}(0) \\
= V^0_{p(L^*,m(L^*))}(0) + (p(L, m(L)) - p(L^*, m(L^*))) \frac{\partial}{\partial x} V^0_x(0)|_{x=p(L,m(L))} \\
\geq V^0_{p(L^*,m(L^*))}(0) + K|L - L^*| = r + K|L - L^*| 
\]

where the first equality follows from Corollary 2, the second equality follows from using the Mean Value Theorem, the inequality follows from the fact that \( \frac{\partial}{\partial x} V^0_x(0) > 0 \) by Lemma 4 and the fact that \( L \mapsto p(L, m(L)) \) has a kink at \( L^* \), the last equality follows from (4) and Corollary 2.

On the other hand, for \( L \not\in (L^* - \delta, L^* + \delta) \) there exists \( K' > 0 \) such that

\[
V^0_{m(L)}(L) = V^0_{p(L,m(L))}(0) \geq V^0_{p(L^*,m(L^*))}(0) + K' = r + K' 
\]

(9)

where the first equality follows from Corollary 2, the inequality follows from the fact that \( p(L, m(L)) - p(L^*, m(L^*)) \) is bounded away from zero in this case, and the second equality follows from (4) and Corollary 2.

(7) implies that \( V^\epsilon_{m(L)}(L) \geq r \) is equivalent to \( V^0_{m(L)}(L) \geq r - T_{m(L),L^*}(L)\epsilon \frac{r - V^0_{m(L)}(L^*)}{1 - (1 - \epsilon)T_{m(L),L^*}(L^*)} \). If \( V^0_{m(L)}(L^*) - r \leq 0 \), then we are done, so suppose that \( V^0_{m(L)}(L^*) - r > 0 \).

Suppose that \( L \in (L^* - \delta, L^* + \delta) \). Then, because \( V^0_{m(L)}(L) \geq r + K|L - L^*| \) by (8), it is enough to show that \( r + K|L - L^*| \geq r - T_{m(L),L^*}(L)\epsilon \frac{r - V^0_{m(L)}(L^*)}{1 - (1 - \epsilon)T_{m(L),L^*}(L^*)} \). This is equivalent to \( K|L - L^*| \geq T_{m(L),L^*}(L)\epsilon \frac{V^0_{m(L)}(L^*) - r}{1 - (1 - \epsilon)T_{m(L),L^*}(L^*)} \).

Observe that \( 1 - (1 - \epsilon)T_{m(L),L^*}(L^*) \geq 1 - T_{m(L),L^*}(L^*) \). Then, because

\[
\frac{T_{m(L),L^*}(L)}{1 - T_{m(L),L^*}(L^*)} \epsilon (V^0_{m(L)}(L^*) - r) \geq T_{m(L),L^*}(L)\epsilon \frac{V^0_{m(L)}(L^*) - r}{1 - (1 - \epsilon)T_{m(L),L^*}(L^*)},
\]

it is enough to show that \( K|L - L^*| \geq \frac{T_{m(L),L^*}(L)}{1 - T_{m(L),L^*}(L^*)} \epsilon (V^0_{m(L)}(L^*) - r) \). Then we need to show that \( \epsilon \leq \)
we require that $\epsilon \leq \inf_{L \in (L^* - \delta, L^* + \delta)} K \frac{|L - L^*|}{T_m(L, L^*)}$. Observe that $r = V^0_m(L^*)$ by (4). Then we require that

$$
\epsilon \leq \inf_{L \in (L^* - \delta, L^* + \delta)} K \frac{1 - T_m(L, L^*)}{T_m(L, L^*)} \frac{|L - L^*|}{V^0_m(L^*) - V^0_m(L^*)}
$$

Because $T_m(L, L^*) \leq 1$, it is enough to show that $\epsilon \leq \inf_{L \in (L^* - \delta, L^* + \delta)} K (1 - T_m(L, L^*)) \frac{|L - L^*|}{V^0_m(L^*) - V^0_m(L^*)}$.

Suppose next that $L \notin (L^* - \delta, L^* + \delta)$. Then, because $V^0_m(L^*) \geq r + K'$ by (9), it is enough to show that $r + K' \geq r - T_m(L, L^*) \epsilon \frac{V^0_m(L^*)}{1 - (1 - \epsilon) T_m(L, L^*)}$. This is equivalent to $K' \geq T_m(L, L^*) \epsilon \frac{V^0_m(L^*)}{1 - T_m(L, L^*) (L^*) - r}$. Then, because $\frac{T_m(L, L^*)}{|L - L^*|} \epsilon \frac{V^0_m(L^*)}{1 - T_m(L, L^*) (L^*) - r}$, it is enough to show that $K' \geq \frac{T_m(L, L^*)}{|L - L^*|} \epsilon \frac{V^0_m(L^*)}{1 - T_m(L, L^*) (L^*) - r}$. Thus we need to show that $\epsilon \leq \inf_{L \notin (L^* - \delta, L^* + \delta)} K' \frac{1 - T_m(L, L^*)}{T_m(L, L^*)} \frac{1 - \epsilon}{b - r}$.

Since $T_m(L, L^*) \leq 1$, it is sufficient to show that $\epsilon \leq \inf_{L \notin (L^* - \delta, L^* + \delta)} K'(1 - T_m(L, L^*)) \frac{1}{b - r}$.

Let $\epsilon_1 = \inf_{L \notin (L^* - \delta, L^* + \delta)} K \frac{1 - T_m(L, L^*)}{T_m(L, L^*)} \frac{|L - L^*|}{V^0_m(L^*) - V^0_m(L^*)}$ and $\epsilon_2 = \inf_{L \notin (L^* - \delta, L^* + \delta)} K' \frac{1 - T_m(L, L^*)}{T_m(L, L^*)} \frac{1}{b - r}$. To have $\min\{\epsilon_1, \epsilon_2\} > 0$ we need to verify that $\inf_{L \notin (L^* - \delta, L^* + \delta)} T_{x, L^*} (L^*) < 1$ and that $\sup_{L \in (L^* - \delta, L^* + \delta)} \frac{V^0_m(L^*) - V^0_m(L^*)}{L - L^*}$ is finite. The fact that $\sup_{L \in (L^* - \delta, L^* + \delta)} \frac{V^0_m(L^*) - V^0_m(L^*)}{L - L^*}$ is finite follows from the fact that $\frac{\partial}{\partial x} V^0(m(L^*))$ and $m'(L)$ are bounded.

Then choosing $\epsilon \in (0, \min\{\epsilon_1, \epsilon_2\})$ delivers the result.

**Proof of Corollary 1.**

Take the example constructed in Proposition 6, and assume that $L_0 > L^*$.\(^{38}\)

---

\(^{38}\)Technically, our definition of $L_0$ requires that $L_0 = 1$, but we can relax this assumption by considering a continuation of the game starting at some $t_0 > 0$, where, by assumption, the number of successes at time $t_0$ is such that the state variable at $t_0$ is $L_0$. This example can be fit into our original framework by redefining the density of prior beliefs $f$ to be the density of the posteriors held by agents when $L = L_0$ and $f$ is as in Proposition 6. With this relabeling, $L_0$ would equal 1 and $L^*$ would shift to some value less than 1. We find it is easier to think in terms of shifting $L_0$
Let \( P_\theta(L_0) \) be the probability that, conditional on starting at \( L_0 \) and the state being \( \theta \in \{G, B\} \), the organization stops experimenting at any finite time \( t < \infty \). We will show that \( P_G(L_0) > P_B(L_0) \) for \( L_0 \) large enough. In fact, we will prove a stronger result: we will show that there is \( C > 0 \) such that \( P_G(L_0) \geq C > 0 \) for all \( L_0 > L^* \), but \( \lim_{L_0 \to \infty} P_B(L_0) = 0 \).

Let \( Q_\theta(L_0, L^*) \) denote the probability that there exists \( t < \infty \) such that \( L_t \in ((\frac{c}{b}) L^*, L^*) \) when the state is \( \theta \in \{G, B\} \). \( Q_\theta(L_0, L^*) \) is the probability that \( L_t \) ever crosses over to the left of \( L^* \).

We claim that \( Q_G(L_0, L^*) = 1 \) for all \( L_0 > L^* \) but \( \lim_{L_0 \to \infty} Q_B(L_0, L^*) = 0 \).

Let \( l(k, t) = \ln L(k, t) \), and note that \( l(k, t) = \ln \left( \left( \frac{c}{b} \right)^k e^{(b-c)t} \right) = k \ln \left( \frac{c}{b} \right) + \ln \left( e^{(b-c)t} \right) = k(\ln(c) - \ln(b)) + (b-c)t \). Let \( l_0 = \ln(L_0) \).

When \( \theta = G \), we then have \( (l_t)_t = l_0 + (b-c)t - [\ln(b) - \ln(c)]N(t) \), where \( (N(t))_t \) is a Poisson process with rate \( b \), that is, \( N(t) \sim P(bt) \). This can be written as a random walk: for integer values of \( t \), \( l_t - l_0 = \sum_{i=0}^{t} S_i \), where \( S_i = b-c-\ln(b) - \ln(c) \) and \( N_t \sim P(b) \) are iid. Note that \( E[S_i] = b-c-\ln(b) - \ln(c) < 0 \). Then, by the strong law of large numbers, we have \( \frac{l_t}{t} \xrightarrow{t \to \infty} E[S_i] < 0 \) a.s., whence \( l_t \xrightarrow{t \to \infty} -\infty \) a.s., implying the first claim.

On the other hand, when \( \theta = B \), we have \( (l_t)_t = l_0 + (b-c)t - \ln(b-c)N(t) \), where \( (N(t))_t \) is a Poisson process with rate \( c \). This can be written as a random walk with positive drift: \( l_t - l_0 = \sum_{i=0}^{t} S_i \), where \( S_i = b-c-\ln(b) - \ln(c) \) and \( N_t \sim P(c) \), and \( E[S_i] = b-c-c(\ln(b) - \ln(c)) > 0 \). As above, by the strong law of large numbers, we have \( l_t \xrightarrow{t \to \infty} \infty \) a.s.

Note that \( Q_B(L, L^*) = q \) is independent of \( L \). Now suppose for the sake of contradiction that \( \lim \sup_{L \to \infty} Q_B(L, L^*) > 0 \). We claim that this implies \( q = 1 \). Suppose towards a contradiction that \( q < 1 \). Fix \( J \in \mathbb{N} \). Then, for \( L_0 \) large enough.

---

Footnote: Let \( \frac{b}{c} = 1 + x \). Then \( b-c-b(\ln(b) - \ln(c)) = c(x-(1+x)\ln(1+x)) \), where \( x-(1+x)\ln(1+x) \) is negative for all \( x > 0 \). Similarly, \( b-c-c(\ln(b) - \ln(c)) = c(x-\ln(1+x)) \), where \( x-\ln(1+x) \) is positive for all \( x > 0 \).
that \((\frac{c}{b})^{2J+1} L_0 > L^*\),

\[
Q_B(L_0, L^*) \leq \prod_{j=0}^{J} Q_B \left( \left( \frac{c}{b} \right)^{2j} L_0, \left( \frac{c}{b} \right)^{2j+1} L_0 \right) = q^{J+1}.
\]

This implies that, whenever \(\limsup_{L \to \infty} Q_B(L, L^*) > 0\), \(q = 1\).

Hence \((l_t)\) is recurrent (Durrett 2010: pp. 190–201), but this contradicts the fact that \(\lim_{t \to \infty} l_t = \infty\) a.s. Therefore, \(\limsup_{L \to \infty} Q_B(L, L^*) = 0\).

This implies that \(P_B(L_0) \leq Q_B(L_0, L^*) \to 0\) as \(L_0 \to \infty\). On the other hand, \(P_G(L_0) \geq Q_G(L_0, L^*) \inf_{L \in ([\frac{c}{b}]L^*, L^*]} P_G(L) > 0\). The first inequality holds for the following reason. With probability 1, if \(L_t = L^*\) for some \(t\), there must be \(t' < t\) such that \(L_{t'} \in ([\frac{c}{b}]L^*, L^*)\), which happens with probability \(Q_G(L_0, L^*)\). Conditional on this event, the probability of hitting state \(L^*\) in the continuation is \(P_G(L_{t'})\). Note that \(\inf_{L \in ([\frac{c}{b}]L^*, L^*]} P_G(L) > 0\) because it is equal to \(P_G((\frac{c}{b}) L^*)\).

\[\blacksquare\]

Proof of Proposition 7.

Fix an equilibrium \(\sigma\) in which organization experiments forever and let \(\mu_t\) be the size of the organization at time \(t\) on the equilibrium path. Let \(g_t = g(\mu_t)\). The first success that happens at time \(t\) yields the per-capita payoff of \(g_t\), and all further successes pay 1 (because all agents enter the organization after the first success).

Let \(P_t = 1 - e^{-bt}\) denote the probability that there is a success by time \(t\) given that the risky technology is good. In the above problem, an agent with belief \(x\) who expects experimentation to continue forever has utility

\[
V_{(g_t)}(x) = x \int_0^{t^*} e^{-\gamma t} \left( P_t b + (1 - P_t) g_t \right) dt + x \int_{t^*}^{\infty} e^{-\gamma t} \left( P_t b + (1 - P_t) s \right) dt + (1 - x) \int_{t^*}^{\infty} e^{-\gamma t} s dt
\]

where \(t^*\) is the time at which the agent leaves, that is, the time when her posterior reaches \(\frac{s}{g_t b}\).

Now consider the case in which \(g_t = g\) for all \(t\). Then the above expression is
equivalent to

\[
V_g(x) = x \int_0^{t^*} e^{-\gamma t} \left( (1 - e^{-bt}) b + e^{-bt} gb \right) dt \\
+ x \int_{t^*}^{\infty} e^{-\gamma t} \left( (1 - e^{-bt}) b + e^{-bt} s \right) dt + (1 - x) \int_{t^*}^{\infty} e^{-\gamma t} s dt \\
= x \left( \frac{b}{\gamma} - \frac{b}{\gamma + b} + \frac{gb (1 - e^{-(\gamma + b)t^*}) - e^{-\gamma t^*} s}{\gamma + b} \right) + (1 - x) \frac{e^{-\gamma t^*} s}{\gamma} \\
= x \left( \frac{b}{\gamma} - \frac{b}{\gamma + b} + \frac{gb (1 - e^{-(\gamma + b)t^*}) - e^{-\gamma t^*} s}{\gamma + b} \right) + \frac{e^{-\gamma t^*} s}{\gamma}
\]

Suppose that \( f = f_{\alpha} \), as in Proposition 2. By the same arguments as in that Proposition, if \( y_t \) satisfies \( p_t(y_t) = \frac{s}{gb} \) for all \( t \), then \( p_t(m_t) \searrow \frac{s}{\lambda(gb - s) + s} \) as \( t \to \infty \), and, by Lemma 3, we have \( t^* = -\frac{\ln(\lambda)}{b} \). Then

\[
V_g \left( \frac{s}{\lambda(gb - s) + s} \right) = \frac{s}{\lambda gb + (1 - \lambda)s} \left( \frac{b}{\gamma} - \frac{b}{\gamma + b} + \frac{gb (1 - \frac{\gamma + b}{\lambda})}{\gamma + b} + \frac{\frac{\lambda + b}{s}}{\gamma + b} - \frac{\lambda \gamma}{s} \right) + \frac{\lambda \gamma s}{\gamma}
\]

Since this is a hyperbola in \( g \), it is either increasing in \( g \) for all \( g > 0 \) or decreasing in \( g \) for all \( g > 0 \). In particular, when the congestion effect is maximal, that is, when \( g \to \infty \), we have

\[
\lim_{g \to \infty} \gamma V_g \left( \frac{s}{\lambda(gb - s) + s} \right) = \frac{b}{\gamma + b} + \frac{s}{\gamma} - \frac{b}{\gamma + b} + \frac{s}{\gamma}
\]

On the other hand, when the economies of scale are maximal, that is, as \( g \to \frac{s}{b} \), we have

\[
\lim_{g \to \frac{s}{b}} \gamma V_g \left( \frac{s}{\lambda(gb - s) + s} \right) = \gamma \left( \frac{b}{\gamma} - \frac{b}{\gamma + b} + \frac{s (1 - \frac{\gamma + b}{\lambda})}{\gamma + b} + \frac{\lambda \gamma s}{\gamma + b} \right) = \frac{b}{\gamma + b} + \frac{s}{\gamma + b}
\]

Thus \( \frac{b}{\gamma + b} + \frac{s \gamma}{\gamma + b} > \frac{b}{\gamma + b} + \frac{s}{\gamma + b} \) is equivalent to \( \lim_{g \to \infty} V_g \left( \frac{s}{\lambda(gb - s) + s} \right) > \frac{b}{\gamma + b} + \frac{s}{\gamma + b} \)

---

\[40]\text{Note that this is the same } t^\ast \text{ as in the baseline model.}\\
\[41]\text{If } g < \frac{s}{b}, \text{ we enter a degenerate case in which the organization becomes empty immediately.}
lim_{g \to \bar{g}} V_g \left( \frac{s}{\lambda \left( g b - s \right) + s} \right). Because V_g \left( \frac{s}{\lambda \left( g b - s \right) + s} \right) is either increasing in g for all g > 0 or decreasing in g for all g > 0, this condition implies that V_g \left( \frac{s}{\lambda \left( g b - s \right) + s} \right) is increasing in g for all g > 0. The argument in the case when the inequality is reversed is similar.

Finally, note that if V_g \left( \frac{s}{\lambda \left( g b - s \right) + s} \right) is increasing in g, then we can guarantee that, with a congestion effect,

\[ V_{(g_t)_{t \geq 1}} (p_t(m_t)) > V_{g_t} (p_t(m_t)) > V_{g_t} \left( \frac{s}{\lambda (g_t b - s) + s} \right) > V \left( \frac{s}{\lambda (b - s) + s} \right) \]

Here the first inequality follows because \( g_t \mapsto V_{g_t, \ldots, g_t, \ldots} (x) \) is increasing, under congestion effect \( \mu \mapsto g(\mu) \) is decreasing and under perpetual experimentation \( t \mapsto \mu_t \) is decreasing, so \( t \mapsto g_t = g(\mu_t) \) is increasing. The second inequality follows because \( x \mapsto V_{g_t} (x) \) is strictly increasing and \( p_t(m_t) \searrow \frac{s}{\lambda (g_b - s) + s} \) as \( t \to \infty \). The last inequality follows because \( g \mapsto V_g \left( \frac{s}{\lambda (g b - s) + s} \right) \) is increasing, under congestion effect \( \mu \mapsto g(\mu) \) is decreasing and in the baseline model we have \( g(\mu) = g(1) = 1 \) for all \( \mu \).

Thus the condition to obtain experimentation forever is slacker under congestion effect than in the baseline model at every \( t \), not just in the limit. By the same argument, the condition for experimentation forever is tighter for all \( t \) under economies of scale.\(^{42}\)

**Proof of Proposition 8.**

We have

\[ \bar{V}_t(x) = x \int_t^{\infty} e^{-\gamma (\tau - t)} \left[ P_\tau b + (1 - P_\tau) (sF(y_\tau) + b(1 - F(y_\tau))) \right] d\tau 
+ (1 - x) \int_t^{\infty} e^{-\gamma (\tau - t)} sF(y_\tau) d\tau, \]

where \( P_\tau = 1 - e^{-b(\tau - t)} \) is the probability that there has been a success by time \( \tau \), conditional on the state being good and there being no success up to time \( t \), and \( F(y_\tau) \) is the fraction of the population that are outsiders at time \( \tau \), conditional on no

\(^{42}\)If \( g \mapsto V_g \left( \frac{s}{\lambda (g b - s) + s} \right) \) is decreasing, it is more difficult to make general statements about what happens away from the limit because in this case the effect of moving away from the limit goes against the result: for instance, under congestion effect the condition becomes tighter in the limit but increasing \( g_t \) slackens the condition.
successes. We can rewrite this equation as

\[ \tilde{V}_t(x) = x \int_t^\infty e^{-\gamma(t-\tau)} [P_\tau b + (1 - P_\tau)(b - (b - s)F(y_\tau))] \, d\tau + (1 - x) \int_t^\infty e^{-\gamma(t-\tau)} sF(y_\tau) \, d\tau \]

We have

\[ \tilde{V}_t(x) = x \int_t^\infty e^{-\gamma(t-\tau)} [b - (1 - P_\tau)(b - s)F(y_\tau)] \, d\tau + (1 - x) \int_t^\infty e^{-\gamma(t-\tau)} sF(y_\tau) \, d\tau \]

\[ = x \int_t^\infty e^{-\gamma(t-\tau)} [b - e^{-b(t-\tau)}(b - s)F(y_\tau)] \, d\tau + (1 - x) \int_t^\infty e^{-\gamma(t-\tau)} sF(y_\tau) \, d\tau \]

\[ = x \frac{b}{\gamma} - x \int_t^\infty e^{-(\gamma+b)(t-\tau)}(b - s)F(y_\tau) \, d\tau + (1 - x) \int_t^\infty e^{-\gamma(t-\tau)} sF(y_\tau) \, d\tau \]

(10)

Let \( F_\tau = F(y_\tau) \) and note that \( \tau \mapsto F_\tau \) is weakly increasing.

The upper bound for \( \tilde{V}_t(x) \) is now obtained as follows. Note that, given \( \tau \geq t \), the derivative of (10) with respect to \( F_\tau \) is proportional to \(-xe^{-b(t-\tau)}(b - s) + (1 - x)s\).

It follows that if the agent could choose \( F_\tau \) everywhere at will to maximize her payoff, she would choose \( F_\tau = 1 \) for \( \tau \geq t(x) \) and \( F_\tau = 0 \) for \( \tau < t(x) \), where \( t(x) \) is defined by the condition \( xe^{-b(t(x)-t)}(b - s) = (1 - x)s \) (obtained by setting the derivative equal to 0). The result of this choice is \( V(x) \), her utility in the private values case, in which she only cares about her own entry and exit decisions and gets to choose them optimally. Because in the common values case the entry and exit decisions of other agents are not optimal from \( x \)'s point of view, \( \tilde{V}_t(x) \) must be weakly lower than \( V(x) \).

As for the lower bound, assume for the sake of argument that \( F_\tau \) is constant for all \( y_\tau \) and equal to \( \bar{F} \in [0, 1] \). Then the expression in (10) is

\[ x \frac{b}{\gamma} - x \frac{(b - s)\bar{F}}{\gamma + b} + (1 - x) \frac{s\bar{F}}{\gamma} \]

which is linear in \( \bar{F} \) and is minimized either when \( \bar{F} = 0 \) or when \( \bar{F} = 1 \). In the first case, the expression equals \( x\frac{b}{\gamma} \). In the second case, it equals \( x\frac{b}{\gamma} - x\frac{(b - s)}{\gamma + b} + (1 - x)\frac{s}{\gamma} \).

To finish the proof, we argue that whenever \( F_\tau \) is weakly increasing in \( \tau \), the expression in 10 is higher than the expression that is obtained when \( F_\tau \) is replaced
by a suitably chosen constant $F$. Hence the lower bound obtained for constant $F_\tau$
applies in all cases.

The argument is as follows. Take $F = F_t(x)$. Then for $\tau > t(x)$, $F_\tau$ is weakly greater than $F$ and $\tilde{V}_t(x)$ is increasing in the value of $F$ at $\tau$. Conversely, for $\tau < t(x)$, $F_\tau$ is weakly lower than $F$ and $\tilde{V}_t(x)$ is decreasing in the value of $F$ at $\tau$. Hence the agent’s utility is weakly higher under $F_\tau$ than under a constant $F_t(x)$. ■

C Value Functions under Imperfectly Informative News

In this section, we describe the result we obtained for the value function of an agent in a model of imperfectly informative news. We show in the Appendix that the value function $U_{x,G}(l)$ of an agent with prior $x$ given that the risky technology is good, the state variable is $l = \ln(L(k,t))$ must satisfy a certain delay differential equation. Lemma 20 shows that the value function must satisfy the boundary condition $\lim_{l \to -\infty} U_{x,G}(l) = \frac{b}{\gamma}$. Lemmas 21 and 23 show that the delay differential equation admits the following solution:

$$U_{x,G}(l) = U_n(l)$$

for $l \in \left[\ln \frac{x(b-s)}{(1-x)(s-c)} - (n-1) \ln \frac{c}{b}, \ln \frac{x(b-s)}{(1-x)(s-c)} - n \ln \frac{c}{b}\right]$, $n \geq 1$, and $U_{x,G}(l) = U_0(l) = De^{\omega_0 l} + b$ for $l \leq \ln \frac{x(b-s)}{(1-x)(s-c)}$, with $U_n(l)$ satisfying

$$U_n(l) = P_n(l)e^{\omega_1 l} + De^{\omega_0 l} + C_n$$

where $P_n$ is a polynomial of degree $n - 1$, $\omega_0$ satisfies $(b-c)\omega_0 = \gamma + b - be^{\omega_0 a}$ for $a = -\ln \frac{c}{b}$, $\omega_1 = \frac{\gamma + b}{b - c}$ and $D$ is a constant. Moreover, $C_n = b - (b - s)\left(1 - \left(\frac{b}{\gamma + b}\right)^n\right)$ for $n \geq 1$ and $(P_n)_n$ satisfies $P_n(l) = -\frac{b}{(b-c)e^{\omega_1 l}}P_{n-1}(l-a)$ for all $n \geq 1$.

The reason we have not been able to show that the solution we obtained for the delay differential equation is the value function of the agent is that, due to the lack of the closed form solution for the polynomials $P_n$, we have not been able to verify that there exists a constant $D$ such that the boundary condition is satisfied.

We use $V_x(L)$ to denote the value function of an agent with prior $x$ for a given value of $L$. We let $V_{x,\theta}(L)$ denote the value function of an agent with prior $x$ given
that the state variable is $L$ and the risky policy is of type $\theta$. Then

$$V_x(L) = xV_{x,G}(L) + (1 - x)V_{x,B}(L)$$

Observe that a (pure strategy) equilibrium can be characterized by a stopping set $\mathcal{L} \subseteq (0, \infty)$ such that, whenever $L \in \mathcal{L}$, the pivotal agent $m(L)$ switches to the safe policy, and experimentation continues for values of $L$ outside of $\mathcal{L}$.

**Lemma 19.** $V_{x,G}(L)$ and $V_{x,B}(L)$ satisfy the following equations for $L \in \mathbb{R}_{\geq 0} \setminus \mathcal{L}$:

$$L(b - c) \frac{\partial V_{x,G}(L)}{\partial L} = \gamma \left( \mathbb{1}_{L \leq \frac{x(b - s)}{(1 - x)(s - c)}} (s - b) - s \right) + (\gamma + b)V_{x,G}(L) - bV_{x,G} \left( \frac{L}{b} \right)$$

$$L(b - c) \frac{\partial V_{x,B}(L)}{\partial L} = \gamma \left( \mathbb{1}_{L \leq \frac{x(b - s)}{(1 - x)(s - c)}} (s - c) - s \right) + (\gamma + c)V_{x,B}(L) - cV_{x,B} \left( \frac{L}{b} \right)$$

$$V_{x,\theta}(L) = \frac{r}{\gamma}$$ for $L \in \mathcal{L}$ and $\theta \in \{B, G\}$. Moreover, the boundary conditions $V_{x,G}(0) = \frac{b}{\gamma}$ and $V_{x,B}(0) = \frac{c}{\gamma}$ are satisfied.

**Proof of lemma 19.**

Because, by lemma 16, an agent with belief $x$ at time $t$ is in the organization at time $t$ if and only if $L \leq \frac{x(b - s)}{(1 - x)(s - c)}$, and, provided that the risky technology is good, an agent’s flow payoff during the time period of length $\epsilon$ is

$$\mathbb{1}_{L \leq \frac{x(b - s)}{(1 - x)(s - c)}} b + \mathbb{1}_{L \geq \frac{x(b - s)}{(1 - x)(s - c)}} s = \mathbb{1}_{L \leq \frac{x(b - s)}{(1 - x)(s - c)}} b + s - \mathbb{1}_{L \leq \frac{x(b - s)}{(1 - x)(s - c)}} s = \mathbb{1}_{L \leq \frac{x(b - s)}{(1 - x)(s - c)}} (b - s) + s$$

Similarly, provided that the risky technology is bad, an agent’s flow payoff during the time period of length $\epsilon$ is $\mathbb{1}_{L \leq \frac{x(b - s)}{(1 - x)(s - c)}} (b - s) + s$.

Provided that the risky technology is good, with probability approximately equal to $e^{-bt}$, a success arrives within the time period of length $\epsilon$, which changes the state from $L$ to $L^c e^{(b - c)\epsilon}$. With probability approximately equal to $e^{-bt}$, a success does not arrive within this time period, which changes the state from $L$ to $Le^{(b - c)\epsilon}$.
Then we have
\[ V_{x,G}(L) \approx (1 - e^{-\gamma \epsilon}) \left( \mathbb{1}_{L \leq \frac{x(b-s)}{(1-x)(s-c)}} (b-s) + s \right) + \\
+ e^{-\gamma \epsilon} \left[ e^{-be} V_{x,G} \left( Le^{(b-c)\epsilon} \right) + (1 - e^{-be}) V_{x,G} \left( L \frac{c}{b} e^{(b-c)\epsilon} \right) \right] \]

Subtracting \( V_{x,G} \left( Le^{(b-c)\epsilon} \right) \) from both sides, we obtain
\[ V_{x,G}(L) - V_{x,G} \left( Le^{(b-c)\epsilon} \right) \approx (1 - e^{-\gamma \epsilon}) \left( \mathbb{1}_{L \leq \frac{x(b-s)}{(1-x)(s-c)}} (b-s) + s \right) + \\
+ \left( e^{-\gamma \epsilon} e^{(\gamma + b)\epsilon} - 1 \right) V_{x,G} \left( Le^{(b-c)\epsilon} \right) \\
+ e^{-\gamma \epsilon} \left( 1 - e^{-be} \right) V_{x,G} \left( L \frac{c}{b} e^{(b-c)\epsilon} \right) \]

Dividing both sides by \( \epsilon \) and taking the limit as \( \epsilon \to 0 \), we find that this simplifies to the desired equation for \( V_{x,G}(L) \). The proof for \( V_{x,B}(L) \) is similar.

We have \( V_{x,\theta}(L) = \frac{x}{\gamma} \) for \( L \in \mathcal{L} \) and \( \theta \in \{B, G\} \) because, whenever \( L \in \mathcal{L} \), the organization switches to the safe technology and uses it forever, which yields a payoff of \( \int_0^\infty e^{-\gamma \tau} r d\tau = \frac{x}{\gamma} \).

The boundary conditions \( V_{x,G}(0) = \frac{b}{\gamma} \) and \( V_{x,B}(0) = \frac{c}{\gamma} \) are satisfied because if \( L = 0 \), then all agents put probability one on the event that the technology is good. This results in the organization experimenting forever, which yields a payoff of \( \int_0^\infty e^{-\gamma \tau} bd\tau = \frac{b}{\gamma} \) if the technology is good and a payoff of \( \int_0^\infty e^{-\gamma \tau} cd\tau = \frac{c}{\gamma} \) if the technology is bad. \( \blacksquare \)

We perform a convenient change of variables, letting \( l = \ln L \) so that \( L = e^l \). Given the change of variables, we let \( \mathcal{L} = \{l = \ln(L) : L \in \mathcal{L}\} \) denote the set of the values of \( l \) for which the organization stops experimentation. For convenience, we rewrite the equations in lemma 19 with \( l = \ln L \) as our state variable. We let \( U_{x,G}(l) \) and \( U_{x,B}(l) \) denote the resulting value functions.

**Lemma 20.** \( U_{x,G}(l) \) and \( U_{x,B}(l) \) satisfy the following equations for \( l \in \mathbb{R} \setminus \overline{\mathcal{L}} \):

\[ (b-c) \frac{\partial U_{x,G}(l)}{\partial l} = \gamma \left( \mathbb{1}_{l \leq \ln \frac{x(b-s)}{(1-x)(s-c)}} (s-b) - s \right) + (\gamma + b) U_{x,G}(l) - b U_{x,G}(l) \]
\[(b - c) \frac{\partial U_{x,B}(l)}{\partial l} = \gamma \left( \mathbf{1}_{l \leq \ln \frac{x(b-s)}{(1-x)(s-c)}} (s - c) - s \right) + (\gamma + b)U_{x,B}(l) - cU_{x,B}(l)\]

\[U_{x,\theta}(l) = \frac{x}{\gamma} \text{ for } l \in \overline{L} \text{ and } \theta \in \{B,G\}. \text{ Moreover, the boundary conditions } \lim_{l \to -\infty} U_{x,G}(l) = \frac{b}{\gamma} \text{ and } \lim_{l \to -\infty} U_{x,B}(l) = \frac{x}{\gamma} \text{ are satisfied.}\]

**Proof of lemma 20.**

Note that \(\ln \left( L^\frac{x}{\gamma} \right) = \ln \left( e^{\ln \frac{x}{\gamma}} \right) = \ln \left( e^{l + \ln \frac{x}{\gamma}} \right) = l + \ln \frac{x}{\gamma} \) and that

\[
\begin{align*}
\frac{\partial U_{x,G}(l)}{\partial l} &= \frac{\partial V_{x,G}(L)}{\partial L} \frac{\partial L}{\partial l} = \frac{\partial V_{x,G}(L)}{\partial L} \frac{\partial \left( e^l \right)}{\partial l} = \\
\frac{\partial V_{x,G}(L)}{\partial L} e^l &= \frac{1}{L} \frac{\partial U_{x,G}(L)}{\partial l}.
\end{align*}
\]

which implies that \(\frac{\partial V_{x,G}(L)}{\partial L} = \frac{1}{L} \frac{\partial U_{x,G}(L)}{\partial l}\). Note also that \(V_{x,G}(L) = U_{x,G}(l)\).

Substituting the formulas for \(\frac{\partial V_{x,G}(L)}{\partial L}\) and \(V_{x,G}(L)\) into the equations from lemma 19, we obtain the desired equation for \(U_{x,G}(l)\) in the statement of lemma 20. The proof for \(U_{x,B}(l)\) is similar.

We introduce several definitions that we find convenient to use in the proofs below. We let \(d = \ln \frac{x(b-s)}{(1-x)(s-c)}\) denote the threshold value of \(l\) such that an agent is in the organization if and only if \(l\) is below this threshold. We let \(a = -\ln \frac{x}{\gamma}\) denote the amount by which \(l\) decreases after the technology experiences a success. We let \(U_0(l) = U_{x,G}(l)\) denote the value function of an agent with prior \(x\) given that \(l \leq d\) and that the technology is good. Finally, we let \(U_n(l) = U_{x,G}(l)\) denote the value function an agent with prior \(x\) given that \(l \in (d + (n - 1)a, d + na]\) for \(n \geq 1\) and given that the technology is good.

**Lemma 21.**

\[U_0(l) = De^{\omega_0l} + C_0\]

for \(\omega_0\) satisfying \((b - c)\omega_0 = \gamma + b - be^{-\omega_0a}, C_0 = b\) and some constant \(D\).

**Proof of lemma 21.**

Note that if \(U_0(l) = De^{\omega_0l} + C_0\), then \(U'_0(l) = D\omega_0e^{\omega_0l}\).
Suppose that \( l \leq d \). Then the equation from lemma 20 can be written as

\[
(b - c)U'_0(l) = (\gamma + b)U_0(l) - bU_0(l - a) - \gamma b
\]

Substituting in the conjectured formula for \( U_0(l) \), we obtain

\[
(b - c)\omega_0 De^{\omega_0l} = (\gamma + b) \left( De^{\omega_0l} + C_0 \right) - b \left( De^{\omega_0(l-a)} + C_0 \right) - \gamma b
\]

In order for the constant terms to cancel out, we need \( 0 = (\gamma + b)C_0 - bC_0 - \gamma b \), which is equivalent to \( C_0 = b \).

Then the equation simplifies to

\[
(b - c)\omega_0 De^{\omega_0l} = (\gamma + b) De^{\omega_0l} - b De^{\omega_0(l-a)}
\]

Canceling \( De^{\omega_0l} \) from both sides, we obtain

\[
(b - c)\omega_0 = \gamma + b - be^{\omega_0a}
\]

which pins down \( \omega_0 \).

Lemma 22. If \( U_0(l) = De^{\omega_0l} + b \) for \( \omega_0 \) satisfying \( (b - c)\omega_0 = \gamma + b - be^{\omega_0a} \) and some constant \( D \), then

\[
U_1(l) = De^{\omega_1l} + a_0 e^{\omega_1l} + C_1
\]

for \( \omega_1 = \frac{\gamma + b}{b-c}, C_1 = \frac{b^2 + \gamma s}{s + b} \) and some constant \( a_0 \).

Proof of lemma 22.

Note that if \( U_1(l) = D_1 e^{\omega_1l} + a_0 e^{\omega_1l} + C_1 \) for some constant \( D_1 \), then \( U'_1(l) = \omega_0 D_1 e^{\omega_0l} + \omega_1 a_0 e^{\omega_1l} \).

Suppose that \( l \in (d, d + a] \), so that \( l - a \in (d - a, d] \). Then the equation from lemma 20 can be written as

\[
(b - c)U'_1(l) = (\gamma + b)U_1(l) - bU_0(l - a) - \gamma s
\]
Substituting in the formulas for $U_1(l)$ and $U'_1(l)$, this is equivalent to

$$(b-c)\left(\omega_0D_1e^{\omega_0l} + \omega_1a_0e^{\omega_1l}\right) = (\gamma+b)\left(D_1e^{\omega_0l} + a_0e^{\omega_1l} + C_1\right) - b\left(De^{\omega_0(l-a)} + C_0\right) - \gamma s$$

In order for the constant terms to cancel out, we need $0 = (\gamma+b)C_1 - bC_0 - \gamma s$.

That is, we need $C_1 = \frac{b^2 + \gamma b}{\gamma + b}$.

Then the equation simplifies to

$$(b-c)\left(\omega_0D_1e^{\omega_0l} + \omega_1a_0e^{\omega_1l}\right) = (\gamma+b)\left(D_1e^{\omega_0l} + a_0e^{\omega_1l}\right) - bD_0e^{\omega_0(l-a)}$$

To match the coefficients, we need that $(b-c)\omega_0D_1e^{\omega_0l} = (\gamma+b)D_1e^{\omega_0l} - bDe^{\omega_0(l-a)}$. This equation holds for all $l$ if $D_1 = D$, and there can only be one value of $D_1$ that works for all $l$, so $D_1 = D$.

Then the equation simplifies to

$$(b-c)\omega_1a_0e^{\omega_1l} = (\gamma+b)a_0e^{\omega_1l}$$

which implies that $\omega_1 = \frac{\gamma + b}{b-c}$.

\[\square\]

**Lemma 23.**

$$U_n(l) = P_n(l)e^{\omega_1l} + D_n e^{\omega_0l} + C_n$$

where $P_n$ is a polynomial of degree $n-1$, $\omega_0$ satisfies $(b-c)\omega_0 = \gamma + b - be^{\omega_0a}$ for $a = -\ln \frac{c}{b}$, $\omega_1 = \frac{\gamma + b}{b-c}$ and $D_n = D$ for some constant $D$ for all $n \geq 1$.

Moreover, $C_n = b - (b-s) \left(1 - \left(\frac{b}{\gamma + b}\right)^n\right)$ for $n \geq 1$ and $(P_n)_n$ satisfies

$$P'_n(l) = -\frac{b}{(b-c)e^{\omega_1a}}P_{n-1}(l-a)$$

for all $n \geq 1$.

**Proof of lemma 23.**

We will prove the lemma by induction.

Lemma 22 shows that the statement is true for $n = 1$. Suppose as an inductive hypothesis that the statement is true for $n = k$, and consider $U_{k+1}$.
We have
\[
U'_{k+1}(l) = P'_{k+1}(l)e^{\omega_1 l} + P_{k+1}(l)\omega_1 e^{\omega_1 l} + D_{k+1}\omega_0 e^{\omega_0 l}
\]
and we want to show that
\[
(b - c)U'_{k+1}(l) = (\gamma + b)U_{k+1}(l) - bU_{k}(l-a) - \gamma s
\]
Substituting in the formulas for \(U'_{k+1}(l), U_{k+1}(l)\) and \(U_{k}(l-a)\), we want to show that
\[
(b - c)\left( P'_{k+1}(l)e^{\omega_1 l} + P_{k+1}(l)\omega_1 e^{\omega_1 l} + D_{k+1}\omega_0 e^{\omega_0 l} \right) =
(\gamma + b)(P_{k+1}(l)e^{\omega_1 l} + D_{k+1}e^{\omega_0 l} + C_{k+1}) - b(P_{k}(l-a)e^{\omega_1(l-a)} + D_{k}e^{\omega_0(l-a)} + C_{k}) - \gamma s
\]
For the constants to cancel out, it must be that \(0 = (\gamma + b)C_{k+1} - bC_{k} - \gamma s\). This pins down \(C_{k+1}\) for all \(n \geq 1\) given that \(C_{0} = b\), and we can check manually that \(C_{n} = b - (b - s)\left(1 - \left(\frac{b}{\gamma+b}\right)^n\right)\) works.

The equation then simplifies to
\[
(b - c)\left( P'_{k+1}(l)e^{\omega_1 l} + P_{k+1}(l)\omega_1 e^{\omega_1 l} + D_{k+1}\omega_0 e^{\omega_0 l} \right) =
(\gamma + b)(P_{k+1}(l)e^{\omega_1 l} + D_{k+1}e^{\omega_0 l}) - b(P_{k}(l-a)e^{\omega_1(l-a)} + D_{k}e^{\omega_0(l-a)})
\]
As in lemma 22, for the terms multiplied by \(e^{\omega_0 l}\) to cancel out, we need that \(D_{k+1} = D\). The equation then simplifies to
\[
(b - c)\left( P'_{k+1}(l)e^{\omega_1 l} + P_{k+1}(l)\omega_1 e^{\omega_1 l} \right) = (\gamma + b)P_{k+1}(l)\omega_1 e^{\omega_1 l} - bP_{k}(l-a)e^{\omega_1(l-a)}
\]
Since \(\omega_1 = \frac{\gamma+b}{b-c}\), we have that \((b - c)P_{k+1}(l)\omega_1 e^{\omega_1 l} = (\gamma + b)P_{k+1}(l)e^{\omega_1 l}\). Then the equation simplifies to
\[
(b - c)P'_{k+1}(l)e^{\omega_1 l} = -bP_{k}(l-a)e^{\omega_1(l-a)}
\]