Endogenous Experimentation in Organizations∗

Germán Gieczewski†      Svetlana Kosterina‡

April 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 in the wider population possess resources of use to the organization and can enter and leave the organization freely. 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 more radical median voter who supports continued experimentation. Conversely, successes induce less enthusiastic agents to join. When successes are imperfectly informative, these selection forces can sustain equilibria where, on average, the organization experiments more when the policy is bad than when it is good. The model is applied to decision-making in cooperatives, civil rights organizations and for-profit firms.

Keywords: experimentation, dynamics, median voter, endogenous population

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

∗We would like to thank Daron Acemoglu, Alessandro Bonatti, Glenn Ellison, Robert Gibbons, Matias Iaryczower, Navin Kartik, Wolfgang Pesendorfer, 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 for helpful comments. All remaining errors are our own.
†Department of Politics, Princeton University.
‡Department of Politics, Princeton University.
are diverse: the members of an organization oftentimes disagree about the merits of its policies. In addition, their membership is fluid: some members leave, disillusioned with the policies pursued, while others join, lured by the promise of greater benefits afforded by the organization’s management of its considerable resources. As the membership of an organization changes, so do the policies they pursue. 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. 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. The members of the non-profit organization vote on which policy to pursue.

In the model, an organization is choosing 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 there exist parameters
under which there is overexperimentation relative to the social optimum in the unique equilibrium. Overexperimentation takes a particularly stark form: the organization experiments forever regardless of the outcome. 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. Success may make the 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.

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 underexperimentation more likely, while the second force pushes the organization to overexperiment. The first major contribution of this paper is showing that the second force can dominate.

More precisely, we prove two results which determine in what cases experimenting forever is the (unique) equilibrium outcome. The first result provides simple and sufficient conditions on the parameters that guarantee perpetual experimentation for several specific densities $f$ of the agents’ prior beliefs, namely, when $f$ is uniform; when it follows a power law; and when it is arbitrary. The second result states that, if there is perpetual experimentation for a density $f$, the same is true for any density that MLRP-dominates $f$. In other words, greater optimism (in the sense of the monotone likelihood ratio property) increases the likelihood of overexperimentation.

Finally, we characterize the set of equilibria when the parameters are such that there does not exist an equilibrium with perpetual experimentation. From the point of view of the initial median voter, both overexperimentation and underexperimentation are possible in these equilibria.
The rest of the paper extends the model to more general settings. The first two extensions concern 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. We show that 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 gradually expands 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 overexperimentation from the point of view of any pivotal agent—because, in this special case, even perpetual experimentation does not constitute overexperimentation—but underexperimentation 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, with modified parameter conditions. In addition, 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 general non-median voting rules and settings in which the members’ flow payoffs depend on the size of the organization. We show that our result of overexperimentation in equilibrium also obtains in these settings. Voting rules requiring greater supermajorities to make policy changes make overexperimentation more likely, while payoffs decreasing in the size of the organization make it more difficult to sustain an equilibrium with overexperimentation.

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 in which success is only imperfectly informative about the quality of the technology. 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 non-profit organizations, cooperatives, civil rights organizations, and publicly traded firms.

One prominent application of our model, the non-profit organizations, has been discussed in the introduction. Our next application is a cooperative. Here agents are individual producers who own factors of production. In 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. In the second case, 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, the cooperative can limit itself to selling mainstream products or it can instead develop a line of premium cheeses that may or may not become popular. 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.

Another application of our model is a civil rights organization as a vehicle for political activism. A citizen desiring to change the government’s policy on an LGBT rights issue can act independently by, for instance, writing to her elected representatives, or she can join a civil rights organization 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 the risky strategies while others support safe ones, the members of the organization can influence the strategy the organization chooses by voting in the organization’s leadership elections.

In a publicly traded firm, agents are the individuals who invest in the firm.\(^1\)

\(^1\)Because, by making public offerings or buying back shares, firms typically also control how many
Having bought the firm’s shares, they gain voting rights which afford them a measure of control over the firm’s decisions. The shareholders influence the policy of the firm 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.

3 Related Literature


Keller, Rady and Cripps (2005) develop a model with multiple agents each controlling a two-armed bandit. In contrast, the present paper considers multiple agents with heterogeneous beliefs experimenting with the same bandit and being free to enter and exit an organization. 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 overexperimentation in equilibrium.

In Strulovici (2010) a community of agents decides by voting at each point in time whether to continue experimenting with a risky technology or to switch to a safe technology. Agents’ payoffs from the risky technology are heterogeneous: under complete information some agents would prefer it to the safe technology, while others would not. The agents learn about their payoff from the risky technology by observing their payoffs while the experimentation continues. Differently from Strulovici, the 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 overexperimentation in our model should still apply there.
present paper considers multiple agents who have the same payoffs but heterogeneous beliefs and are free to enter and exit an organization.

Strulovici finds that there is too little experimentation in equilibrium because agents fear being trapped into using the risky technology that turns out to be bad for them. The same incentive to underexperiment as in Strulovici is present in our model under the natural assumption that the return to investing outside the organization is strictly lower than the return to investing with the organization provided the organization uses the safe technology. This is because then, even though the agent who has become pessimistic about the risky technology can always exit the organization and invest outside, she would strictly prefer to remain in the organization and have the organization pursue the safe policy. Consider an agent who would prefer to experiment today but not tomorrow if she was always in control of the organization. Suppose that this agent is in control today and anticipates that the agent who will gain control tomorrow should the experimentation continue will experiment. Then she may choose not to experiment today if the benefits from investing in the organization which pursues the safe policy rather than outside of it are great enough.

In contrast to Strulovici’s results, in the present paper under some parameters there is a unique equilibrium exhibiting overexperimentation. The reason for the difference lies in the ability of the agents in our model to freely enter and exit the organization and the agents’ heterogeneous prior beliefs about the risky technology. Because the agents can freely exit the organization, they cannot become trapped in an organization pursuing the policies they disagree with, unlike in Strulovici, though, as explained above, the incentive to underexperiment due to the fear of the loss of control is still present in our model. Moreover, because agents differ in their beliefs about the risky technology, only the most optimistic agents stay in the organization. Since these are precisely the agents most likely to want to continue to experiment, this produces incentives for the organization to overexperiment. We show that there are parameters under which the incentives to overexperiment dominate the incentives to underexperiment in our model.

The literature on the dynamics of decision-making in clubs considers the dynamics of policy-making in a setting where 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, Bai and Lagunoff
the feature that the policy chosen by the decision-maker in control of the organization’s policy today affects the identities of the future decision makers. 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 from the aforementioned papers in considering agents whose preferences are the same but who differ in their prior beliefs about the risky policy and in studying a setting in which the organization can experiment with a policy, causing 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 $[0, 1]$ is given by 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. We use $X_t \subseteq [0, 1]$ to denote the subset of the population that belongs to the organization at time $t$. We write $\beta_t(x) = 1$ if $x \in X_t$ and $\beta_t(x) = 0$ otherwise. 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

---

2This notation rules out mixed membership strategies, but the restriction is without loss of generality.
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.

It follows that the expected utility of an agent with a prior belief $x$ is given by

$$
xE\left[\int_0^\infty e^{-\gamma t} (b\pi_t \beta_t(x) + r(1 - \pi_t)\beta_t(x) + s(1 - \beta_t(x))) dt|\theta = G\right] + \\
(1 - x)E\left[\int_0^\infty e^{-\gamma t} (r(1 - \pi_t)\beta_t(x) + s(1 - \beta_t(x))) dt|\theta = B\right]
$$

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. It then follows from the Bayes’ rule
that the posterior of an agent with prior $x$ after seeing $k$ successes during a length of
time $\tau$ spent experimenting is

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

where $L(k, \tau) = \mathbb{1}_{k=0}e^{br}$. Since $L(k, \tau)$ serves as a sufficient statistic for the inform-
ation observed so far, suppressing the dependence of $L(k, t)$ on $k$ and $t$, 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 $L$. We let \(\tilde{L}_t\) denote the stochastic process induced by experimenting without stopping starting at
time 0.

Recall that, at each time $t$, a subset of the population $X_t$ belongs to the organi-
bzation. We assume that the median voter of this set chooses whether the organization
should continue to experiment at that instant. 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

\[\text{Note that, in particular, if } k = 0 \text{ and } \tau = 0, \text{ then the posterior is } x.\]

\[\text{For the median of this set to be well-defined, we require } X_t \text{ to be Lebesgue-measurable. It can be shown that in any equilibrium } X_t \text{ is an interval.}\]
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 $\alpha_t$ 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 $\alpha_t$—are the incumbent members at time $t + \epsilon$. The median of this set of members then chooses $\alpha_{t+\epsilon}$. The small delay between joining the organization and voting on the policy rules out equilibria involving self-fulfilling prophecies. These are the 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.

The above requirements are incorporated into the following notions of strategy profile and equilibrium:

**Definition 1.** A strategy profile is given by a collection of membership functions $\beta(x, L, \pi) : [0, 1] \times \mathbb{R}_+ \times \{0, 1\} \to \{0, 1\}$, a policy function $\alpha(L, \pi) : \mathbb{R}_+ \times \{0, 1\} \to \{0, 1, [0, 1]\}$, and a (stochastic) path of play given by a policy, information and pivotal voter path $(\pi_t, L_t, m_t)_t$, which satisfies:

(a) $(L_t, \pi_t)_t$ is a progressively measurable Markov process with cadlag paths.

(b) $L_t = \tilde{L}_{n(t)}$, where $(\tilde{k}_\tau)_{\tau}$ is an underlying Poisson process with rate $b$ or $0$ if $\theta = G$ or $B$ respectively; $(\tilde{L}_\tau)_{\tau}$ is given by $\tilde{L}_\tau = L(\tilde{k}_\tau, \tau)$; an $n(t) = \int_0^t \pi_s$ is the amount of time that the organization has experimented for up to time $t$.

(c) Letting $s_t = \lim_{t \to t_0} \pi_t$ be the *incumbent policy* at time $t_0$, for every $t_0 \geq 0$, $\pi_{t_0} \in \alpha(L_{t_0}, s_t)$.

(d) $m_t = m_{\pi_t}(L_t)$, where $m_{\pi_t}(L_t)$ is the median voter of the set of members of the organization given that the information is $L_t$ and the policy is $\pi_t$.

An equilibrium $\sigma$ is a strategy profile such that:

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

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

We have $\beta(x, L, \pi) = 1$ if agent $x$ chooses to be a member of the organization given information $L$ and policy $\pi$, and $\beta(x, L, \pi) = 0$ otherwise. $\alpha(L, \pi)$ is the probability that the pivotal decision-maker chooses to employ the risky policy, given information $L$ and existing policy $\pi$.

Parts (c) and (d) of the definition of strategy profile say that the policy chosen along the path of play can only change when the pivotal decision-maker given the existing policy wants to change it. 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. Finally, 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 in equilibrium depending on what has happened in the game so far. We use these observations to provide insight into the structure of the equilibria. We then state our first main result, which shows that the organization may experiment forever and provides a simple sufficient condition for this to happen.

---

$V_x(L, \pi)$ is $x$'s continuation utility starting from $t_0$ with information $L_{t_0} = L$ when the current policy is $\pi$ and the equilibrium strategies are played thereafter. $\nabla_x(L, \pi, \epsilon)$ is $x$'s continuation utility starting from $t_0$ with information $L_{t_0} = L$ when $\pi$ is played during $(t, t + \epsilon)$ irrespective of the equilibrium strategies and the equilibrium strategies are played thereafter.
(Propositions 1 and 2). Finally, in Proposition 3 we characterize the equilibria in cases when the sufficient condition for obtaining experimentation forever fails.

We start by making several useful observations about the composition of the set of members at different histories of the game. 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 members forever.

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 also have higher posteriors 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. 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 \( t^* \).\(^6\) 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, an equilibrium can be described by a set \( t_0 < t_1 < t_2 < \ldots \) of stopping times, in the following sense. 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

\(^6\)If \( t^* = +\infty \), the risky policy is used forever.
switches to the safe policy at $t_n$.\footnote{$t_0$ is the only stopping time on the equilibrium path.}

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 of its results.

To state Proposition 1, we will need the following definitions. $V(x)$ is the continuation utility of an agent with posterior belief $x$ at time $t$, provided that she expects experimentation to continue for all $s \geq t$. $m_t$ is 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 $V(p_t(m_t)) > \frac{r}{\gamma}$ for all $t$, then there is a unique equilibrium. In this equilibrium, if the risky policy is used at $t = 0$, the organization experiments forever. If $\inf_{t \geq 0} V(p_t(m_t)) < \frac{r}{\gamma}$, there is no equilibrium in which the organization experiments forever.

![Figure 1: Median voter, indifferent voter, and marginal member on the equilibrium path](image)

The intuition behind Proposition 1 is illustrated in Figure 1. As the organization experiments unsuccessfully on the equilibrium path, all agents become more pessimistic, i.e., $p_t(x)$ is decreasing in $t$ for fixed $x$. In particular, if we define $x_t$ as the agent indifferent about continuing experimentation at time $t$ (i.e., $V(p_t(x_t)) = \frac{r}{\gamma}$), then $x_t$ 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 (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 \((y_t \text{ increases})\), and so does the median member \((m_t \text{ increases})\). If \(m_t \geq x_t\) for all \(t\), 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 implies 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 \(\frac{e^{-1}}{1 + e^{-1}}\). 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 concrete 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

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

(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
\[ \lambda = \frac{1}{2^{\frac{1}{\pi^2}}} \]. Then

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

(iii) Let \( f \) be any density with support \([0, 1]\). Then

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

Proposition 2 shows that the sufficient condition for obtaining experimentation forever provided in Proposition 1 is not difficult to satisfy. It is more 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 \( s = 2 \), when \( b \geq 6 \), the condition holds regardless of \( \gamma \), when \( 10^3 < b < 6 \), it holds for low enough \( \gamma \), and when \( b \leq 10^3 \), it cannot hold.

The logic behind the bounds provided by Proposition 2 can be explained as follows. When \( f \) is uniform or follows a power law, \( p_t(m_t) \) 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 \( V(\lim p_t(m_t)) \geq \frac{s}{\gamma} \). Since the marginal member \( y_t \) always has posterior belief \( \frac{s}{b} \), and the most optimistic member has posterior belief 1, \( \lim p_t(m_t) \) must be between these values, but its position depends on the position of \( m_t \) in the interval \([y_t, 1]\). When \( f \) is uniform, \( m_t \) is the midpoint between \( y_t \) and 1; when \( f \) decreases steeply near \( x = 1 \), \( m_t \) is closer to \( y_t \) than 1, resulting in a value of \( \lim p_t(m_t) \) closer to \( \frac{s}{b} \), and vice-versa. Even if \( f \) decreases faster than any polynomial, we can guarantee that the median member is at least as optimistic as the marginal member, so \( p_t(m_t) \geq \frac{s}{b} \).

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 \( \tau(t) \). If there is no such \( t \), then \( \tau(t) = +\infty \). \(^8\)

Proposition 3 characterizes the 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 \( t_0(\sigma) < t_1(\sigma) < t_2(\sigma) < \ldots \) such that \( t_n(\sigma) = \tau(t_{n-1}(\sigma)) \) for all \( n > 0 \) and \( t_0(\sigma) \leq \tau(0) \).

In particular, there can be at most one pure strategy equilibrium given a value of \( t_0 \).

Moreover, if \( \tau \) is 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 continuing 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 the set of equilibria all equilibria must still be of the form specified in Proposition 3 but it may be that, for some times \( t \in (0, \tau(0)) \), there does not exist an equilibrium in which experimentation continues for time \( t \) on the equilibrium path.

Lastly, 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.

\(^8\)It can be shown that \( \tau(t) \) is unique.
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 greatly simplify the analysis, allowing us to provide closed-form solutions and detailed characterizations of the equilibria. When the 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. Given an equilibrium, we let $p_t(m_t)$ denote the probability that the median member of the organization at time $t$ assigns at time $t$ to the event that the risky technology is good provided that that the organization has been experimenting from time 0 to time $t$ and no failures have been observed. We let $W_{T-t}(x)$ 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 $\tau > t$, $\frac{\partial}{\partial t} W_{\tau-t}(p_t(m_t)) \neq 0$
whenever $W_{\tau-t}(p_t(m_t)) = \frac{r}{\gamma}$.

Assumption 1 states that the parameters of the problem—$b$, $r$, $\gamma$ and $f$—are such that agents’ value functions are well-behaved: that is, for each $\tau$, the function $t \mapsto W_{\tau-t}(p_t(m_t))$ does not have its derivative equal zero at any point where it crosses the threshold $\frac{r}{\gamma}$.\footnote{Observe that this implies that $W_{\tau-t}(p_t(m_t))$ can only equal $\frac{r}{\gamma}$ for a finite number of times $t$.}

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 $I_0 = [t_0, t_1]$, $I_1 = [t_2, t_3]$, ..., $I_n$ 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 $V \left( \frac{2s}{b+s} \right) \geq \frac{r}{\gamma}$, 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. In a model of bad news, so long as no failures are observed, all agents become more optimistic over time 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 either continues forever 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 an analogous 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.

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 $p_t(m_t) \geq \frac{2s}{b+s}$ 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 single-agent version of the 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 during \([0, t]\) and achieve \( k \) successes is\(^1\)

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

As before, we suppress the dependence of \( L(k, t) \) on \( k \) and \( t \) and use \( L \) to denote our sufficient statistic.

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 \( V(x) = V_x(0) \) 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 \( V\left( \frac{2(s-c)}{(b-c)+s} \right) \geq \frac{r}{\gamma} \), there is a unique equilibrium in which the organization experiments forever. If \( V\left( \frac{2(s-c)}{(b-c)+s} \right) < \frac{r}{\gamma} \), there is no equilibrium in which the organization experiments forever.

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

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 \( \mathcal{L} \).

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.

\(^1\)The agent’s posterior only depends on \( k \) and \( t \), and not on the timing of the successes.
Proposition 6. There are values of $b$, $r$, $s$, $c$ and $f$ for which there exists an equilibrium such that if $L = L^*$, then there exists $\epsilon \in (0, 1]$ such that the organization stops experimenting with probability $\epsilon$, and, if $L \neq L^*$, then the organization continues experimenting with probability one.

The intuition behind the equilibrium is as follows. Suppose that the density of prior beliefs $f$ is such that $m(L)$ increases in $L$ quickly to the right of a certain value $L^*$, but slowly to its left—for instance, because $f(x)$ is high for $x < m(L^*)$ and low for $x > m(L^*)$—and that, as a result, $L \mapsto p(L, m(L))$ is decreasing for $L < L^*$ but increasing for $L > L^*$. It may then be that $p(L, m(L))$ 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 the pivotal decision-makers.

The striking feature of this equilibrium is that stopping only happens for an intermediate value of $L$. In particular, if $L^* < L(0, 0) = 0$, 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 state of the world is bad than when it is good.

Observe that in the equilibrium we just described persistent failure makes organizations more radical, as reflected in their willingness to experiment forever, while success may make organizations more conservative and more prone to stop experimentation.

Interestingly, the result in Proposition 6 can be obtained only for certain densities of prior beliefs. In particular, it can be shown that if the density of prior beliefs is increasing, then the posterior belief of the median $p(L, m(L))$ cannot be decreasing at any value of $L$. Thus the equilibrium in which the experimentation is more likely to continue forever if the risky policy is bad cannot obtain if optimistic beliefs about the risky technology are prevalent in the abovementioned sense.
To conclude 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 19 shows that the value function must satisfy the boundary condition $\lim_{l \to -\infty} U_{x,G}(l) = \frac{b}{\gamma}$. Lemmas 20 and 22 show that the delay differential equation admits the following solution:

$$U_{x,G}(l) = U_n(l) \quad \text{for} \quad l \in \left(\ln \frac{x(b-s)}{(1-x)(s-c)} - (n-1) \ln \frac{c}{b}, \ln \frac{x(b-s)}{(1-x)(s-c)} - n \ln \frac{c}{b}\right], \quad n \geq 1,$$

$$U_{x,G}(l) = D e^{\omega_0 l} + b \quad \text{for} \quad l \leq \ln \frac{x(b-s)}{(1-x)(s-c)}; \quad \text{with} \quad U_n(l) \text{ satisfying}$$

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

where $P_n$ is a polynomial of degree $n - 1$, $\omega_0$ satisfies $(b-c)\omega_0 = \gamma + b - b e^{\omega_0 a}$ 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 $P'_n(l) = -\frac{b}{(b-s)e^{-\omega_1}} P_{n-1}(l-a)$ 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.

7 Other Extensions

7.1 General Voting Rules

Our analysis can be extended to other voting rules under which the agent whose preferences are equal to the $q$-th percentile of the preferences of the members makes policy decisions. It can be shown that, in this case, if the organization has experimented continuously for time $t$ and has not experienced a success, then, provided that $f$ is uniform, the posterior belief of the decision-maker $q_t$ in the organization at time $t$ is given by $p_t(q_t) = \frac{q s + q(b-s)e^{-bt}}{qs+(1-q)b+q(b-s)e^{-at}}$. Observe that this implies that, as $t \to \infty$, the posterior of the decision-maker converges to $\frac{s}{qs+(1-q)b}$. It is straightforward to extend our arguments for show that for any voting rule $q \in (0, 1)$ there exist parameters such that in the unique equilibrium the organization experiments forever.
Observe that, because \( q \mapsto \frac{\alpha q}{qs + (1-q)b} \) is increasing, more stringent majority requirements produce more optimistic leadership of the organization. It can be shown that this makes it easier to sustain an equilibrium with over-experimentation.

### 7.2 Payoffs Depending on the Size of the Organization

In some settings the payoffs that members of an organization obtain may depend on the size of the organization. In this section we discuss how our results are affected by this dependence. We consider organizations that are subject to a congestion effect. The congestion effect captures the idea that in some settings the organization produces a fixed amount of benefits that then have to be divided among the members, and the share that each member gets decreases with the size of the organization.

We assume that if the flow payoff generated by the technology that the organization uses in period \( t \) is \( v \) and the measure of agents who are members of the organization at time \( t \) is \( \mu_t \), a member receives a flow payoff of \( \frac{v}{\alpha + (1-\alpha)\mu_t} \). Here \( \alpha \) measures the magnitude of the congestion effect: if \( \alpha = 1 \), then there is no congestion effect, while if \( \alpha = 0 \), the congestion effect is the strongest.

Lemma 23 in the Appendix shows that in an equilibrium in which the organization experiments forever in the limit as \( t \to \infty \) the posterior of the marginal member \( p_t(y_t) \) converges to \( \frac{s}{b} \), and the posterior of the median \( p_t(m_t) \) converges to \( \frac{2s}{b + sa} \). Thus when there is no congestion effect (\( \alpha = 1 \)), these are the marginal belief \( \frac{s}{b} \) and the limit posterior of the median \( \frac{2s}{b + sa} \) from the baseline model. Lemma 27 in the Appendix shows that for any \( \alpha \in \left( \frac{1}{2}, 1 \right] \), we can find parameters under which there exists an equilibrium in which the organization experiments forever.

The congestion effect makes it more difficult to sustain the equilibrium with perpetual experimentation. This is because in the equilibrium in which the organization experiments forever pessimistic agents need to leave the organization fast enough to ensure that the median is sufficiently optimistic. Because members of a smaller organization receive a higher share of payoffs, the congestion effect makes agents who anticipate that the organization will shrink reluctant to leave.
Appendix

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 an 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

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

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 \( 2(1 - m_t)^{\alpha + 1} = (1 - y_t)^{\alpha + 1} \). 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

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

In particular, if \( \alpha = 0 \), \( f \) is uniform and

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

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 $s/b$, at which time she would leave the organization. Then

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

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

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

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

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

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

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 he time it takes for an agent’s posterior belief to go from
\[ t^y(x) = -\ln\left(\frac{y}{1-y} \frac{1-x}{x}\right) \quad t(x) = -\frac{\ln\left((1-x)/(b-s)x\right)}{b} \]

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

**Proof of Lemma 3.**

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

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

**Lemma 4.** \( x \mapsto V(x) \) is continuous and strictly increasing.

**Proof of Lemma 4.** The continuity follows trivially from Lemma 2. That \( V \) is strictly increasing is a consequence of the more general Lemma 12.

**Lemma 5.** 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 \( T \). For all interior beliefs \( x \), the following is true: if \( V(x) > \frac{\xi}{\gamma} \), then \( W_T(x) > \frac{\xi}{\gamma} \) for all \( T > 0 \).

**Proof of Lemma 5.** Note that

\[ W_T(x) - V(x) = e^{-rT} P_T (W_0(p_T(x)) - V(p_T(x))) = e^{-rT} P_T \left( \frac{r}{\gamma} - V(p_T(x)) \right), \]

where \( P_T \) is the probability that there will be no success up to time \( T \), based on a prior belief \( x \) about the risky policy, and \( p_T(x) \) is the agent’s posterior belief after time \( T \) in this case.

In other words, the continuation value of an agent with belief \( x \) under experimentation up to time \( T \) is the same as under experimentation forever, except for that, if no successes are observed up to time \( T \), then in the first case the remaining continuation value is that of the safe policy, \( \frac{\xi}{\gamma} \), while in the second case it is \( V(p_T(x)) \). Now, by Lemma ??, we have \( V(x) > V(p_T(x)) \) since \( p_T(x) < x \). Assume
that $V(x) > \frac{r}{\gamma} \geq W_T(x)$. Then

$$\frac{r}{\gamma} \geq W_T(x) = V(x) + e^{-rT} P_T \left( \frac{r}{\gamma} - V(p_T(x)) \right) = e^{-rT} P_T (V(x) - V(p_T(x))) + (1 - e^{-rT} P_T) V(x) + e^{-rT} P_T \frac{r}{\gamma} > \frac{r}{\gamma},$$

a contradiction.

Lemma 6. Let $m(L)$ and $\tilde{m}(L)$ denote the median voter when the state variable is $L$ and the density is $f$ or $\tilde{f}$, respectively. Suppose that $\tilde{f}$ MLRP-dominates $f$. Then, for any $L$, $\tilde{m}(L) \geq m(L)$.

**Proof of Lemma 6.** Let $y(L)$ be the indifferent agent given information $L$ under either density (note that $y(L)$ is given by the condition $p(L, y(L)) = \frac{s}{b}$, independently of the density). By definition,

$$\int_{y(L)}^{m(L)} f(x)dx = \int_{m(L)}^{1} f(x)dx.$$

Now, suppose that $\tilde{m}(L) < m(L)$. This is equivalent to

$$\int_{y(L)}^{\tilde{m}(L)} s(x) f(x)dx = \int_{y(L)}^{m(L)} \tilde{f}(x)dx > \int_{m(L)}^{\tilde{m}(L)} \tilde{f}(x)dx = \int_{\tilde{m}(L)}^{m(L)} s(x) f(x)dx,$$

where $s(x) = \frac{\tilde{f}(x)}{f(x)}$. Since $\tilde{f}$ MLRP-dominates $f$, $s(x)$ is weakly increasing. Thus

$$\int_{y(L)}^{\tilde{m}(L)} s(\tilde{m}(L)) f(x)dx \geq \int_{y(L)}^{\tilde{m}(L)} s(x) f(x)dx > \int_{m(L)}^{1} s(x) f(x)dx \geq \int_{m(L)}^{1} s(\tilde{m}(L)) f(x)dx,$$

a contradiction.

Lemma 7. 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 $y(L) \rightarrow 1$ as $L \rightarrow \infty$. Then $\frac{1 - \tilde{m}(L)}{1 - m(L)} \rightarrow 1$ as $L \rightarrow \infty$.

**Proof of Lemma 7.**

Given a state variable $l$ and the marginal member $y(L)$ corresponding to it, let $f_0 = f(y(L))$ and $f_1 = f(1)$. By Lemma 6, 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 \( 1 - \hat{m}(L) = \frac{f_{0L}(1 - y(L))}{f_{0L} + f_1} \) 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 \( \frac{1 - \hat{m}(L)}{1 - m(L)} \to 1 \).

**Lemma 8.** 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 \( t \to \infty \). Assume that \( \frac{1 - x_t}{1 - \tilde{x}_t} \to 1 \). Then \( \frac{p_t(\tilde{x}_t)}{p_t(x_t)} \to 1 \).

**Proof of Lemma 8.** By our usual formula for posterior beliefs,

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

Since \( x_t \to 1 \) and \( \tilde{x}_t \geq x_t \) for all \( t \), \( \tilde{x}_t \to 1 \), whence \( \frac{\tilde{x}_t}{x_t} \to 1 \). In addition, since \( \frac{1 - x_t}{1 - \tilde{x}_t} \to 1 \), \( \frac{(1 - x_t)L_t}{(1 - \tilde{x}_t)L_t} \to 1 \). 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 9 (Recursive Decomposition).** Let \( (\theta_t)_t \) be a right-continuous measurable Markov process, \( f \) 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 \( (s_t)_t \) mapping into \( (A \cup \{\emptyset\}) \) as follows: \( s_t = \emptyset \) if \( \exists t' \leq t \) such that \( \theta_{t'} = \theta \) and \( \theta_{t''} \notin A \) for all \( t'' < t' \). If this is not true for any \( \theta \in A, \, s_t = \emptyset \).\(^{11}\) Then

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

\(^{11}\)In other words, \( s_t \) takes the value of the first \( \theta \in A \) that \( (\theta_t)_t \) hits.
where $P_s$ is defined as follows: $P_{s_t}$ is the probability measure on $A \cup \emptyset$ induced by $s_t$, and $P_s = \gamma \int_0^{+\infty} e^{-\gamma t} P_{s_t} dt$.

**Proof of Lemma 9.**

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

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 $z$ mapping into $(A \times [0, +\infty)) \cup \{\emptyset\}$ as follows: $z = (\theta, t)$ if $\theta_t = \theta$ and $\theta_t' \notin A$ for all $t' < t$. If this is not true for any $\theta \in A$ and $t \geq 0$, $z = \emptyset$. Define $P_z$ as the probability measure on $(A \times [0, +\infty)) \cup \{\emptyset\}$ induced by $z$.

Then we can write

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

as desired. ■

**Lemma 10.** In any equilibrium, for any realization of the stochastic process $\hat{L}$, if the policy path satisfies $\pi_{t_0} = 1$ and $\pi_{t_1} = 0$ for some $t_1 > t_0$, then $\pi_t = 0$ for all $t \geq t_1$. 

**Proof.**

Note that, for all $L$, there is $\rho \in [0, 1]$ independent of $x$ such that

\[
U_x(L, 0) = \rho \frac{r}{\gamma} + (1 - \rho) U_x(L, 1).
\]

\[\text{12} \text{In other words, } z \text{ takes the value of the first } \theta \in A \text{ that } (\theta_t)_t \text{ hits, and the time when it hits.}\]
This follows from Lemma 9, with the added observation that $\rho$ (that is, $P_s$ in the notation of Lemma 9) is independent of $x$ in this case because the stochastic process governing $(L, \pi)$ is independent of $x$ so long as $\pi = 0$.

$\rho$ is a function of the switching rate from $(L, 0)$ to $(L, 1)$. Because the path of play is always cadlag, it must be that $\rho > 0$.\(^{13}\)

Let $\hat{t} = \inf\{t \in [t_0, t_1] : \pi_t = 0\}$. Because the path of play must be cadlag, we have that $\pi_{\hat{t}} = 0$; $\lim_{t \downarrow \hat{t}} \pi_t = 0$; and $\lim_{t \uparrow \hat{t}} \pi_t = 1$.

By Condition (c), the second and third statements imply that $0 \in \alpha(L_{\hat{t}}, 0)$ and $0 \in \alpha(L_{\hat{t}}, 1)$, respectively.

Suppose there is $\omega$ for which all the above is true and $\pi_{\tilde{t}} = 1$ for some $\tilde{t} > \hat{t}$. Without loss of generality, assume that $\tilde{t} = \inf\{t \geq \hat{t} : \pi_t = 1\}$. Then Condition (c) implies that $1 \in \alpha(L_{\tilde{t}}, 0)$, i.e., $\alpha(L_{\tilde{t}}, 0) = [0, 1]$.

By Condition (ii), this implies that $U_{m(L_{\tilde{t}}, 0)}(L_{\tilde{t}}, 1) = U_{m(L_{\tilde{t}}, 0)}(L_{\tilde{t}}, 0)$, which in turn implies $U_{m(L_{\tilde{t}}, 0)}(L_{\tilde{t}}, 1) = U_{m(L_{\tilde{t}}, 0)}(L_{\tilde{t}}, 0) = \frac{r}{\gamma}$ since $\rho_{L_{\tilde{t}}} > 0$.

For any $L > 0$, we have $m(L, 1) > m(L, 0)$. In particular, $m(L_{\tilde{t}}, 1) > m(L_{\tilde{t}}, 0)$ which implies $U_{m(L_{\tilde{t}}, 1)}(L_{\tilde{t}}, 1) > U_{m(L_{\tilde{t}}, 0)}(L_{\tilde{t}}, 1)$ by Claim ??\(^{14}\) Thus

\[
U_{m(L_{\tilde{t}}, 1)}(L_{\tilde{t}}, 1) > U_{m(L_{\tilde{t}}, 0)}(L_{\tilde{t}}, 1) = \frac{r}{\gamma}
\]

\[
U_{m(L_{\tilde{t}}, 1)}(L_{\tilde{t}}, 1) > \rho_{L_{\tilde{t}}} \frac{r}{\gamma} + (1 - \rho_{L_{\tilde{t}}}) U_{m(L_{\tilde{t}}, 1)}(L_{\tilde{t}}, 1) = U_{m(L_{\tilde{t}}, 1)}(L_{\tilde{t}}, 0),
\]

which implies $\alpha(L_{\tilde{t}}, 1) = 1$ by Condition (ii), a contradiction.

\[\blacksquare\]

**Lemma 11.** In any equilibrium, if $L_{t_0} = 0$, then $\pi_t = 1$ for all $t \geq t_0$.

**Proof.**

Note that, if $L_{t_0} = 0$, then $k(t_0) \geq 1$, which implies $L_t = 0$ for all $t \geq t_0$ no matter what policy path is followed. In addition, $\beta(x, 0, \pi) = 1$ for all $x > 0$ no

\(^{13}\)As the path of play is right-continuous, starting at $(L, 0)$, the policy must remain at 0 for a positive amount of time w.p. 1, which implies the policy remains at 0 for a positive expected length of time, so $\rho > 0$.

\(^{14}\)The path starting at $(L_{\tilde{t}}, 1)$ cannot have the future policy be 0 w.p. 1 by the right-continuity of the policy path.
matter what $\pi$ is, as the flow payoff from being a member in this case will be either $b$ or $r$ depending on $\pi$, and hence always higher than $s$.

As in Lemma 10, for all $L$, there is $\rho \in (0, 1]$ independent of $x$ such that

$$U_x(L, 0) = \rho \frac{r}{\gamma} + (1 - \rho)U_x(L, 1).$$

In particular, this is true for $L = 0$.

By the same logic, there is $\rho' \in (0, 1]$ such that, for all $x > 0$,

$$U_x(0, 1) = \rho' \frac{b}{\gamma} + (1 - \rho')U_x(0, 1).$$

Taken together, the two equations imply that $U_x(0, 1) > U_x(0, 0)$ for all $x > 0$. In particular, $U_{m_0}(0, 1) > U_{m_0}(0, 0)$, where $m_0 = m(0, 0) = m(0, 1)$ is the population median. Hence $\alpha(0, 0) = \alpha(0, 1) = 1$. From Condition (c) and the fact that $L_t = 0$ for $t \geq t_0$, it follows that $\pi_t = 1$ for all $t \geq t_0$.

**Lemma 12.** Let $V_x(L)$ denote the value function of the agent with prior $x$ when the state variable is $L$. $x \mapsto V_x(L)$ is strictly increasing for all $L$ for all agents with prior $x$ that are in the experimenting organization at a set of times of a strictly positive measure (on the equilibrium path).

**Proof of lemma 12.**

Consider agents with priors $x' > x$. Let $V_{x'}^x(L)$ 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 $L$, $x'$’s expected flow payoff is $p(L, x')b + (1 - p(L, x'))c$. Because $x' > x$, we have $p(L, x') > p(L, x)$ and thus $p(L, x')b + (1 - p(L, x'))c > p(L, x)b + (1 - p(L, x))c$, 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 $V_{x'}^x(L) > V_x(L)$ for all agents with prior $x$ that are in the experimenting organization at a set of times of a strictly positive measure. Because
\[ V'_x(L) \geq V'_x(L), \text{ we then have } V'_x(L) \geq V'_x(L) > V_x(L), \text{ which implies that } V'_x(L) > V_x(L) \text{ for all agents with prior } x \text{ that are in the experimenting organization at a set of times of a strictly positive measure, as required.} \]

**Corollary 2.** \( V_x(L) = V_{p(L,x)}(0) \), which is increasing in \( p(L,x) \).

**Proof of Proposition 1.**

We first argue that, if \( \inf_{t \geq 0} V(p_t(m_t)) \geq \xi \), then experimenting forever is an equilibrium if the organization is experimenting at time \( t_0 = 0 \).

Consider the following strategy profile: \( \alpha(L,1) = 1 \) for all \( L \in \{0\} \cup [1, +\infty) \), \( \alpha(L,0) = 1 \) if \( V(p(L,m_0(L))) > \frac{\xi}{\gamma} \), and \( \beta(x,L,\pi) \) is given by (i) in the definition of 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 \( \pi_t = 1 \) for all \( t \geq t_0 \) if \( V(p(L_{t_0},m(L_{t_0},0))) > \frac{\xi}{\gamma} \) and \( \pi_t = 0 \) for all \( t \geq t_0 \) if \( V(p(L_{t_0},m(L_{t_0},0))) \leq \frac{\xi}{\gamma} \).

We can then check that Conditions (a)-(d) and (i)-(iii) hold.

Hence, 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. (In particular, the equilibrium is unique if there is no \( L \) for which \( p(L_{t_0},m(L_{t_0},0))) = \frac{\xi}{\gamma} \).

Let \( \sigma \) be an equilibrium. By Lemma 11, the risky policy must always be used after a success. By Lemma 10, after a switch from the risky policy to the safe policy, the safe policy will be used forever.

Now assume that \( \sigma \) is not such that, starting with policy 1, policy 1 will be used forever for sure. In other words, there is \( L \) for which \( 0 \in \alpha(L,1) \). By Condition (ii), this requires that \( \frac{\xi}{\gamma} \geq V_{m(L,1)}(L,1) \).

Let \( T \) be the (possibly random) time until the policy switches to 0, starting in state \( (L,1) \). Then \( V_{m(L,1)}(L,1) = E(W_T(p(L,m(L,1)))) \), where \( W_T(x) \) is as in Lemma 5 and the expectation is taken over \( T \). Now, recall that \( V(p(L,m(L,1))) > \frac{\xi}{\gamma} \) by assumption, and then \( W_T(p(L,m(L,1)))) > \frac{\xi}{\gamma} \) for all \( T > 0 \) by Lemma 5. Since equilibrium paths must be right-continuous, \( T > 0 \) w.p. 1. Hence \( E(W_T(p(L,m(L,1)))) > \)
Finally, suppose that \( \inf_{t \geq 0} V(p_t(m_t)) < \frac{x}{\gamma} \), so \( V(p_{t_0}(m_{t_0})) < \frac{x}{\gamma} \) for some \( t_0 \), and suppose there is an equilibrium where \( \pi_t = 1 \) for all \( t \) starting in state \((1, 1)\). In particular, this requires \( 1 \in \alpha(L_{t_0}, 1) \), which requires \( V(p_{t_0}(m_{t_0})) \geq \frac{x}{\gamma} \) 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, \( p_t(m_t) \geq \frac{2s}{b+s} \). For any \( \alpha > 0 \), \( p_t(m_t) \geq \frac{s}{\lambda b + (1-\lambda)s} \) for \( \lambda = \frac{1}{2\pi(t)} \) when \( f = f_\alpha \). Finally, \( p_t(m_t) \geq \frac{s}{b} \) if \( f \) has full support. The first two claims follow from Lemma 1, which in fact gives us that \( p_t(m_t) \downarrow \frac{2s}{b+s} \) as \( t \to \infty \) when \( f \) is uniform, and \( p_t(m_t) \downarrow \frac{s}{\lambda b + (1-\lambda)s} \) when \( f = f_\alpha \). The last claim is implied directly 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 6 and the fact that the function \( p_t(x) \) is strictly increasing in \( x \).

Third, we observe that, since \( V(x) \) is strictly increasing and continuous in \( x \) (Lemma 4), \( \inf_{t \geq 0} V(p_t(m_t)) = V(\inf_{t \geq 0} p_t(m_t)) \). 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

\[
t \left( \frac{s}{\lambda b + (1-\lambda)s} \right) = \frac{\ln \left( \frac{s}{b} \frac{\lambda(b-s)}{s} \right)}{b} = -\frac{\ln \lambda}{b}
\]

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

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

35
In particular, for $\alpha = 0$, this becomes
\[
\gamma V \left( \frac{2s}{b + s} \right) = \frac{2bs}{b + s} + \left( \frac{1}{2} \right)^{\frac{s}{b + s}} \frac{(b - s)s}{b + s} \frac{b}{\gamma + b}.
\]

On the other hand, for $x = \frac{s}{b}$, $t(x) = 0$. Substituting 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. Denote $\tilde{m}_t$ to be the median at time $t$ under $f$, and $m_t$ the median at time $t$ under a uniform density. It is sufficient to show that the asymptotic posterior of the median is exactly $\frac{2s}{b + s}$ 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 7 tells us that $\frac{1 - \tilde{m}_t}{1 - m_t} \to 1$. Moreover, Lemma 8, applied to the sequences $p_t(\tilde{m}_t)$ and $p_t(m_t)$, guarantees that the ratio $\frac{p_t(\tilde{m}_t)}{p_t(m_t)}$ converges to 1. \hfill ■

**Lemma 13.** There exist parameters such that $V \left( \frac{2s}{b + s} \right) \geq \frac{x}{\gamma}$.

**Proof of Lemma 13.**

It is easy to show that there exist parameters such that $\frac{2bs}{b + s} + \left( \frac{1}{2} \right)^{\frac{s}{b + s}} \frac{s(b - s)}{b + s} \frac{b}{\gamma + b} \geq r$ is satisfied. For example, suppose that $\gamma \approx \infty$. Then we need that $\frac{2bs}{b + s} \geq r$. For this, it is sufficient to have $s > \frac{r}{2}$ and $b \geq \frac{sr}{2s - r}$. \hfill ■

**Proof of Proposition 3.**

We first argue that $\tau$ is well-defined.

Recall the definition of $W_T(x)$ from Lemma 5. Let $t$ be the current time, $x = p_t(m_t)$ and let $t^*$ be the time at which $m_t$ would choose to stop if she had complete control over the policy. By an argument analogous to that in Lemma 5 it can be shown that if $t^* > t$ (that is, $W_{t^*}(x) > \frac{x}{\gamma}$), then $W_T(x)$ is strictly increasing in $T$ for $T \in [t, t^*]$ and strictly decreasing in $T$ for $T > t^*$. Hence $W_T(x) > \frac{x}{\gamma}$ for all $T \in (t, t^*)$. In addition, we know that $W_T(x) \to V(x)$ as $t \to +\infty$. It follows that there is a unique $\tau(t)$ for which $W_{\tau(t)}(x) = \frac{x}{\gamma}$ unless $V(x) \geq \frac{x}{\gamma}$, in which case $\tau(t) = +\infty$. 36
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, $t_0(\sigma) \leq \tau(0)$, as otherwise $m_0$ would switch to the safe policy at time 0. As in previous equilibria, 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. In a pure strategy equilibrium, there must be a time $t_1(\sigma) > t_0(\sigma)$ for which experimentation stops in this continuation (or $t_1(\sigma) = \infty$). If $t_1(\sigma)$ is finite, then it must be that $t_1(\sigma) = \tau(t_0(\sigma))$. To see why, suppose that $t_1(\sigma) > \tau(t_0(\sigma))$. In this case, for small $\epsilon > 0$, $m_{t_0(\sigma)+\epsilon}$ would strictly prefer to stop experimenting, a contradiction. On the other hand, if $t_1(\sigma) < \tau(t_0(\sigma))$, then $\tau_0(\sigma)$ would strictly prefer to deviate from the equilibrium path. If $t_1(\sigma) = +\infty$, we have a contradiction unless $V(p_{t_0(\sigma)}(m_{t_0(\sigma)})) = \frac{x}{\gamma}$, and in this case it must still be the case that $t_1(\sigma) = \tau(t_0(\sigma))$.

We now 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 must show that, for $t$ not in the sequence of stopping times, $m_t$ weakly prefers to continue experimenting. Let $t \in (t_n(\sigma), t_{n+1}(\sigma))$. Since $t > t_n(\sigma)$, we have $\tau(t) \geq \tau(t_n(\sigma)) = t_{n+1}(\sigma)$. Hence $W_{t_{n+1}(\sigma)}(p_t(m_t)) \geq \frac{x}{\gamma}$, as desired. ■

B A Model of Bad News

Lemma 14. 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

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

Proof of lemma 14.

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 $1 - P_t = (1 - x)(1 - e^{-bt})$.

Then

$$V(x) = \int_0^{\infty} (xb + (1 - P_t)r) e^{-\gamma \tau} d\tau$$
$$= \int_0^{\infty} (xb + (1 - x)(1 - e^{-bt})r) e^{-\gamma \tau} d\tau$$
$$= (xb + (1 - x)r) \int_0^{\infty} e^{-\gamma \tau} d\tau - (1 - x)r \int_0^{\infty} e^{-(\gamma+b)\tau} d\tau$$
$$= (xb + (1 - x)r) \frac{1}{\gamma} - (1 - x)r \frac{1}{\gamma + b}$$

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 $m_t = \frac{1}{2} \frac{12s+(b-s)e^{bt}}{s+(b-s)e^{bt}}$. 


Substituting the above formula for \( m_t \) into \( p_t(m_t) = \frac{m_t}{m_t + e^{-bt}(1-m_t)} \), we obtain

\[
p_t(m_t) = \frac{1}{2} \left( \frac{2s + (b-s)e^{bt}}{s + (b-s)e^{bt}} \right) + e^{-bt} \left( 1 - \frac{1}{2} \frac{2s + (b-s)e^{bt}}{s + (b-s)e^{bt}} \right) = \frac{2s + (b-s)e^{bt}}{2s + (b-s)(1 + e^{bt})}
\]

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 \( p'(e) \propto (2s + (b-s)(1 + e))(b-s) - (2s + (b-s)e)(b-s) \propto b-s > 0 \). Moreover, \( t \mapsto e^{bt} \) is strictly increasing. It follows that \( t \mapsto p_t(m_t) \) is strictly increasing.

Suppose first that \( f \) is non-decreasing and \( V \left( \frac{2s}{b+s} \right) \geq \frac{\zeta}{\gamma} \). Because \( \frac{b+s}{2b} > \frac{2s}{s+b} \), the fact that \( V \left( \frac{2s}{b+s} \right) \geq \frac{\zeta}{\gamma} \) implies that \( V \left( \frac{b+s}{2b} \right) > \frac{\zeta}{\gamma} \). Because, by claim 1, \( p_0(m_0) = \frac{b+s}{2b} \) and \( t \mapsto p_t(m_t) \) is strictly increasing and lemma 14 implies that \( V \) is increasing, this implies that \( V(p_t(m_t)) > \frac{\zeta}{\gamma} \) for all \( t \).

Then, using the fact that \( f \) is non-decreasing, we can use an argument similar to the one used in the proof of Proposition 1 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 proof of Proposition 1 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 \( m \mapsto p_t(m) \) 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 1.

We now show that there exists \( T \) such that for all \( t \geq T \), if no failures have been observed during \([0, t]\), then \( V(p_t(m_t)) \geq \frac{\zeta}{\gamma} \). First note that, because in a model of bad news agents do not leave the organization, \( \lim_{t \to \infty} y_t < 1 \). Moreover, \( \lim_{t \to \infty} e^{-bt} = 0 \). This implies that \( \lim_{t \to \infty} p_t(m_t) = \lim_{t \to \infty} \frac{y_t}{y_t + e^{-bt}(1-y_t)} = 1 \). Then, provided that no failures have been observed during \([0, t]\), we have \( \lim_{t \to \infty} p_t(m_t) = \frac{b+s}{2b} \).
\[ \lim_{t \to \infty} \frac{y_t e^{bt}}{y_t + e^{bt}} = 1, \quad \lim_{t \to \infty} V(p_t(m_t)) = V(1) \] 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 \( t_+ \) 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

\[
\int_0^{t_+} (xb + (1 - P_\tau)r) e^{-\gamma \tau} d\tau + \int_{t_+}^{\infty} re^{-\gamma \tau} d\tau \\
= \int_0^{t_+} (xb + (1 - x)(1 - e^{-bt})r) e^{-\gamma \tau} d\tau + \int_{t_+}^{\infty} re^{-\gamma \tau} d\tau \\
= (xb + (1 - x)r) \int_0^{t_+} e^{-\gamma \tau} d\tau - (1 - x)r \int_{t_+}^{t_+} e^{-(\gamma + b)\tau} d\tau + \int_{t_+}^{\infty} re^{-\gamma \tau} d\tau \\
= (xb + (1 - x)r) \frac{1 - e^{-\gamma t_+}}{\gamma} - (1 - x)r \frac{1 - e^{-(\gamma + b)t_+}}{\gamma + b} + e^{-\gamma t_+} \frac{r}{\gamma}
\]

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

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

Proof of claim 3.

Suppose for the sake of contradiction that this is not the case. Then there exists a non-empty subset \( B \subseteq [t_-, t_0) \) such that for all \( t \in B, m_t \) continues experimenting. Let \( t^1 = \sup\{t : t \in B\} \).
Then for all $\epsilon > 0$ sufficiently small there exist $\tau$ and $t^2$ such that $m_\tau$ continues experimenting, $m_{t^2}$ stops experimenting and $t^2 - \tau \in (0, \epsilon]$. In particular, if $t^1 = \max\{t : t \in B\}$, then take $\tau = t^1$ and $t^2 = t^1 + \epsilon$ for some $\epsilon < t^0 - t^1$. If $t^1 \neq \max\{t : t \in B\}$, then, because $m_{t^0}$ stops experimenting, we have $t^1 < t^0$. Moreover, in this case, we have $t^1 \not\in B$ and, by definition of the supremum, for all $\epsilon > 0$ there exists $\tau \in B$ such that $t^1 - \tau \in (0, \epsilon)$. Then take $t^2 = t^1$ and we are done.

Thus for all $\epsilon' > 0$ sufficiently small there exist $\tau$ and $t^2$ such that $m_\tau$ continues experimenting, $m_{t^2}$ stops experimenting and $t^2 - \tau = \epsilon$ for some $\epsilon < \epsilon'$.

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

The payoff to stopping experimentation is $\frac{\tau}{\gamma}$. Then if $W_{t^2-\tau}(p_\tau(m_\tau)) < \frac{\tau}{\gamma}$, the median strictly prefers to stop experimenting. This is equivalent to

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

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

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

By L’Hospital’s rule, when we take the limit as $\epsilon \to 0$, we obtain $\frac{p_\tau(m_\tau)(b - r)}{(1 - p_\tau(m_\tau))r} \frac{\gamma + b}{\gamma} < 1$. Equivalently, $\frac{p_\tau(m_\tau)(b - r)}{(1 - p_\tau(m_\tau))r} < 1$, or $p_\tau(m_\tau)b < r$. 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. ■

Claim 4. If for all $t \in [t, t^0)$ 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 $T' \subseteq [t, t^0)$ 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.
Note that, because the median switches to the safe policy after a time period of length \( t_n \), the payoff to \( m_t \) from continuing experimentation is

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

The payoff to stopping experimentation is \( \frac{\xi}{r} \). Then if \( \frac{1 - e^{-\gamma t+}}{\gamma} p_t(m_t) b + \frac{e^{-\gamma t+}}{\gamma} r > \frac{\xi}{r} \), 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. \( \blacksquare \)

Let \( t_{2n+1} = \sup \left\{ t : V(p_t(m_t)) \leq \frac{\xi}{r} \right\} \) denote the largest time for which the median stops experimenting.

Let \( T^1 = \{ t : p_t(m_t)b \leq r \} \) and \( T^2 = \{ t : p_t(m_t)b > r \} \). Our genericity assumption (Assumption 1) implies that \( T^1 \) and \( T^2 \) are finite collections of intervals. Enumerate the intervals such that \( T^1 = \bigcup_{i=0}^{n} [t_i, t_{i+1}] \).

Suppose first that \( p_t(m_t)b \leq r \) 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 \( W_{t_{2n+1}}(x) \) denotes the value function of an agent with belief \( x \) (at time \( t \)) given this continuation equilibrium path. We then let \( t_{2n-1} = \sup \left\{ t < t_{2n} : W_{t_{2n+1}}(p_t(m_t)) \leq \frac{\xi}{r} \right\} \).

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

Since \( F \) admits a continuous density, \( t \mapsto W_{t_{2n}}(p_t(m_t)) \) is continuous, which implies that we must have \( t_{2n-1} = \max \left\{ t < t_{2n} : W_{t_{2n}}(p_t(m_t)) \leq \frac{\xi}{r} \right\} \). Note that it is then consistent with equilibrium for the median \( m_{t_{2n}} \) to stop experimenting.
Now note that if \( W_{t_{2n-1}}(p_t(m_{t_{2n-1}})) = \frac{x}{n} \), then \( p_{t_{2n-1}}(m_{t_{2n-1}})b < r \). By continuity, there exists an interval \([t_i, t_j]\) in \( T^1\) such that \( t_{2n-1} \in [t_i, t_j]\) (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 \( W_{t_{2n-2}}(p_t(m_t)) \leq \frac{x}{n} \). 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 \( W_{t_{2k+2}}(p_t(m_{t_{2k+1}})) = \frac{x}{n} \) but also \( \frac{\partial}{\partial t} W_{t_{2k+2}}(p_t(m_t))|_{t=t_{2k+1}} > 0 \), that is, \( W_{t_{2k+2}}(p_t(m_t)) < \frac{x}{n} \) for all \( t < t_{2k+1} \) close enough to \( t_{2k+1} \). Thus, even if we allow \( m_{t_{2k+1}} \) 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 \( p_{t_{2k}}(m_{t_{2k}})b - r = 0 \) but also \( \frac{\partial}{\partial t} p_t(m_t)|_{t=t_{2k}} < 0 \), that is, \( p_t(m_t)b - r > 0 \) for all \( t < t_{2k} \) close enough to \( t_{2k} \). Thus, even if we allow \( m_{t_{2k}} \) to stop experimenting, all agents in \((t_{2k} - \epsilon, t_{2k})\) must stop as they strictly prefer to do so.

## C Imperfectly Informative Experimentation Technology

**Lemma 15.** 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 \( L(k, t) \leq \frac{x(b-s)}{(1-x)(s-c)} \).

**Proof of lemma 15.**

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 \( L(k, t) \leq \frac{x(b-s)}{(1-x)(s-c)} \).

**Lemma 16.** 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 16.**

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 $F(z) = \int_0^z (1 - x)\alpha c = \frac{z - c}{\alpha + 1} (1 - (1 - z)\alpha + 1)$.

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 $1 - m(L) = (1 - y(L))2^{-\frac{1}{\alpha + 1}}$, or $m(L) = 1 - \kappa + \kappa y(L)$ for $\kappa = 2^{-\frac{1}{\alpha + 1}}$.

Note that $p(L, m(L)) = \frac{1}{1 + (\frac{1}{m(L)} - 1)L}$. 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 $\frac{\partial}{\partial L} \left(1 + \left(\frac{1}{m(L)} - 1\right)L\right) = \frac{\partial}{\partial L} \left(\left(\frac{1}{m(L)} - 1\right)L\right) = \frac{1}{m(L)} - 1 - \frac{L}{(m(L))^{\alpha + 1}} m'(L)$.

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

**Lemma 17.** 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 17.**

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,
\[ p(L, m(L)) = \frac{2Ls + 5 - \pi}{L(s + b) + b - \pi}, \] which is decreasing in \( L \). Moreover, with the notation we have defined, the formula for \( y(L) \) can be written as \( y = \frac{Ls}{Ls + b - \pi} \).

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, \( m = \frac{1 + b_1}{2} - \frac{a_1b_1}{2a_2} + \frac{a_1}{2a_2}y \). Equivalently, \( m = \left(1 - \frac{1}{2a_2}\right) + \frac{a_1}{2a_2}y = \left(1 - \frac{1}{2a_2}\right) + \frac{a_1}{2a_2} \frac{Ls}{Ls + b - \pi} \). Then

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

For \( L < L_1 \), we have \( a_1(m - y) = a_1(b_1 - m) + a_2(1 - b_1) \), that is, \( 2a_1m = a_1b_1 + a_2(1 - b_1) + a_1y = 1 + a_1y \), so \( m = \frac{1}{2a_1} + \frac{1}{2}y \), and

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

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

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

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

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

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

so \( L_1 \) is a global minimizer if we take \( a_1 \in \left(1, \frac{3}{2}\right) \).\[\square\]
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 \( p(L, y(L)) = \frac{y(L)}{y(L)+(1-y(L))L} \), this is equivalent to \( \frac{y(L)}{y(L)+(1-y(L))L} = \frac{s-c}{b-c} \). Solving for \( y(L) \), we obtain \( y(L) = \frac{s-c}{s-c+(b-s)\frac{1}{L}} \).

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

Because \( p(L, m(L)) = \frac{1}{1+\frac{1}{m(L)-1}} \), 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 \( \lim_{L \to \infty} p(L, m(L)) = \lim_{L \to \infty} \frac{2L(s-c)+b-s}{L(2(s-c)+b-s)+b-s} = \frac{2(s-c)}{(b-c)+(s-c)} \).

By lemma 16, 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 (the proof of Proposition 1). In particular, because, by lemma 12, \( x \mapsto V(x) \) 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 \( V(p(L, m(L))) \geq \frac{\gamma}{7} \) for all \( L \), it is enough to ensure that \( \lim_{L \to \infty} V(p(L, m(L))) \geq \frac{\gamma}{7} \). 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 \( V\left(\frac{2(s-c)}{(b-c)+(s-c)}\right) \geq \frac{\gamma}{7} \).

Next, given that we have shown that \( V\left(\frac{2(s-c)}{(b-c)+(s-c)}\right) \geq \frac{\gamma}{7} \) implies that \( V(p(L, m(L))) \geq \frac{\gamma}{7} \) for all \( L \) for a uniform distribution of priors, by an argument similar to the one in the proof of Proposition 1 the hypothesis that \( f \) is non-decreasing ensures that that we have \( V(p(L, m(L))) \geq \frac{\gamma}{7} \) for all \( L \) under \( f \).
To show that if $V \left( \frac{2(s-c)}{(b-c)+(s-c)} \right) < \frac{r}{\gamma}$, there is no equilibrium in which the organization experiments forever, we can use the result in lemma 8 that $\lim_{L \to \infty} p(L, m(L)) = \lim_{L \to \infty} p(L, \tilde{m}(L)) = \frac{2(s-c)}{(b-c)+(s-c)}$ and an argument similar to the one used in the proof of Proposition 1.

Finally, we show that there exist parameter values such that $V \left( \frac{2(s-c)}{(b-c)+(s-c)} \right) \geq \frac{r}{\gamma}$ 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 $V(x) \geq x \frac{b}{\gamma} + (1 - x) \frac{c}{\gamma}$, 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 $V \left( \frac{2(s-c)}{(b-c)+(s-c)} \right) \geq \frac{r}{\gamma}$ 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 $V_x^{\epsilon}(L)$ 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 $V_{x}(L)$ 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

$$V_{m(L)}^{\epsilon}(L) = p(L, m(L))V_{m(L),G}^{\epsilon}(L) + (1 - p(L, m(L)))V_{m(L),B}^{\epsilon}(L)$$

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) =
\( V^0_{p(L,m(L))}(0) \) and \( x \mapsto V_x(0) \) 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 17.15

Thus we fix \( f \) constructed in the proof of Lemma 17. Observe that the construction is such that \( L \mapsto p(L,m(L)) \) has a kink at \( L^* \).

Note that \( V^0_{m(L)}(L) \) does not depend on \( r \) and the density \( f \) constructed in the proof of Lemma 17 does not depend on \( r \), which implies that \( L^* \) does not depend on \( r \). Then we can choose \( r \) such that

\[
V^0_{m(L^*)}(L^*) = r \tag{1}
\]

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

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

\[
\hat{V}_{x,L'}(L) = \int_0^\infty \gamma e^{-\gamma t} E[u_x(h^t)\mathbb{1}_{\exists s \in [0,t] : L_s = L'|L_0 = L}]dt
\]

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

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 \( \left( \tilde{L}_\tau \right)_\tau \).

Intuitively, \( T_{x,L'}(L) \) is the weighted discounted probability that the stochastic

---

15Technically, we also need the condition that \( V^0_{m(L^*)}(L^*) < \lim_{L \to \infty} V^0_{m(L)}(L) \), but this is also satisfied by the construction in Lemma 17.
process \((L_s)_s\) hits the value \(L'\) at least once, \(V_{x,L'}(L)\) 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 \(V_{x,L'}(L)\) 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

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

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

Let \(T_{x,L^+}(L) = \int_0^\infty \gamma e^{-\gamma t}Pr[3s \in (0, t) : L_s = L'|L_0 = L]dt\), \(\tilde{V}_{x,L^+}(L) = \int_0^\infty \gamma e^{-\gamma t}E[u_x(h^t)1_{3s \in (0, t); L_s = L'}|L_0 = L]dt\) and \(\tilde{V}_{x,L^*}(L^+) = \frac{V_{x,L^+}(L^+)}{1 - T_{x,L^+}(L^+)}\). Observe that \(\tilde{V}_{x,L^*}(L^+) = \lim_{L \searrow L^*} \tilde{V}_{x,L^*}(L)\) and \(T_{x,L^*}(L^+) = \lim_{L \searrow L^*} T_{x,L^*}(L)\). Let \(\tilde{W}_x^\epsilon = \tilde{V}_{x,L^*}(L^+)\).

Let \(W_x^\epsilon = \lim_{L \searrow L^*} V_x^\epsilon(L)\). \(W_x^\epsilon\) 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,

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

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

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

where the second equality follows from the fact that \(\tilde{W}_x^\epsilon = V_x^0(L^*)\) because \(\tilde{W}_x^\epsilon\) 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 (3) into (2), we obtain

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

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

(4)

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 12 and the fact that \( L \mapsto p(L, m(L)) \) has a kink at \( L^* \), the last equality follows from (1) and Corollary 2.

On the other hand, for \( L \notin (L^* - \delta, L^* + \delta) \) there exists \( K' > 0 \) such that

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

(6)

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 (1) and Corollary 2.

(4) implies that \( V^\varepsilon_{m(L)}(L) \geq r \) 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 (5), 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 \( \frac{T_{m(L),L^*}(L)}{1 - T_{m(L),L^*}(L^{*+})} \epsilon \left( V^{0}_{m(L)}(L^*) - r \right) \geq T_{m(L),L^*}(L)\epsilon \frac{V^{0}_{m(L)}(L^*) - r}{1 - (1 - \epsilon)T_{m(L),L^*}(L^{*+})} \), it is enough to show that \( K|L - L^*| \geq \frac{T_{m(L),L^*}(L)}{1 - T_{m(L),L^*}(L^{*+})} \epsilon \left( V^{0}_{m(L)}(L^*) - r \right) \). Then we need to show that \( \epsilon \leq \)
we require that \( \epsilon \leq \inf_{L \in (L^* - \delta, L^* + \delta)} K^{1 - T_{m(L), L^*}(L^*)} \frac{|L - L^*|}{T_{m(L), L^*}(L^*)} \). Observe that \( r = V^0_{m(L^*)}(L^*) \) by (1). Then we require that

\[
\epsilon \leq \inf_{L \in (L^* - \delta, L^* + \delta)} K^{1 - T_{m(L), L^*}(L^*)} \frac{|L - L^*|}{T_{m(L), L^*}(L^*)} V^0_{m(L)}(L^*) - V^0_{m(L^*)}(L^*)
\]

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

Suppose next that \( L \notin (L^* - \delta, L^* + \delta) \). Then, because \( V^0_{m(L)}(L) \geq r + K' m\) by (6), it is enough to show that \( r + K' \geq r - T_{m(L), L^*}(L) \epsilon \frac{r - V^0_{m(L^*)}(L^*)}{1 - T_{m(L), L^*}(L^*)} \). This is equivalent to \( K' \geq T_{m(L), L^*}(L) \epsilon \frac{V^0_{m(L)}(L^*) - r}{1 - T_{m(L), L^*}(L^*)} \). Then, because \( T_{m(L), L^*}(L) \epsilon \frac{V^0_{m(L)}(L^*) - r}{1 - T_{m(L), L^*}(L^*)} \), it is enough to show that \( K' \geq \frac{T_{m(L), L^*}(L)}{1 - T_{m(L), L^*}(L^*)} \epsilon (V^0_{m(L)}(L^*) - r) \). Thus we need to show that \( \epsilon \leq \inf_{L \notin (L^* - \delta, L^* + \delta)} K' \frac{1 - T_{m(L), L^*}(L^*)}{1 - T_{m(L), L^*}(L^*)} \frac{|L - L^*|}{V^0_{m(L^*)}(L^*)} - V^0_{m(L^*)}(L^*) \).

Because \( T_{m(L), L^*}(L) \leq 1 \), it is enough to show that \( \epsilon \leq \inf_{L \notin (L^* - \delta, L^* + \delta)} K'(1 - T_{m(L), L^*}(L^*)) \frac{1}{b - r} \).

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

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

**Proof of Corollary 1.**

Take the example constructed in Proposition 6, and assume that \( L_0 > L^* \).\(^{16}\) Let
$P_\theta(L_0)$ be the probability that there exists $t < \infty$ such that the organization stops experimenting at $t$ when the state is $\theta \in \{G, B\}$. 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 $L_t \in \left(\left(\frac{b}{c}\right) L^*, L^*\right]$ 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 $L^*$.

It can be shown that $Q_G(L_0, L^*) = 1$ for all $L_0 > L^*$, but $\lim_{L_0 \to \infty} Q_B(L_0, L^*) = 0$.

To show this, note that, when $\theta = G$, $(l_t)_t = l_0 + (b - c) t - [\ln(b) - \ln(c)] N(t)$, where $(N(t))_t$ is a Poisson process with rate $b$. This can be written as a random walk: $l_t - l_0 = \sum_{i=0}^{t} S_i$, where $S_i = b - c - [\ln(b) - \ln(c)] N_i$, and $N_i \sim P(b)$. Note that $E(S_i) = b - c - b (\ln(b) - \ln(c)) < 0$\footnote{Denote $\frac{b}{c} = 1 + x$. 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 $x > 0$.}. Then, by the strong law of large numbers, $\frac{l_t}{t} \xrightarrow{t \to \infty} E(S_i) < 0$ a.s., whence $l_t \xrightarrow{t \to \infty} -\infty$ a.s., implying the first claim.

On the other hand, when $\theta = B$, $(l_t)_t = l_0 + (b - c) t - \ln(b - c) N(t)$, 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 $S_i = b - c - [\ln(b) - \ln(c)] N_i$, $N_i \sim P(c)$, and $E(S_i) = b - c - c (\ln(b) - \ln(c)) > 0$. By the strong law of large numbers, $l_t \xrightarrow{t \to \infty} +\infty$ a.s. Now, suppose that $\lim_{L \to \infty} Q_B(L, L^*) > 0$. It follows that $\lim_{L \to \infty} Q_B(L, \frac{b}{c}L) = 1$. Since in fact $Q_B(L, \frac{b}{c}L)$ must be independent of $L$, it follows that $Q_B(L, \frac{b}{c}L) = 1$ and hence $(l_t)_t$ is recurrent (Durrett 2010), but this contradicts the fact that $l_t \xrightarrow{t \to \infty} +\infty$ a.s.

This implies that $P_B(L_0) \leq Q_B(L_0) \to 0$ as $L_0 \to \infty$. On the other hand, $P_G(L_0) \geq Q_G(L_0) \inf_{L \in \left(\left(\frac{b}{c}\right)L^*, L^*\right]} P_G(L) > 0$. \hfill \blacksquare

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

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 $L_0$ and leaving $f$ unchanged.
that the state variable is \( L \) and the risky policy is of type \( \theta \). Then

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

**Lemma 18.** \( V_{x,G}(L) \) and \( V_{x,B}(L) \) satisfy the following equations for \( L \in \mathbb{R}_{\geq 0} \setminus \mathcal{L} \):

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

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

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

**Proof of lemma 18.**

Because, by lemma 15, 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

\[
\mathbb{I}_{L \leq \frac{x(b - s)}{(1 - x)(s - c)}} \frac{b + \mathbb{I}_{L \geq \frac{x(b - s)}{(1 - x)(s - c)}} \frac{s}{(1 - x)(s - c)}}{(1 - x)(s - c)} \frac{b + s - \mathbb{I}_{L \leq \frac{x(b - s)}{(1 - x)(s - c)}} \frac{s}{(1 - x)(s - c)}}{(1 - x)(s - c)} \frac{b}{(1 - x)(s - c)} + s
\]

Similarly, provided that the risky technology is bad, an agent’s flow payoff during the time period of length \( \epsilon \) is

\[
\mathbb{I}_{L \leq \frac{x(b - s)}{(1 - x)(s - c)}} \frac{b}{(1 - x)(s - c)} + s.
\]

Provided that the risky technology is good, with probability approximately equal to \( e^{-b\epsilon} \), a success arrives within the time period of length \( \epsilon \), which changes the state from \( L \) to \( L e^{(b-c)\epsilon} \). With probability approximately equal to \( e^{-b\epsilon} \), a success does not arrive within this time period, which changes the state from \( L \) to \( L e^{(b-c)\epsilon} \).

Then we have

\[
V_{x,G}(L) \approx (1 - e^{-\gamma \epsilon}) \left( \mathbb{I}_{L \leq \frac{x(b - s)}{(1 - x)(s - c)}} (b - s) + s \right) + e^{-\gamma \epsilon} \left[ e^{-b\epsilon} V_{x,G} \left( Le^{(b-c)\epsilon} \right) \right]
\]

\[
+ e^{-\gamma \epsilon} \left[ e^{-b\epsilon} V_{x,G} \left( Le^{(b-c)\epsilon} \right) \right]
\]

53
Subtracting \( V_{x,G}(Le^{(b-c)\epsilon}) \) from both sides, we obtain

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

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 \( V_{x,\theta}(L) = \frac{\zeta}{\gamma} \) for \( \theta \in \{B, G\} \) because, whenever \( L = 0 \), 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 \( \int_0^\infty e^{-\gamma\tau} rd\tau = \frac{r}{\gamma} \).

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.

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

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

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

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

**Proof of lemma 19.**
Note that \( \ln(L_c^e) = \ln(e^{l+\ln \frac{1}{b}}) = l + \ln \frac{e}{b} \) and that
\[
\frac{\partial U_{x,G}(l)}{\partial l} = \frac{\partial V_{x,G}(L)}{\partial L} \frac{\partial L}{\partial l} = \frac{\partial V_{x,G}(L)}{\partial L} \frac{\partial}{\partial l} \left( e^l \right) = \frac{\partial V_{x,G}(L)}{\partial L} e^l = \frac{\partial V_{x,G}(L)}{\partial L} L
\]
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 18, we obtain the desired equation for \( U_{x,G}(l) \) in the statement of lemma 19. 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
\[ d = \ln \frac{x(b-c)}{(1-x)(d-c)} \]
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 \( a = -\ln \frac{e}{b} \) 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 20.**
\[ U_0(l) = De^{\omega_0 l} + C_0 \]
for \( \omega_0 \) satisfying \( (b-c)\omega_0 = \gamma + b - be^{-\omega_0 a} \), \( C_0 = b \) and some constant \( D \).

**Proof of lemma 20.**

Note that if \( U_0(l) = De^{\omega_0 l} + C_0 \), then \( U_0'(l) = D\omega_0 e^{\omega_0 l} \).

Suppose that \( l \leq d \). Then the equation from lemma 19 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
\[
(b-c)\omega_0 De^{\omega_0 l} = (\gamma + b) \left( De^{\omega_0 l} + C_0 \right) - b \left( De^{\omega_0 (l-a)} + C_0 \right) - \gamma b
\]
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

$$(b - c)\omega_0 D_0 e^{\omega_0 l} = (\gamma + b)D e^{\omega_0 (l-a)}$$

Canceling $D e^{\omega_0 l}$ from both sides, we obtain

$$(b - c)\omega_0 = \gamma + b - be^{-\omega_0 a}$$

which pins down $\omega_0$.

Lemma 21. If $U_0(l) = D e^{\omega_0 l} + b$ for $\omega_0$ satisfying $(b - c)\omega_0 = \gamma + b - be^{-\omega_0 a}$ and some constant $D$, then

$$U_1(l) = De^{\omega_0 l} + a_0 e^{\omega_1 l} + C_1$$

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

Proof of lemma 21.

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

Suppose that $l \in (d, d + a]$, so that $l - a \in (d - a, d]$. Then the equation from lemma 19 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

$$(b - c)\left(\omega_0 D_1 e^{\omega_0 l} + \omega_1 a_0 e^{\omega_1 l}\right) = (\gamma + b)\left(D_1 e^{\omega_0 l} + a_0 e^{\omega_1 l} + C_1\right) - b\left(D e^{\omega_0 (l-a)} + C_0\right) - \gamma s$$

In order for the constant terms to cancel out, we need $0 = (\gamma + b)C_1 - bC_0 - \gamma s$. That is, we need $C_1 = \frac{b^2 + \gamma s}{\gamma + b}$.
Then the equation simplifies to

\[(b - c) \left( \omega_0 D_1 e^{\omega_0 l} + \omega_1 a_0 e^{\omega_1 l} \right) = (\gamma + b) \left( D_1 e^{\omega_0 l} + a_0 e^{\omega_1 l} \right) - b D_0 e^{\omega_0 (l-a)} \]

To match the coefficients, we need that \((b - c) \omega_0 D_1 e^{\omega_0 l} = (\gamma + b) D_1 e^{\omega_0 l} - b D_0 e^{\omega_0 (l-a)}\). This equation holds for all \(l\) if \(D_1 = D_0 \), and there can only be one value of \(D_1\) that works for all \(l\), so \(D_1 = D_0\).

Then the equation simplifies to

\[(b - c) \omega_1 a_0 e^{\omega_1 l} = (\gamma + b) a_0 e^{\omega_1 l} \]

which implies that \(\omega_1 = \frac{\gamma + b}{b-c}\). \(\blacksquare\)

**Lemma 22.**

\[ U_n(l) = P_n(l) e^{\omega_1 l} + D_n e^{\omega_0 l} + C_n \]

where \(P_n\) is a polynomial of degree \(n - 1\), \(\omega_0\) satisfies \((b - c) \omega_0 = \gamma + b - b e^{\omega_0 a}\) for \(a = -\ln \frac{\epsilon}{b}\), \(\omega_1 = \frac{\gamma + b}{b-c}\) 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 \(P_n\) satisfies

\[ P'_n(l) = -\frac{b}{(b-c) e^{\omega_1 a}} P_{n-1}(l-a) \]

for all \(n \geq 1\).

**Proof of lemma 22.**

We will prove the lemma by induction.

Lemma 21 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

\[ U'_{k+1}(l) = P'_{k+1}(l) e^{\omega_1 l} + P_{k+1}(l) \omega_1 e^{\omega_1 l} + D_{k+1} \omega_0 e^{\omega_0 l} \]

for all \(n \geq 1\).
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
\[(b - c) \left( P'_{k+1}(l)e^{\omega_1 l} + P_{k+1}(l)\omega_1 e^{\omega_1 l} + D_{k+1}\omega_0 e^{\omega_0 l} \right) =
(\gamma + b)\left( P_{k+1}(l)e^{\omega_1 l} + D_{k+1}e^{\omega_0 l} + C_{k+1} \right) - b(P_k(l - a)e^{\omega_1(l-a)} + De^{\omega_0(l-a)} + C_k) - \gamma s\]

For the constants to cancel out, it must be that \(0 = (\gamma + b)C_{k+1} - bC_k - \gamma s\). That is, we need \(C_{k+1} = \frac{bC_k + \gamma s}{\gamma + b}\). This pins down \(C_n\) 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
\[(b - c) \left( P'_{k+1}(l)e^{\omega_1 l} + P_{k+1}(l)\omega_1 e^{\omega_1 l} + D_{k+1}\omega_0 e^{\omega_0 l} \right) =
(\gamma + b)(P_{k+1}(l)e^{\omega_1 l} + D_{k+1}e^{\omega_0 l}) - b(P_k(l - a)e^{\omega_1(l-a)} + De^{\omega_0(l-a)})\]

As in lemma 21, for the terms multiplied by \(e^{\omega_0 l}\) to cancel out, we need that \(D_{k+1} = D\). The equation then simplifies to
\[(b - c) \left( P'_{k+1}(l)e^{\omega_1 l} + P_{k+1}(l)\omega_1 e^{\omega_1 l} \right) = (\gamma + b)P_{k+1}(l)e^{\omega_1 l} - bP_k(l - a)e^{\omega_1(l-a)}\]

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
\[(b - c)P'_{k+1}(l)e^{\omega_1 l} = -bP_k(l - a)e^{\omega_1(l-a)}\]

\[\blacksquare\]

**Lemma 23.** Suppose that organization is subject to a congestion effect. Then in the
equilibrium in which the organization experiments forever we have

\[ y_t = \frac{e^{-bt}b - s(e^{-bt} + \alpha - 2) - \sqrt{(e^{-bt}b - s(e^{-bt} + \alpha - 2))^2 - 4s^2(1 - e^{-bt})(1 - \alpha)}}{2s(1 - e^{-bt})(1 - \alpha)} \]

and

\[ p_t(y_t) = \frac{e^{-bt}(e^{-bt}b - s(e^{-bt} + \alpha - 2) - \sqrt{(e^{-bt}b - s(e^{-bt} + \alpha - 2))^2 - 4s^2(1 - e^{-bt})(1 - \alpha)})}{(1 - e^{-bt})(2s(1 - \alpha) - (e^{-bt}b - s(e^{-bt} + \alpha - 2) - \sqrt{(e^{-bt}b - s(e^{-bt} + \alpha - 2))^2 - 4s^2(1 - e^{-bt})(1 - \alpha)})} \].

Moreover, \( \lim_{t \to \infty} p_t(y_t) = \frac{s}{b + \alpha} \) and \( \lim_{t \to \infty} p_t(m_t) = \frac{2\alpha}{b + \alpha} \).

**Proof of lemma 23.**

In the equilibrium in which the organization experiments forever an agent with belief \( x \) deciding whether to be a member of the organization based on her flow payoffs is indifferent if \( s = x\frac{b}{\alpha + (1 - \alpha)(1 - y_t)} \). For the marginal member \( x = p_t(y_t) \), so that \( s = p_t(y_t)\frac{b}{\alpha + (1 - \alpha)(1 - y_t)} \). Because \( p_t(y_t) = \frac{y_t e^{-bt}}{y_t e^{-bt} + 1 - y_t} \), this is equivalent to

\[ s(1 - e^{-bt})(1 - \alpha)(y_t)^2 + y_t(s(e^{-bt} + \alpha - 2) - e^{-bt}b) + s = 0. \]

This equation has solutions \( y_t^- = \frac{e^{-bt}b - s(e^{-bt} + \alpha - 2) - \sqrt{(e^{-bt}b - s(e^{-bt} + \alpha - 2))^2 - 4s^2(1 - e^{-bt})(1 - \alpha)}}{2s(1 - e^{-bt})(1 - \alpha)} \)

and \( y_t^+ = \frac{e^{-bt}b - s(e^{-bt} + \alpha - 2) + \sqrt{(e^{-bt}b - s(e^{-bt} + \alpha - 2))^2 - 4s^2(1 - e^{-bt})(1 - \alpha)}}{2s(1 - e^{-bt})(1 - \alpha)} \).

Note that \( y_t^+ > 1 \). This is because \( y_t^+ > 1 \) is implied by \( \frac{e^{-bt}b - s(e^{-bt} + \alpha - 2)}{2s(1 - e^{-bt})(1 - \alpha)} > 1 \), which is equivalent to \( e^{-bt}(b + s - 2as) > -\alpha s \). This simplifies to \( b > s(2\alpha - 1) \), which is satisfied because \( b > s \) and \( 2\alpha - 1 \leq 1 \). Then, because \( y_t \leq 1 \), we cannot have \( y_t = y_t^+ \).

\( y_t^- \leq 1 \) is equivalent to \( \sqrt{(e^{-bt}b - s(e^{-bt} + \alpha - 2))^2 - 4s^2(1 - e^{-bt})(1 - \alpha)} \geq \alpha e^{-bt}(s + b) - s(2e^{-bt} - 1) \). This simplifies to \( b \geq \alpha s \), which is satisfied. Then \( y_t = y_t^- \) is the valid solution. Thus \( p_t(y_t) = \frac{e^{-bt}(e^{-bt}b - s(e^{-bt} + \alpha - 2) - \sqrt{(e^{-bt}b - s(e^{-bt} + \alpha - 2))^2 - 4s^2(1 - e^{-bt})(1 - \alpha)})}{(1 - e^{-bt})(2s(1 - \alpha) - (e^{-bt}b - s(e^{-bt} + \alpha - 2) - \sqrt{(e^{-bt}b - s(e^{-bt} + \alpha - 2))^2 - 4s^2(1 - e^{-bt})(1 - \alpha)})} \).

Let \( Q_1 = e^{-bt}b - s(e^{-bt} + \alpha - 2) \) and let \( Q_2 = (e^{-bt}b - s(e^{-bt} + \alpha - 2))^2 - 4s^2(1 - e^{-bt})(1 - \alpha) \). Then \( y_t = \frac{Q_1 - \sqrt{Q_2}}{2s(1 - \alpha - (Q_1 - \sqrt{Q_2}))} \) and \( p_t(y_t) = \frac{e^{-bt}(Q_1 - \sqrt{Q_2})}{(1 - e^{-bt})(2s(1 - \alpha) - (Q_1 - \sqrt{Q_2}))} \).

Note that \( \frac{\partial}{\partial e^{-bt}} Q_2 = 2(e^{-bt}b - s(e^{-bt} + \alpha - 2))(b - s) + 4s^2(1 - \alpha) \), \( \lim_{e^{-bt} \to 0} \frac{\partial}{\partial e^{-bt}} Q_2 = 2s(2b - \alpha b - \alpha s) \), \( \lim_{e^{-bt} \to 0} \sqrt{Q_2} = \alpha s \) and \( \lim_{e^{-bt} \to 0}(2s(1 - \alpha) - (Q_1 - \sqrt{Q_2})) = 0 \). Then L’Hospital’s rule applies.
We have
\[
\frac{\partial}{\partial(e^{-bt})}((1 - e^{-bt})(2s(1 - \alpha) - (Q_1 - \sqrt{Q_2}))) = -(2s(1 - \alpha) - (Q_1 - \sqrt{Q_2})) - \\
(1 - e^{-bt}) \left( b - s - \frac{1}{2} \frac{1}{\sqrt{Q_2}} (2(e^{-bt}b - s(e^{-bt} + \alpha - 2))(b - s) + 4s^2(1 - \alpha)) \right)
\]

Moreover,
\[
\lim_{e^{-bt} \to 0} \frac{\partial}{\partial(e^{-bt})}((1 - e^{-bt})(2s(1 - \alpha) - (Q_1 - \sqrt{Q_2}))) = \\
\frac{1}{\alpha} (2b - b\alpha - s\alpha) - (b - s)
\]

We also have \(\frac{\partial}{\partial(e^{-bt})}(e^{-bt}(Q_1 - \sqrt{Q_2})) = Q_1 - \sqrt{Q_2} + e^{-bt}(b - s - \frac{1}{2} \frac{1}{\sqrt{Q_2}} (2(e^{-bt}b - s(e^{-bt} + \alpha - 2))(b - s) + 4s^2(1 - \alpha)))\). Moreover, \(\lim_{e^{-bt} \to 0} \frac{\partial}{\partial(e^{-bt})}(e^{-bt}(Q_1 - \sqrt{Q_2})) = 2s(1 - \alpha)\).

Then, by L'Hospital's rule, \(\lim_{t \to \infty} p_t(y_t) = \lim_{e^{-bt} \to 0} p_t(y_t) = \frac{2s(1 - \alpha)}{\frac{1}{\alpha} (2b - b\alpha - s\alpha) - (b - s)} = \frac{s\alpha}{b}\).

Because \(m_t = \frac{1}{2}(1 + y_t)\), we have \(m_t = \frac{e^{-bt}(b - s(3 - 2\alpha)) + s(4 - 3\alpha) - \sqrt{Q_2}}{4s(1 - e^{-bt})(1 - \alpha)}\). Then \(p_t(m_t) = \frac{e^{-bt}(b - s(3 - 2\alpha)) + s(4 - 3\alpha) - \sqrt{Q_2}}{(1 - e^{-bt})(4s(1 - \alpha) - (e^{-bt}(b - s(3 - 2\alpha)) + s(4 - 3\alpha) - \sqrt{Q_2}))}\). Observe that \(\lim_{e^{-bt} \to 0}(1 - e^{-bt})(4s(1 - \alpha) - (e^{-bt}(b - s(3 - 2\alpha)) + s(4 - 3\alpha) - \sqrt{Q_2})) = 0\). Then L'Hospital's rule applies.

We have
\[
\frac{\partial}{\partial(e^{-bt})} \left( (1 - e^{-bt}) \left( 4s(1 - \alpha) - (e^{-bt}(b - s(3 - 2\alpha)) + s(4 - 3\alpha) - \sqrt{Q_2}) \right) \right) = \\
-(4s(1 - \alpha) - (e^{-bt}(b - s(3 - 2\alpha)) + s(4 - 3\alpha) - \sqrt{Q_2})) - \\
(1 - e^{-bt}) \left( b - s(3 - 2\alpha) - \frac{1}{2} \frac{1}{\sqrt{Q_2}} \frac{\partial Q_2}{\partial(e^{-bt})} \right)
\]
Moreover,

\[
\lim_{e^{-bt}\to0} \frac{\partial}{\partial (e^{-bt})} \left( (1 - e^{-bt}) \left( 4s(1 - \alpha) - (e^{-bt}(b - s(3 - 2\alpha)) + s(4 - 3\alpha) - \sqrt{Q_2}) \right) \right) = \\
-(4s(1 - \alpha) - s(4 - 3\alpha) + s\alpha) - (b - s(3 - 2\alpha) - \frac{1}{2s\alpha}2s(2b - \alpha b - \alpha s)) = \\
\frac{1}{\alpha}(2b - \alpha b - \alpha s) - b + 3s - 2s\alpha
\]

We also have

\[
\lim_{e^{-bt}\to0} \frac{\partial}{\partial (e^{-bt})} (e^{-bt}(b - s(3 - 2\alpha)) + s(4 - 3\alpha) - \sqrt{Q_2}) = e^{-bt}(b - s(3 - 2\alpha)) + s(4 - 3\alpha) - \sqrt{Q_2} + e^{-bt} \frac{\partial}{\partial e^{-bt}} (e^{-bt}(b - s(3 - 2\alpha)) + s(4 - 3\alpha) - \sqrt{Q_2}).
\]

Moreover, \(\lim_{e^{-bt}\to0} \frac{\partial}{\partial (e^{-bt})} (e^{-bt}(b - s(3 - 2\alpha)) + s(4 - 3\alpha) - \sqrt{Q_2}) = 4s(1 - \alpha)\).

Then, by L’Hospital’s rule, \(\lim_{t\to\infty} p_t(m_t) = \frac{\frac{4s(1-\alpha)}{\alpha(2b-\alpha b-\alpha s)-b+3s-2s\alpha}}{\frac{4s(1-\alpha)}{2(2b+\alpha)\alpha(1-\alpha)}} = \frac{2s\alpha}{b+\alpha}\).

**Lemma 24.** Suppose that organization is subject to a congestion effect. Then in the equilibrium in which the organization experiments forever we have \(\inf_{t\in[0,\infty)} p_t(m_t) = \frac{2s\alpha}{b+\alpha}\).

**Proof of lemma 24.**

We will show that \(p_t(y_t) \geq \alpha_x^\sharp\) for all \(t\). Because \(\lim_{t\to\infty} p_t(y_t) = \alpha_x^\sharp\) and \(\lim_{t\to\infty} p_t(m_t) = \frac{2s\alpha}{b+\alpha}\) by lemma 23, this will prove the claim in the lemma.

Lemma 23 implies that this is equivalent to \(((1 - e^{-bt})\alpha s + e^{-bt}b)(\alpha e^{-bt}b - s(e^{-bt} + \alpha s)) - ((1 - e^{-bt})\alpha s + e^{-bt}b)\) \(\leq\) \(\alpha^2 s^2(1 - e^{-bt})(1 - \alpha)\). This is equivalent to \(((1 - e^{-bt})\alpha s + be^{-bt})(\alpha s - be^{-bt}) \leq \alpha^2 s^2(1 - e^{-bt})\). This simplifies to \(\alpha s \leq b\), which is satisfied.

**Lemma 25.** Suppose that organization is subject to a congestion effect. Then in the equilibrium in which the organization experiments forever

\[
V(x) \geq V_1(x) = xb \frac{1}{\gamma}
\]

**Proof of lemma 25.**

Observe that the agent’s value is greater than the payoff \(V_0(x) = x \int_0^\infty \frac{b}{\alpha + (1-\alpha)(1-y_t)} e^{-\gamma \tau} d\tau\) obtained when the agent never leaves the organization. Because \(\frac{b}{\alpha + (1-\alpha)(1-y_t)} \geq b\) for all \(\alpha \in [0,1]\) and for all \(y_t \in [0,1]\), we have \(V_0(x) \geq V_1(x)\), where \(V_1(x) = \frac{1}{\gamma}\).
\[ x \int_0^\infty b e^{-\gamma \tau} d\tau = xb^{\frac{1}{\gamma}}. \]

**Lemma 26.** If \( V_1(x) = xb^{\frac{1}{\gamma}} \) and \( \alpha \in (\frac{1}{2}, 1] \), then there exist parameters such that \( V_1\left(\frac{2as}{b+as}\right) \geq \frac{r}{\gamma} \).

**Proof of lemma 26.**

Because \( V_1\left(\frac{2as}{b+as}\right) = \frac{1}{\gamma} \cdot \frac{2as}{b+as} \cdot b \), \( V_1\left(\frac{2as}{b+as}\right) \geq \frac{r}{\gamma} \) is equivalent to \( \alpha \geq \frac{rb}{s(2b-r)} \). We can write this as \( b(2s\alpha - r) > r s \alpha \). To have \( b(2s\alpha - r) > r s \alpha \) it is sufficient that we have \( \alpha > \frac{r}{2s} \) and \( b > \frac{rs\alpha}{2s\alpha - r} \). Observe that, because \( \alpha > \frac{1}{2} \), we can choose \( s \) and \( r \) such that \( r > s \) and \( \alpha > \frac{r}{2s} \).

**Lemma 27.** Suppose that organization is subject to a congestion effect and \( \alpha \in (\frac{1}{2}, 1] \). Then there exist parameters under which there is an equilibrium in which the organization experiments forever.

**Proof of lemma 27.**

To show that there exist parameters under which there is an equilibrium in which the organization experiments forever, it is sufficient to show that \( \inf_t V(p_t(m_t)) \geq \frac{r}{\gamma} \). By slightly modifying the proof of lemma 12, we can establish that \( x \mapsto V(x) \) is increasing. Then, because \( x \mapsto V(x) \) is increasing, it is enough to show that \( V(\inf_t p_t(m_t)) \geq \frac{r}{\gamma} \). Because \( \inf_{t \in [0, \infty)} p_t(m_t) = \frac{2s\alpha}{b+as} \) by lemma 24, it is enough to show that \( V\left(\frac{2as}{b+as}\right) \geq \frac{r}{\gamma} \). Because \( V_1\left(\frac{2as}{b+as}\right) \geq \frac{r}{\gamma} \) for \( V_1(x) = xb^{\frac{1}{\gamma}} \) by lemma 26 and \( V(x) \geq V_1(x) \) by lemma 25, this is satisfied.

**References**


