# Endogenous Experimentation in Organizations*

**Germán Gieczewski**[†] **Svetlana Kosterina**[‡]

August 2018

**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 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, taking their resources with them. We show that for a wide range of parameters there is too much experimentation in equilibrium relative to the social optimum. The potential change in membership and control due to experimentation lowers the incentives to experiment. At the same time, self-selection into the organization plays a countervailing role: when experimentation is unsuccessful, only the staunch optimists most supportive of experimentation choose to remain. For some parameters this force dominates, yielding equilibria with overexperimentation. We apply the model to decision-making in cooperatives, civil rights organizations and for-profit firms.

***Keywords:*** experimentation, dynamics, median voter, endogenous population

PRELIMINARY AND INCOMPLETE.

## 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, Wolfgang Pesendorfer, Juuso Toikka, and 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. Membership of organizations 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 membership of organizations changes, so do the policies they pursue. This paper aims to understand the dynamics of experimentation in the environments that have these features – that is, in the environments where the membership of an organization is in flux, the beliefs of the members are diverse, and the decision-making 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 the individuals within and outside the organization disagree about the degree to which the risky poverty alleviation policy is promising. 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 one at each point in time. The safe policy yields a flow payoff known by everyone, while the risky policy yields a flow payoff that the agent is uncertain about. 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 they invest with the organization, they obtain 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 an organization, they remain voting members of the organization and vote on the policy the organization 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.

Our main result is that if success is perfectly informative about the quality of the technology, then there are parameters under which there is overexperimentation in equilibrium. In particular, under some parameters there is a unique equilibrium in which the organization experiments forever. 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. On the other hand, as time passes and no successes are observed, only the most optimistic members remain in the organization, and these are precisely the members who are mostly likely to want to experiment. The first force makes underexperimentation more likely, while the second force pushes the organization to overexperiment. One contribution our paper makes is showing that the second force can dominate.

Furthermore, we shed light on the conditions under which experimenting forever is an equilibrium. We provide a simple condition on the fundamentals such that if this condition is satisfied, then under any distribution of the agents' beliefs with an increasing density there is a unique equilibrium, and the organization experiments forever in this equilibrium. This implies that, in particular, if there is an equilibrium in which the organization experiments forever under the uniform distribution of the agents' priors, then there is an equilibrium in which the organization experiments forever under any distribution of the agents' priors with an increasing density. In this manner, greater optimism (in the sense of the monotone likelihood ratio property) increases the likelihood of overexperimentation.

We next consider more general experimentation technologies under which success is only imperfectly informative about the quality of the technology. In this case, we are able to provide a delay differential equation characterizing the evolution of the continuation value of an agent with a fixed belief. We have several conjectures (that we have not proven yet) about the results that obtain in this case. In particular, we conjecture that for some parameters there is an equilibrium in which an organization stops experimenting with a strictly positive probability if and only if a success is observed. A consequence of this would be 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.

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 basic model. Section 5 analyzes the set of equilibria. Section 6 deals with the case in which success is only imperfectly informative about the quality of the technology.

## 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 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 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

The paper is related to two broad strands of literature: the literature on strategic experimentation with multiple agents (Keller, Rady, and Cripps 2005, Keller and Rady 2010, Keller and Rady 2015, Strulovici 2010) and the literature on the dynamics of decision-making in clubs (Acemoglu, Egorov, and Sonin 2008, Bai and Lagunoff 2011, Acemoglu, Egorov, and Sonin 2012, Acemoglu, Egorov, and Sonin 2015, Gieczewski 2017).

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

2011, Acemoglu, Egorov, and Sonin 2012, Acemoglu, Egorov, and Sonin 2015) 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 *good* with probability $\rho$ and *bad* with probability $1 - \rho$. 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 $f$ defined 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 $t$.[2] We write $\alpha_t(x) = 1$ if $x \in X_t$ and $\alpha_t(x) = 0$ 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 ($\sigma_t = 0$) all members receive a guaranteed flow payoff $r$. When the risky policy is used ($\sigma_t = 1$) its payoffs depend on the state of the world. If the risky policy is good, it succeeds according to a Poisson process with rate $b$. If the risky policy is bad, it never succeeds. Each time the risky

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

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

$$U_x(\sigma,\alpha) = xE\left[\int_0^\infty e^{-\gamma t}\left(b\sigma_t\alpha_t(x) + r(1-\sigma_t)\alpha_t(x) + s(1-\alpha_t(x))\right)dt|\theta = G\right] +$$
$$+(1-x)E\left[\int_0^\infty e^{-\gamma t}\left(r(1-\sigma_t)\alpha_t(x) + s(1-\alpha_t(x))\right)dt|\theta = B\right]$$

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. Let $k_t$ denote the number of successes observed by time $t$, let $q(t)$ denote the amount of time the organization has used the risky policy up to time $t$, and let $z_x(k,\tau)$ be the posterior of an agent with prior $x$ after seeing $k$ successes during a length of time $\tau$ spent experimenting. It then follows from Bayes' rule that

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

where $L(k,\tau) = \mathbb{1}_{k=0}e^{b\tau}$.[3] Since $L(k,\tau)$ serves as a sufficient statistic for the information observed so far, we hereafter define $z_x(k,\tau) = z_x(L(k,\tau))$ and, suppressing the dependence of $L(k,t)$ on $k$ and $t$, we take $L$ to be a state variable in our model.

Recall that, at each time $t$, a subset of the population $X_t$ belongs to the organization. We assume that the median voter of this set, $m(X_t)$, chooses whether the organization should continue to experiment at that instant.[4] Since there is a continuum of agents, an agent obtains no value from her ability to vote and behaves as a policy-taker with respect to her membership decision. That is, she joins the organization when she prefers the expected flow payoff it offers to that of investing independently.

[3] Note that, in particular, $z_x(0,0) = x$.

[4] For $m(X_t)$ to be well-defined, we require $X_t$ to be Lebesgue-measurable. It can be shown that in any equilibrium $X_t$ is an interval.

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 $t \in \{0, \epsilon, 2\epsilon, \dots\}$ with $\epsilon > 0$ small. In this discrete-time game, at each time $t$ in $\{0, \epsilon, 2\epsilon, \dots\}$, first the incumbent median chooses a policy $\sigma_t$ 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 $\sigma_t$—are the incumbent members at time $t+\epsilon$. The median of this set of members, $m(X_{t+\epsilon})$, then chooses $\sigma_{t+\epsilon}$. 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.** An strategy profile is given by a collection of membership functions $\alpha(x) : [0,1] \times \mathbb{R}_+ \times \{0,1\} \to \{0,1\}$ and a policy function $\sigma : \mathbb{R}_+ \times \{0,1\} \to [0,1]$.

Given a strategy profile and a right-continuous path for the underlying stochastic process $(\tilde{L}_\tau)_\tau$, the path of play is given by a policy, information and pivotal voter path $(p_t, L_t, m_t)$, which satisfies:[5]

(a) $L_t = \tilde{L}_{q(t)}$, where $q(t) = \int_0^t p_t$ is the amount of time that the organization has experimented for.

(b) Whenever $\sigma(L, 1 - p_{t_0}) = 1 - p_{t_0}$ for all $l$ in a neighborhood of $L_{t_0}$, and $\tilde{L}_\tau$ is left-continuous at $\tau = q(t_0)$, $p_t$ is left-continuous at $t_0$.

(c) Whenever $\sigma(L, p_{t_0}) = p_{t_0}$ for all $l$ in a neighborhood of $L_{t_0}$, $p_t$ is right-continuous at $t_0$.

(d) $m_t = m(L_t, p_t)$, where $m(L, p)$ is the median agent in the set of members $X(L, p)$.

An equilibrium is a strategy profile such that:

[5] We have $p_t = 1$ when the risky policy is being used and $p_t = 0$ when the safe policy is used.

(i) $\alpha(x, L, 1) = 1$ if $z_x(L)b > s$, $\alpha(x, L, 1) = 0$ if $z_x(L)b < s$ and $\alpha(x, L, 0) = 1$ if $r > s$.

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

(iii) If $V_{m(L,p)}(L, 1) = V_{m(L,p)}(L, 0)$ but $V_{m(L,p)}(L, p', \epsilon) - V_{m(L,p)}(L, 1 - p', \epsilon) > 0$ for all $\epsilon > 0$ small enough, then $\sigma(L, p) = p'$.[6]

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

Parts (b) and (c) of the definition 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 going to be in equilibrium depending on what has happened so far in the game. 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

[6] $V_x(L, p, \epsilon)$ is $x$'s continuation utility starting from some time $t_0$ with information $L_{t_0} = L$, when $p$ 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 4 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 $z_x(L)b \geq s$. It is clear that, for any $L \geq 1$ (i.e., if no successes have been observed), $z_x(L)$ 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, $m([0,1])$. 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 $m(X_t)$, the median voter of the organization. Thus, if $m(X_t)$ preferred to switch to the safe policy, so does $m([0,1])$. Because no further learning happens when the safe policy is used, a switch to the safe policy is permanent.

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

This implies that 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

[7]If $t^* = +\infty$, the risky policy is used forever.

switches to the safe policy at $t_n$.[8]

Proposition 1 states our first main result. The result provides a simple condition on the parameters sufficient for overexperimentation to arise in equilibrium. More specifically, if this condition is satisfied, then the organization uses the risky policy forever regardless of its results.

**Proposition 1.** *Suppose that $f$ is non-decreasing. Let $V(x)$ denote the continuation utility of an agent with posterior belief $x$ at time $t$ provided that she expects experimentation to continue for all $s \geq t$.*

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

*Moreover, the value function satisfies*

$$\gamma V\left(\frac{2s}{b+s}\right) = \frac{2bs}{b+s} + \left(\frac{1}{2}\right)^{\frac{\gamma}{b}} \frac{s(b-s)}{b+s}\frac{b}{\gamma+b}.$$

The sufficient condition for obtaining experimentation forever provided in the Proposition is not difficult to satisfy. It is is more likely to hold when $b$ is high relative to $r$ and $s$, that is, when the returns from good risky technology are high, and when $\gamma$ is low, that is, when the agents are sufficiently patient. For example, if $r = 3$ and $s = 2$, when $b \geq 6$ it holds regardless of $\gamma$, when $\frac{10}{3} < b < 6$ it holds for low enough $\gamma$, and when $b \leq \frac{10}{3}$ it cannot hold.

Proposition 1 implies that over-experimentation is a possible outcome in our model. Since agents with different priors prefer different levels of experimentation, the notion of over-experimentation is not straightforward in our setting. Consider an alternative model in which an agent with a fixed prior $x$ controls the policy at all times. It can be shown that, whenever $0 < x < 1$, the agent would experiment until some finite time $t^*$ such that her posterior belief at time $t^*$ equals $\frac{r}{b}\frac{1}{1+\frac{b-r}{\gamma}}$. Thus an equilibrium path on which unsuccessful experimentation continues forever constitutes over-experimentation from the point of view of all agents except those with prior belief equal to 1.

[8] $t_0$ is the only stopping time on the equilibrium path.

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 $m(X_t)$ 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, $m(X_t)$ knows that, if she chooses to continue experimenting, the organization will stop at a time chosen by some other agent, which $m(X_t)$ likely considers suboptimal. This force encourages $m(X_t)$ 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 long as experimentation continues and no successes are observed, all agents become more pessimistic about the risky policy, and the marginal members with lower priors quit when they perceive their flow payoff from the risky policy to be lower than that of the outside option. As a result, the set of members at time $t$ is an interval from $y_t$ to 1, with $y_t$ increasing in $t$. This implies that $m_t$ is also increasing in $t$. Thus the more pessimistic an observer with a fixed prior is 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 converges not to zero but rather to $\frac{2s}{b+s}$. It is for this reason that the median voter may choose to continue experimenting when no successes have been observed for an arbitrarily long time.

The above discussion helps us explain why we require $f$ to be non-decreasing in Proposition 1. After many failures, only an agent with a prior belief very close to 1 would choose to continue experimenting. The requirement that $f$ be non-decreasing guarantees that there are enough optimists in the right tail of the distribution to ensure that the median is extreme enough, and hence optimistic enough, to do this.

We next show that Proposition 1 can be generalized. Indeed, although the validity of the exact parameter conditions imposed by Proposition 1 hinge on $f$ being non-decreasing, analogous conditions can be provided more generally as a function of how quickly $f(x)$ decreases near $x = 1$.

**Proposition 2.** *Given* $\alpha > 0$, *let* $f_\alpha(x)$ *denote a density with support* $[0, 1]$ *such that* $f_\alpha(x) = (\alpha + 1)(1 - x)^\alpha$ *for* $x \in [0, 1]$.

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

*If* $V\left(\frac{s}{\lambda b + (1-\lambda)s}\right) \geq \frac{r}{\gamma}$, *then there is a unique equilibrium. In this equilibrium, if the risky policy is used at* $t = 0$, *the organization experiments forever.*

*Moreover, the value function satisfies*

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

**Proposition 3.** *Let* $f$ *be a density with support* $[0, 1]$.

*If* $V\left(\frac{s}{b}\right) \geq \frac{r}{\gamma}$, *then there is a unique equilibrium. In this equilibrium, if the risky policy is used at* $t = 0$, *the organization experiments forever.*

*Moreover,*

$$\gamma V\left(\frac{s}{b}\right) = s + \frac{s(b-s)}{\gamma + b}.$$

The intuition for the result in Propositions 2 and 3 is as follows. Observe that, at any time, the pivotal decision-maker is the median of a set where the most pessimistic member has posterior $\frac{s}{b}$ and the most optimistic has posterior 1. The higher $\alpha$ is, the more quickly $f_\alpha(x)$ goes to zero as $x \to 1$, meaning that there are fewer optimists in the right tail of the prior distribution. As a result, the distribution of posteriors is also more right-skewed, so median voter's posterior is closer to $\frac{s}{b}$—the marginal voter's posterior—than it is when $f$ is uniform or increasing.

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 the *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 $\tau(t) = +\infty$.[9] Proposition 4 characterizes the equilibria in this setting.

**Proposition 4.** *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 4 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 4 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.

[9]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 $V_{t'}(x)$ denote the value function of an agent with prior $x$ 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 $t' > t$, $\frac{d}{dt}V_{t'}(p_t(m_t)) \neq 0$*

*whenever* $V_{t'}(p_t(m_t)) = \frac{r}{\gamma}$.

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 $t'$, the function $t \mapsto V_{t'}(p_t(m_t))$ does not have its derivative equal zero at any point where it crosses the threshold $\frac{r}{\gamma}$. (In particular, this implies that $V_{t'}(p_t(m_t))$ can only equal $\frac{r}{\gamma}$ for a finite number of times $t$.)

Proposition 5 characterizes the equilibrium in our model that features both bad news and endogenous membership.

**Proposition 5.** *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* $V^b\left(\frac{2s}{b+s}\right) \geq \frac{r}{\gamma}$, *the organization experiments forever unless a failure is observed.*

Proposition 5 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 5. 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 $m_{t'}$ for $t' < t$. 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, i.e., 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, using $V^b(x)$ to denote the value function of an agent with prior $x$ provided that the organization experiments forever in a model of bad news, if $V^b(p_t(m_t)) > \frac{r}{\gamma}$ for all $t$ then the organization must experiment forever. The last part of Proposition 5 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 as 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 5 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 $b > r > s > c > 0$.

Because we now need to study continuations of a game where varying numbers of successes have occurred at different times, our description of equilibrium will make explicit use of our notation for information gathered up to time $t$, $L(k,t)$. 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. We define

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

Then, letting $z_x(k,t)$ denote 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,[10] it is still the case that

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

Henceforth, we suppress the dependence of $L(k,t)$ on $k$ and $t$ and use $L$ to denote our sufficient statistic. In addition, we denote $l = \ln L$, which will sometimes be easier to work with, and $z_x(l) = z_x(l(k,t))$.

We also restrict our attention to equilibrium strategies that condition only on $L$. A (pure strategy) equilibrium can then 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 $\mathcal{L}$.

We let $V_x^i(l)$ denote the value function of an agent with prior $x$ given that the state is $l = \ln L$ in a model of imperfectly informative news provided that on the equilibrium path experimentation continues forever. We also let $V^i(x) = V_x^i(0)$ 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 6.** *If $f$ is non-decreasing and $V^i\left(\frac{2(s-c)}{(b-c)+(s-c)}\right) \geq \frac{r}{\gamma}$, there is a unique equilibrium in which the organization experiments forever. If $V^i\left(\frac{2(s-c)}{(b-c)+(s-c)}\right) < \frac{r}{\gamma}$, there is no equilibrium in which the organization experiments forever.*

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

[10] The agent's posterior only depends on $k$ and $t$, and not on the timing of the successes.

*that* $V^i\left(\frac{2(s-c)}{(b-c)+(s-c)}\right) \geq \frac{r}{\gamma}$.

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

**Proposition 7.** *There are values of* $b$, $r$, $s$, $c$ *and* $f$ *for which there exists an equilibrium such that, if* $l = l^*$, *then the organization stops experimenting with probability* $\epsilon > 0$, *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, $l \mapsto z_{m(l)}(l)$ is decreasing in $l$ for $l < l^*$ but increasing in $l$ for $l > l^*$. It may then be that $z_{m(l)}(l)$ has a global minimum at $l = l^*$, i.e., 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 decisionmaker.

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 7 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.*

# Appendix

**Lemma 1.** *Suppose that the initial distribution of priors is uniform. 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*

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

**Proof of lemma 1.**

The posterior belief of a marginal member of the organization is given by $p_t(y_t) = \frac{y_t e^{-bt}}{y_t e^{-bt} + 1 - y_t}$. Using the fact that $p_t(y_t) = \frac{s}{b}$, we set $\frac{y_t e^{-bt}}{y_t e^{-bt} + 1 - y_t} = \frac{s}{b}$. Solving for $y_t$, we obtain

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

Substituting the above formula for $y_t$ into $m_t = \frac{1+y_t}{2}$, we obtain

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

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

$$\begin{aligned} p_t(m_t) = \quad & \frac{\frac{2s+(b-s)e^{-bt}}{s+(b-s)e^{-bt}} e^{-bt}}{\frac{2s+(b-s)e^{-bt}}{s+(b-s)e^{-bt}} e^{-bt} + 2 - \frac{2s+(b-s)e^{-bt}}{s+(b-s)e^{-bt}}} = \\ & \frac{(2s + (b-s)e^{-bt})e^{-bt}}{(2s + (b-s)e^{-bt})e^{-bt} + 2(s + (b-s)e^{-bt}) - 2s - (b-s)e^{-bt}} = \\ & \frac{(2s + (b-s)e^{-bt})e^{-bt}}{(2s + (b-s)e^{-bt})e^{-bt} + (b-s)e^{-bt}} = \\ & \frac{2s + (b-s)e^{-bt}}{2s + (b-s)e^{-bt} + b - s} = \\ & \frac{2s + (b-s)e^{-bt}}{b + s + (b-s)e^{-bt}} \end{aligned}$$

■

We now provide a formula for $V(x)$.

**Lemma 2.**

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

**Proof of lemma 2.**

Let $t(x)$ denote the time it will take for an agent's posterior belief to go from $x$ to $\frac{s}{b}$, at which time she would leave the organization. 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

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

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 obtaining 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

$$\begin{aligned}
V(x) &= x\int_0^{t(x)} be^{-\gamma\tau}d\tau + \int_{t(x)}^{\infty} se^{-\gamma\tau}d\tau + \int_{t(x)}^{\infty} P_\tau(b-s)e^{-\gamma\tau}d\tau \\
&= xb\frac{1-e^{-\gamma t(x)}}{\gamma} + e^{-\gamma t(x)}\frac{s}{\gamma} + x(b-s)\int_{t(x)}^{\infty}\left(e^{-\gamma\tau} - e^{-(\gamma+b)\tau}\right)d\tau \\
&= xb\frac{1-e^{-\gamma t(x)}}{\gamma} + e^{-\gamma t(x)}\frac{s}{\gamma} + x(b-s)\left(\frac{e^{-\gamma t(x)}}{\gamma} - \frac{e^{-(\gamma+b)t(x)}}{\gamma+b}\right) \\
&= xb\frac{1}{\gamma} + (1-x)e^{-\gamma t(x)}\frac{s}{\gamma} - x(b-s)\frac{e^{-(\gamma+b)t(x)}}{\gamma+b}
\end{aligned}$$

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 $\int_t^\infty e^{-\gamma\tau}d\tau = \frac{e^{-\gamma t}}{\gamma}$ and $\int_t^\infty e^{-(\gamma+b)\tau}d\tau = \frac{e^{-(\gamma+b)t}}{\gamma+b}$. ■

**Lemma 3.** *Let $t_y(x)$ denote he time it takes for an agent's posterior belief to go from $x$ to $y$. Then*

$$t_y(x) = -\frac{\ln\left(\frac{y}{1-y}\frac{1-x}{x}\right)}{b} \qquad t(x) = -\frac{\ln\left(\frac{s(1-x)}{(b-s)x}\right)}{b}$$

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

**Proof of lemma 3.**

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

Recall that $t(x)$ is the time it takes for an agent's posterior belief to go from $x$ to $\frac{s}{b}$. Substituting $y = \frac{s}{b}$ in, we obtain $t(x) = -\frac{\ln\left(\frac{s(1-x)}{(b-s)x}\right)}{b}$. Substituting $x = \frac{2s}{b+s}$ into $e^{-bt(x)} = \frac{s(1-x)}{(b-s)x}$ and simplifying, we obtain $e^{-bt} = \frac{1}{2}$. ■

**Lemma 4.** $x \mapsto V(x)$ *is increasing.*

**Proof of lemma 4.**

By lemma 3, we have

$$t(x) = -\frac{\ln\left(\frac{s(1-x)}{(b-s)x}\right)}{b}$$

Thus

$$t'(x) = -\frac{1}{b}\frac{(b-s)x}{s(1-x)}\frac{s}{b-s}\left(-\frac{1}{x^2}\right) = \frac{1}{b}\frac{1}{x(1-x)} > 0$$

Then we have

$$\begin{aligned}\gamma V'(x) &= b - e^{-\gamma t(x)}s - \frac{\gamma}{\gamma+b}(b-s)e^{-(\gamma+b)t(x)} \\ &+t'(x)e^{-\gamma t(x)}s + t'(x)\frac{\gamma}{\gamma+b}(b-s)(\gamma+b)e^{-(\gamma+b)t(x)} \\ &\geq b - s - \frac{\gamma}{\gamma+b}(b-s) = \frac{(b-s)b}{\gamma+b} > 0\end{aligned}$$

where the first inequality follows because $e^{-\gamma t(x)} \in [0,1]$, $e^{-(\gamma+b)t(x)} \in [0,1]$ and $t'(x) > 0$. ■

**Lemma 5.** *Suppose that the initial distribution of priors is uniform. Then, whenever* $V\left(\frac{2s}{b+s}\right) \geq \frac{r}{\gamma}$ *is satisfied, there exists an equilibrium in which the organization experiments forever.*

**Proof of lemma 5.**

Recall that the continuation utility of an agent with belief $x$ conditional on experimentation continuing forever is $V(x)$. Observe that the payoff to using the safe technology forever is $\frac{r}{\gamma}$. We then require that $V(p_t(m_t)) \geq \frac{r}{\gamma}$ for all $t$. That is, we need to check that, whenever $V\left(\frac{2s}{b+s}\right) \geq \frac{r}{\gamma}$, for all $t$ the median agent $m_t$ prefers to continue experimenting at $t$ (which leads to experimentation forever) instead of stopping.

Because the fact that the initial distribution of priors is uniform implies that $t \mapsto p_t(m_t)$ is decreasing and lemma 4 implies that $x \mapsto V(x)$ is increasing, in order to have $V(p_t(m_t)) \geq \frac{r}{\gamma}$ for all $t$, it is sufficient to have $\lim_{t\to\infty} V(p_t(m_t)) \geq \frac{r}{\gamma}$. Because lemma 1 implies that $p_t(m_t) = \frac{2s+(b-s)e^{-bt}}{b+s+(b-s)e^{-bt}}$, we have $\lim_{t\to\infty} p_t(m_t) = \lim_{t\to\infty} \frac{2s+(b-s)e^{-bt}}{b+s+(b-s)e^{-bt}} = \frac{2s}{b+s}$. Then $\lim_{t\to\infty} V(p_t(m_t)) \geq \frac{r}{\gamma}$ is equivalent to $V\left(\frac{2s}{b+s}\right) \geq \frac{r}{\gamma}$. ■

**Lemma 6.** *There exist parameters such that* $V\left(\frac{2s}{b+s}\right) \geq \frac{r}{\gamma}$.

**Proof of lemma 6.**

Lemma 3 implies that if $x = \frac{2s}{b+s}$, then $e^{-bt(x)} = \frac{1}{2}$. Observe that we have $\left(\frac{1}{2}\right)^{\frac{\gamma}{b}} = \left(e^{-bt(x)}\right)^{\frac{\gamma}{b}} = e^{-\gamma t(x)}$, which implies that $e^{-\gamma t(x)} = \left(\frac{1}{2}\right)^{\frac{\gamma}{b}}$.

Substituting in $e^{-bt(x)} = \frac{1}{2}$ and $e^{-\gamma t(x)} = \left(\frac{1}{2}\right)^{\frac{\gamma}{b}}$ into the formula for $V(x)$ from lemma 2, we obtain

$$\begin{aligned} V(x) &= xb\frac{1}{\gamma} + (1-x)\left(\frac{1}{2}\right)^{\frac{\gamma}{b}}\frac{s}{\gamma} - x(b-s)\left(\frac{1}{2}\right)^{\frac{\gamma}{b}}\left(\frac{1}{2}\right)\frac{1}{\gamma+b} \\ &= xb\frac{1}{\gamma} + \left(\frac{1}{2}\right)^{\frac{\gamma}{b}}\left[(1-x)\frac{s}{\gamma} - x(b-s)\left(\frac{1}{2}\right)\frac{1}{\gamma+b}\right] \end{aligned}$$

Then we have

$$\begin{aligned} V\left(\frac{2s}{b+s}\right) &= \frac{2bs}{b+s}\frac{1}{\gamma} + \left(\frac{1}{2}\right)^{\frac{\gamma}{b}}\left[\frac{s(b-s)}{b+s}\frac{1}{\gamma} - \frac{s(b-s)}{b+s}\frac{1}{\gamma+b}\right] \\ &= \frac{2bs}{b+s}\frac{1}{\gamma} + \left(\frac{1}{2}\right)^{\frac{\gamma}{b}}\frac{s(b-s)}{b+s}\left[\frac{1}{\gamma} - \frac{1}{\gamma+b}\right] \end{aligned}$$

Multiplying both sides by $\gamma$, we obtain

$$\begin{aligned} \gamma V\left(\frac{2s}{b+s}\right) &= \frac{2bs}{b+s} + \left(\frac{1}{2}\right)^{\frac{\gamma}{b}}\frac{s(b-s)}{b+s}\left[1 - \frac{\gamma}{\gamma+b}\right] \\ &= \frac{2bs}{b+s} + \left(\frac{1}{2}\right)^{\frac{\gamma}{b}}\frac{s(b-s)}{b+s}\frac{b}{\gamma+b} \end{aligned}$$

It is easy to show that there exist parameters such that $\frac{2bs}{b+s}+\left(\frac{1}{2}\right)^{\frac{\gamma}{b}}\frac{s(b-s)}{b+s}\frac{b}{\gamma+b}\geq r$ 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 $b\geq\frac{sr}{2s-r}$. ■

**Lemma 7.** *Let $W_T(x)$ denote the value function of an agent with belief $x$ in an equilibrium in which the organization stops the experimentation after time $T$ passes. For all interior beliefs $x$, the following is true: if $V(x)>\frac{r}{\gamma}$, then $W_T(x)>\frac{r}{\gamma}$ for all $T$.*

**Proof of lemma 7.**

We will use a discrete-time approximation. We let $\Delta>0$ denote the length of a time period. We use $W_T^{\Delta}(x)$ to denote the value function of an agent with belief $x$ in an equilibrium in which the organization stops the experimentation after $T$ time periods pass provided that the length of time period is $\Delta$. Note that $W_0^{\Delta}(x)=\frac{r}{\gamma}$.

We can write $W_T^{\Delta}$ recursively as follows:

$$W_T^{\Delta}(x)=x\left(1-e^{-b\Delta}\right)+e^{-\gamma\Delta}\left(x\left(1-e^{-b\Delta}\right)\frac{b}{\gamma}+\left(1-x\left(1-e^{-b\Delta}\right)\right)W_{T-1}^{\Delta}(x_{T-1})\right)$$

where $x_{T-1}$ is the belief of an agent who had the prior $x$ when $T$ periods were left until the experimentation stopped (and $x_{T-1}$ is the belief this agent has when $T-1$ periods are left until the organization stops experimenting).

Similarly, we can write $V^{\Delta}$ recursively as follows:

$$V^{\Delta}(x)=x\left(1-e^{-b\Delta}\right)+e^{-\gamma\Delta}\left(x\left(1-e^{-b\Delta}\right)\frac{b}{\gamma}+\left(1-x\left(1-e^{-b\Delta}\right)\right)V^{\Delta}(x_{T-1})\right)$$

Then, by induction on the number of periods left until the organization stops experimenting, we have

$$V^{\Delta}(x)=(1-Q)Y+QV^{\Delta}(x_0)\qquad W_T^{\Delta}(x)=(1-Q)Y+QW_0^{\Delta}(x_0)$$

for some $Q\in(0,1)$ and $Y>0$.

Observe that

$$V^{\Delta}(x)=(1-Q)Y+QV^{\Delta}(x_0)\leq(1-Q)Y+QV^{\Delta}(x)$$

where the inequality follows from the fact that $V^\Delta(x) \geq V^\Delta(x_0)$ (which holds because $x > x_0$ and $V^\Delta(\cdot)$ is increasing by lemma 4).

This implies that $Y \geq V^\Delta(x)$.

Because $V^\Delta(x) > \frac{r}{\gamma}$, it follows that that

$$Y > \frac{r}{\gamma} \tag{1}$$

Then

$$W_T^\Delta(x) = (1-Q)Y + QW_0^\Delta(x_0) > (1-Q)\frac{r}{\gamma} + Q\frac{r}{\gamma} = \frac{r}{\gamma}$$

where the inequality follows from the fact that $Y > \frac{r}{\gamma}$ by (1) and the fact that $W_0^\Delta(x_0) = \frac{r}{\gamma}$.

Thus we have established that for all $\Delta > 0$, whenever $V^\Delta(x) > \frac{r}{\gamma}$, we have that $W_T^\Delta(x) > \frac{r}{\gamma}$. Because $\lim_{\Delta\to 0} V^\Delta(x) = V(x)$ and $\lim_{\Delta\to 0} W_T^\Delta(x) = W_T(x)$, it follows that whenever $V(x) > \frac{r}{\gamma}$, we have that $W_T(x) > \frac{r}{\gamma}$. ■

**Lemma 8.** *For $L = \left(\frac{c}{b}\right)^k e^{(b-c)t}$, let $m(L)$ and $\tilde{m}(L)$ denote the median voters when the state variable is $L$ under the uniform density and a non-decreasing density $f$ respectively. Suppose that $y(L) \to 1$ as $L \to \infty$. Then $\frac{1-\tilde{m}(L)}{1-m(L)} \to 1$ as $L \to \infty$.*

**Proof of lemma 8.**

Given a state variable $L$ and its corresponding marginal member $y(L)$, let $f_{0L} = f(y(L))$ and $f_1 = f(1)$. Consider $\tilde{m}(L)$. Given that $f$ is increasing on $[y(L), 1]$, 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 9.** *Consider the model of imperfectly informative news. 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, and let $z_m(L)$ and $z_{\tilde{m}}(L)$ denote the posteriors of these voters. Then* $\lim_{L\to\infty} z_m(L) = \lim_{L\to\infty} z_{\tilde{m}}(L) = \frac{2(s-c)}{(b-c)+(s-c)}$.

**Proof of lemma 9.**

Recall that $m(L) = \frac{2L(s-c)+b-s}{2L(s-c)+2(b-s)}$, so that $\lim_{L\to\infty} m(L) = 1$ and $1 - m(L) = \frac{b-s}{2L(s-c)+2(b-s)}$, so that $(1-m(L))L = \frac{L(b-s)}{2L(s-c)+2(b-s)}$ and $\lim_{L\to\infty}(1-m(L))L = \frac{1}{2}\frac{b-s}{s-c}$.

Note that $\lim_{L\to\infty} \tilde{m}(L) = 1$. The fact that, by lemma 8, we have $\lim_{L\to\infty} \frac{1-\tilde{m}(L)}{1-m(L)} = 1$ implies that $\lim_{L\to\infty}(1-\tilde{m}(L))L = \frac{1}{2}\frac{b-s}{s-c}$.

Then, because

$$\lim_{L\to\infty} z_{\tilde{m}}(L) = \lim_{L\to\infty} \frac{\tilde{m}(L)}{\tilde{m}(L) + (1-\tilde{m}(L))L} = \frac{1}{1+\frac{1}{2}\frac{b-s}{s-c}} = \frac{2(s-c)}{(b-c)+(s-c)}$$

we have $\lim_{L\to\infty} z_{\tilde{m}}(L) = \lim_{L\to\infty} z_m(L) = \frac{2(s-c)}{(b-c)+(s-c)}$, as required. ■

**Proof of Proposition 1.**

We first write the proof for when $f$ is uniform, and then consider a general non-decreasing $f$.

Suppose $f$ is uniform. Lemma 5 guarantees that, whenever $V\left(\frac{2s}{b+s}\right) \geq \frac{r}{\gamma}$ is satisfied, there exists an equilibrium in which the organization experiments forever. We will show that this equilibrium is unique. It is sufficient to show that there does not exist an equilibrium in which experimentation continues with probability one until some time $T < \infty$ and stops at time $T$ with probability one.

Let $W_T(x)$ denote the value function of an agent with belief $x$ in such an equilibrium provided that the organization stops the experimentation after time $T$ passes. Suppose that $V\left(\frac{2s}{b+s}\right) \geq \frac{r}{\gamma}$. Then, because $V(\cdot)$ is strictly increasing, we have $V(x) > \frac{r}{\gamma}$ for all $x > \frac{2s}{b+s}$. Lemma 7 shows that for all interior beliefs $x$, if $V(x) > \frac{r}{\gamma}$, then $W_T(x) > \frac{r}{\gamma}$ for all $T$. Then $W_T(x) > \frac{r}{\gamma}$ for all $T$ and for all $x > \frac{2s}{b+s}$. Thus an agent with posterior belief $x > \frac{2s}{b+s}$ strictly prefers to experiment, contradicting our hypothesis that the organization stops the experimentation after time $T$ passes.

Now consider a non-decreasing $f$.

Let $m_t$ denote the prior of the median at time $t$ under the uniform distribution, and let $\tilde{m}_t$ denote the prior of the median at time $t$ under the distribution $F$.

Suppose that under the uniform distribution there is an equilibrium where the organization experiments forever, that is, that $V\left(\frac{2s}{b+s}\right) \geq \frac{r}{\gamma}$.

We will argue that $p_t(\tilde{m}_t) \geq p_t(m_t)$ for all $t$. Because $m \mapsto p_t(m)$ is strictly increasing, it is sufficient to show that $\tilde{m}_t \geq m_t$ for all $t$.

Since $m_t = \frac{1}{2}(1 + y_t)$ and $\tilde{m}_t = F^{-1}\left(\frac{1}{2}(1 + F(y_t))\right)$, we want to show that $F^{-1}\left(\frac{1}{2}(1 + F(y_t))\right) \geq \frac{1}{2}(1 + y_t)$ for all $t$. This is equivalent to $\frac{1}{2}(1 + F(y_t)) \geq F\left(\frac{1}{2}(1 + y_t)\right)$ for all $t$. Observe that, since $f$ is increasing, $F$ is convex. In addition, we have $F(1) = 1$ and, by Jensen's inequality, $\frac{1}{2}(F(1) + F(y_t)) \geq F\left(\frac{1}{2}(1 + y_t)\right)$. Hence $\tilde{m}_t \geq m_t$ for all $t$, as desired, and $p_t(\tilde{m}_t) \geq p_t(m_t) \geq \frac{2s}{b+s}$ for all $t$. There is then a unique equilibrium,[11] involving experimentation forever, under $F$ by an argument analogous to the argument above.

Now suppose that $V\left(\frac{2s}{b+s}\right) < \frac{r}{\gamma}$. Suppose that $f$ is uniform. Because $V(\cdot)$ is continuous, we have $V(x) < \frac{r}{\gamma}$ for some $x > \frac{2s}{b+s}$ such that $x - \frac{2s}{b+s}$ is sufficiently small. Since $\frac{2s}{b+s}$ is the asymptotic posterior of the median, there exists a (possibly large) time $t$ such that, if experimentation continues without success until time $t$, then the posterior of the median at time $t$ will be $x$. As $V(x) < \frac{r}{\gamma}$, this median strictly prefers to switch to the safe policy.

Next, take $f$ to be any non-decreasing density. Let $m_t$ be the median voter at time $t$ under a uniform density, and $\tilde{m}_t$ be the median voter at time $t$ under $f$. It is sufficient to show that $\frac{2s}{b+s}$ is still the asymptotic posterior of the median, that is, that $\lim_{t\to\infty} p_t(\tilde{m}_t) = \lim_{t\to\infty} p_t(m_t) = \frac{2s}{b+s}$. The rest follows from an argument similar to the argument above.

Note that, because the baseline model is a special case of the model with imperfectly informative news with $c = 0$, lemma 8 implies that $\frac{1-\tilde{m}_t}{1-m_t} \to 1$.

Recall that $p_t(m) = \frac{me^{-bt}}{me^{-bt}+1-m}$, and check that $(1 - m_t)e^{bt} \to \frac{b-s}{2s}$ as $t \to \infty$.

[11] Observe that the part of the proof of Proposition 1 that shows that if $V\left(\frac{2s}{b+s}\right) \geq \frac{r}{\gamma}$, then there does not exist any equilibrium that does not involve experimenting forever (in particular, the part of the proof that uses lemma 7) does not rely on the distribution of priors being uniform.

Then lemma 8 implies that $(1-\tilde{m}_t)e^{bt} \to \frac{b-s}{2s}$ as well, and

$$p_t(\tilde{m}_t) = \frac{\tilde{m}_t e^{-bt}}{\tilde{m}_t e^{-bt} + 1 - \tilde{m}_t} = \frac{\tilde{m}_t}{\tilde{m}_t + (1-\tilde{m}_t)e^{bt}} \to_{t\to\infty} \frac{1}{1+\frac{b-s}{2s}} = \frac{2s}{b+s}.$$

Finally, the formula for $V\left(\frac{2s}{b+s}\right)$ follows from Lemma 6. ■

**Proof of Proposition 2.**

As in Proposition 1, we write the proof in two steps. First, for a given $\alpha$, we prove the result for the density $f_\alpha$ given by $f_\alpha(x) = \frac{(1-x)^\alpha}{\alpha+1}$ for $x \in [0,1]$. Afterwards we argue that the argument extends to other densities that MLRP-dominate $f_\alpha$.

First note that, as in Lemma 1, it is still the case that

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

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 $p_t(x)$ from Lemma 1, we obtain

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

It is clear that the above expression is decreasing in $t$ and converges to $\frac{s}{\lambda b+(1-\lambda)s}$ as $t\to\infty$. By the same argument as in Proposition 1, if $V\left(\frac{s}{\lambda b+(1-\lambda)s}\right) \geq \frac{r}{\gamma}$, then there is an equilibrium in which experimentation continues forever, and no other strategy profiles are compatible with equilibrium.

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

$$t_{\frac{s}{b}}\left(\frac{s}{\lambda b+(1-\lambda)s}\right) = -\frac{\ln\left(\frac{s}{b-s}\frac{\lambda(b-s)}{s}\right)}{b} = -\frac{\ln\lambda}{b}$$

Thus $e^{-bt\left(\frac{s}{\lambda b+(1-\lambda)s}\right)} = \lambda$. Substituting this into the formula from Lemma 2, we obtain

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

■

**Proof of Proposition 4.**

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

Recall the definition of $W_T(x)$ from Lemma 7. 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 7 it can be shown that if $t^* > t$ (that is, $W_{t^*}(x) > \frac{r}{\gamma}$), then $W_T(x)$ is strictly increasing in $T$ for $T \in [t, t^*]$ and strictly decreasing in $T$ for $T > t^*$. Hence $W_T(x) > \frac{r}{\gamma}$ 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 $W_{\tau(t)}(x) = \frac{r}{\gamma}$ unless $V(x) \geq \frac{r}{\gamma}$, in which case $\tau(t) = +\infty$.

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 $V(p_{t_0(\sigma)}(m_{t_0(\sigma)}) = \frac{r}{\gamma}$, 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 $m_{t_n(\sigma)}$ 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 $W_{t_{n+1}(\sigma)}(p_t(m_t)) \geq \frac{r}{\gamma}$, as desired.

■

# B A Model of Bad News

**Lemma 10.** *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*

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

**Proof of lemma 10.**

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

$$\begin{aligned}
V^b(x) &= \int_0^\infty (xb + (1-P_\tau)r)\, e^{-\gamma\tau} d\tau \\
&= \int_0^\infty \left(xb + (1-x)(1-e^{-b\tau})r\right) e^{-\gamma\tau} d\tau \\
&= (xb + (1-x)r)\int_0^\infty e^{-\gamma\tau} d\tau - (1-x)r\int_0^\infty e^{-(\gamma+b)\tau} d\tau \\
&= (xb + (1-x)r)\frac{1}{\gamma} - (1-x)r\frac{1}{\gamma+b}
\end{aligned}$$

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 5.**

**Claim 7.1.** 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 7.1.**

Observe that in a model of bad news, we have $p_t(y_t) = \frac{y_t}{y_t + e^{-bt}(1-y_t)}$. Because $p_t(y_t) = \frac{s}{b}$ must be satisfied, using the formula for $p_t(y_t)$ and solving for $y_t$, we obtain $y_t = \frac{s}{s+(b-s)e^{bt}}$.

If the density is uniform, the median is given by $m_t = \frac{1+y_t}{2}$. Substituting the above formula for $y_t$ into $m_t = \frac{1+y_t}{2}$, we obtain $m_t = \frac{1}{2}\frac{2s+(b-s)e^{bt}}{s+(b-s)e^{bt}}$.

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

$$\begin{aligned}
p_t(m_t) = & \frac{\frac{1}{2}\frac{2s+(b-s)e^{bt}}{s+(b-s)e^{bt}}}{\frac{1}{2}\frac{2s+(b-s)e^{bt}}{s+(b-s)e^{bt}} + e^{-bt}\left(1 - \frac{1}{2}\frac{2s+(b-s)e^{bt}}{s+(b-s)e^{bt}}\right)} \\
& = \frac{2s + (b-s)e^{bt}}{2s + (b-s)e^{bt} + e^{-bt}(2(s+(b-s)e^{bt}) - (2s+(b-s)e^{bt}))} \\
& = \frac{2s + (b-s)e^{bt}}{2s + (b-s)(1+e^{bt})}
\end{aligned}$$

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 $V^b\left(\frac{2s}{b+s}\right) \geq \frac{r}{\gamma}$. Because $\frac{b+s}{2b} > \frac{2s}{s+b}$, the fact that $V^b\left(\frac{2s}{b+s}\right) \geq \frac{r}{\gamma}$ implies that $V^b\left(\frac{b+s}{2b}\right) > \frac{r}{\gamma}$. Because, by claim 7.1, $p_0(m_0) = \frac{b+s}{2b}$ and $t \mapsto p_t(m_t)$ is strictly increasing and lemma 10 implies that $V^b$ is increasing, this implies that $V^b(p_t(m_t)) > \frac{r}{\gamma}$ for all $t$.

Then, using the fact that that $f$ is non-decreasing, we can use an argument similar to the one used in the 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 $V^b(p_t(m_t)) \geq \frac{r}{\gamma}$. 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 $\lim_{t\to\infty} p_t(m_t) = \lim_{t\to\infty} \frac{y_t}{y_t+e^{-bt}(1-y_t)} = 1$, $\lim_{t\to\infty} V^b(p_t(m_t)) = V^b(1)$ because $V^b$ is continuous, and $V^b(1) > \frac{r}{\gamma}$.

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

**Claim 7.2.** 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 7.2.**

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

$$\begin{aligned}
&\int_0^{t_+} (xb + (1-P_\tau)r)\, e^{-\gamma\tau} d\tau + \int_{t_+}^{\infty} re^{-\gamma\tau} d\tau \\
&= \int_0^{t_+} \big(xb + (1-x)(1-e^{-bt})r\big)\, e^{-\gamma\tau} d\tau + \int_{t_+}^{\infty} re^{-\gamma\tau} d\tau \\
&= (xb + (1-x)r) \int_0^{t_+} e^{-\gamma\tau} d\tau - (1-x)r \int_0^{t_+} e^{-(\gamma+b)\tau} d\tau + \int_{t_+}^{\infty} re^{-\gamma\tau} d\tau \\
&= (xb + (1-x)r) \frac{1-e^{-\gamma t_+}}{\gamma} - (1-x)r \frac{1-e^{-(\gamma+b)t_+}}{\gamma+b} + e^{-\gamma t_+} \frac{r}{\gamma}
\end{aligned}$$

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 7.3.** Suppose that in some equilibrium $\sigma$ the median $m_{t^0}$ stops experimenting. If for all $t \in [\underline{t}, t^0)$ we have $p_t(m_t)b < r$, then for all $t \in [\underline{t}, t^0)$, $m_t$ stops experimenting.

**Proof of claim 7.3.**

Suppose for the sake of contradiction that this is not the case. Then there exists a non-empty subset $B \subseteq [\underline{t}, t^0)$ 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 $\bar{t}$ and $t^2$ such that $m_{\bar{t}}$ continues experimenting, $m_{t^2}$ stops experimenting and $t^2 - \bar{t} \in (0, \epsilon]$. In particular, if $t^1 = \max\{t : t \in B\}$, then take $\bar{t} = t^1$ 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 $t^1 \notin B$ and, by definition of the supremum, for all $\epsilon > 0$ there exists $\bar{t} \in B$ such that $t^1 - \bar{t} \in (0, \epsilon)$. Then take $t^2 = t^1$ and we are done.

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

By claim 7.2, the payoff to $m_{\bar{t}}$ from continuing experimentation is $V_{\bar{t}} = (p_{\bar{t}}(m_{\bar{t}})b + (1 - p_{\bar{t}}(m_{\bar{t}}))r) \frac{1-e^{-\gamma\epsilon}}{\gamma} - (1 - p_{\bar{t}}(m_{\bar{t}}))r \frac{1-e^{-(\gamma+b)\epsilon}}{\gamma+b} + e^{-\gamma\epsilon} \frac{r}{\gamma}$.

The payoff to stopping experimentation is $\frac{r}{\gamma}$. Then if $V_{\bar{t}} < \frac{r}{\gamma}$, the median strictly prefers to stop experimenting. This is equivalent to

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

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

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

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

**Claim 7.4.** If for all $t \in [\underline{t}, t^0)$ we have $p_t(m_t)b > r$, then in any equilibrium, for all $t \in [\underline{t}, t^0)$, $m_t$ continues experimenting.

**Proof of claim 7.4.**

Suppose for the sake of contradiction that this is not the case. Then there exists a non-empty subset $T' \subseteq [\underline{t}, t^0)$ 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 $t_+$, the payoff to $m_t$ from continuing experimentation is

$$\begin{aligned} V_t &= \int_0^{t_+} (p_t(m_t)b + (1 - P_\tau)r)e^{-\gamma\tau}d\tau + \int_{t_+}^{\infty} re^{-\gamma\tau}d\tau \\ &\geq \int_0^{t_+} p_t(m_t)be^{-\gamma\tau}d\tau + \int_{t_+}^{\infty} re^{-\gamma\tau}d\tau = p_t(m_t)b\frac{1 - e^{-\gamma t_+}}{\gamma} + r\frac{e^{-\gamma t_+}}{\gamma} \end{aligned}$$

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

Let $t_{2n+1} = \sup\left\{t : V^b(p_t(m_t)) \leq \frac{r}{\gamma}\right\}$ 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 $T^1 = \cup_{i=0}^{n}[\underline{t}_i, \bar{t}_i]$.

Suppose first that $p_t(m_t)b \leq r$ for all $t < t_{2n+1}$. In this case, by claim 7.3, 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 7.3 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 $t_{2n}$. We let $V_{t_{2n}}(x)$ denote the value function of an agent with prior $x$ given this continuation equilibrium path. We then let $t_{2n-1} = \sup\left\{t < t_{2n} : V_{t_{2n}}(p_t(m_t)) \leq \frac{r}{\gamma}\right\}$.

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

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

Now note that, if $V_{t_{2n}}(p_{t_{2n-1}}(m_{t_{2n-1}})) = \frac{r}{\gamma}$, then $p_{t_{2n-1}}(m_{t_{2n-1}})b < r$. By continuity, there exists an interval $[\underline{t}_i, \bar{t}_i]$ in $T^1$ such that $t_{2n-1} \in [\underline{t}_i, \bar{t}_i]$ (and $\underline{t}_i$ satisfies $\underline{t}_i = \min\{t < t_{2n-1} : p_t(m_t)b \leq r\}$).

Set $t_{2n-2} = \underline{t}_i$. Because $p_t(m_t)b \leq r$ for all $t \in [t_{2n-2}, t_{2n-1}]$, claim 7.4 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 $V_{t_{2n-2}}(p_t(m_t)) \leq \frac{r}{\gamma}$. Because $T^1$ is finite collection of intervals, the induction terminates in a finite number of steps.

The equilibrium is generically unique for the following reason. Under Assumption 1, each $t_{2k+1}$ satisfies not only $V_{t_{2k+2}}(p_{t_{2k+1}}(m_{t_{2k+1}})) = \frac{r}{\gamma}$ but also $\frac{d}{dt}V_{t_{2k+2}}(p_t(m_t))|_{t_{2k+1}} > 0$, i.e., $V_{t_{2k+2}}(p_t(m_t)) < \frac{r}{\gamma}$ 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 $\frac{d}{dt}p_t(m_t)|_{t_{2k}} < 0$, i.e., $p_t(m_t)b - r > 0$ 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 11.** *An agent with belief $x_t$ at time $t$ 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 11.**

Because agents make their membership decisions based on the expected flow payoffs, an agent with belief $x_t$ at time $t$ 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 12.** *$x \mapsto V_x^i$ is strictly increasing.*

**Proof of lemma 12.**

Consider agents with priors $x' > x$ and suppose that there exists an equilibrium in which an agent with prior $x$ plays some strategy $\sigma_x$. Suppose that the agent with prior $x'$ copies the strategy of the agent with prior $x$. Let $V_{x'}^{ix}$ denote the payoff to the agent with prior $x'$ from copying the 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, at a continuation where $l_t = l$, $x'$'s expected flow payoff is $z_{x'}(l)b + (1 - z_{x'}(l))c$. Because $x' > x$, we have $z_{x'}(l) > z_x(l)$ and thus $z_{x'}(l)b + (1 - z_{x'}(l))c > z_x(l)b + (1 - z_x(l))c$, so the flow payoff of $x'$ is higher than the flow payoff of $x$ when $x$ and $x'$ are inside the organization. Thus $V_{x'}^{ix} > V_x^i$.

Because $V_{x'}^i \geq V_{x'}^{ix}$, we then have $V_{x'}^i \geq V_{x'}^{ix} > V_x^i$, which implies that $V_{x'}^i > V_x^i$, as required. ■

**Corollary 2.** *$V_x^i(l) = V_{z_x(l)}^i(0)$, which is increasing in $z_x(l)$.*

**Proof of Proposition 6.**

Let $z_m(L)$ denote the posterior of the median given $L$. Let $p(L,x)$ denote the posterior belief of an agent with prior $x$ when the state variable is $L$.

**Claim 7.5.** For a uniform distribution of priors, $\lim_{L\to\infty} z_m(L) = \frac{2(s-c)}{(b-c)+(s-c)}$ and $L \mapsto z_m(L)$ is decreasing.

**Proof of claim 7.5.**

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

By lemma 18, if the distribution of priors is power law, i.e., $f(x) = (1-x)^{\alpha}c$, then $z_m(L)$ is decreasing in $L$. In particular, this applies to the uniform distribution if we take $\alpha = 0$, $c = 1$. ■

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, $x \mapsto V^i(x)$ is strictly increasing and, by claim 7.5, $L \mapsto z_m(L)$ is decreasing for a uniform distribution of priors, $L \mapsto V^i(z_m(L))$ is decreasing for a uniform distribution. Thus to ensure that $V^i(z_m(L)) \geq \frac{r}{\gamma}$ for all $L$, it is enough to ensure that $\lim_{L\to\infty} V^i(z_m(L)) \geq \frac{r}{\gamma}$. Because, by claim 7.5, $\lim_{L\to\infty} z_m(L) = \frac{2(s-c)}{(b-c)+(s-c)}$ and $L \mapsto V^i(z_m(L))$ is continuous (because $x \mapsto V^i(x)$ is continuous and $L \mapsto z_m(L)$ is continuous for a uniform distribution), it is enough to ensure that $V^i\left(\frac{2(s-c)}{(b-c)+(s-c)}\right) \geq \frac{r}{\gamma}$.

Next, given that we have shown that $V^i\left(\frac{2(s-c)}{(b-c)+(s-c)}\right) \geq \frac{r}{\gamma}$ implies that $V^i(z_m(L)) \geq \frac{r}{\gamma}$ 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 $V^i(z_m(L)) \geq \frac{r}{\gamma}$ 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 9 that $\lim_{L\to\infty} z_m(L) = \lim_{L\to\infty} z_{\tilde{m}}(L) = \frac{2(s-c)}{(b-c)+(s-c)}$ and an argument similar to the one used in the proof of Proposition 1.

Finally, we show that there exist parameter values such that $V^i\left(\frac{2(s-c)}{(b-c)+(s-c)}\right) \geq \frac{r}{\gamma}$ 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 $V^i(x) \geq x\frac{b}{\gamma} + (1-x)\frac{c}{\gamma}$, which implies that $V^i\left(\frac{2(s-c)}{(b-c)+(s-c)}\right) \geq \frac{1}{\gamma}\frac{(b-c)s+(s-c)b}{(b-c)+(s-c)}$, as required.

Then to show that there exist parameter values such that $V^i\left(\frac{2(s-c)}{b-c+s-c}\right) \geq \frac{r}{\gamma}$ 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 7.**

For convenience, we multiply the all value functions in this proof by $\gamma$.

Let $V^\epsilon_{x,G}(l)$ denote the value function of an agent with prior belief $x$ given that the state is $\ln L(k,t) = l$, the technology is good and the behavior on the equilibrium path is as described in the Proposition. Let $V^\epsilon_{x,B}(l)$ denote the analogous value function given that the technology is bad. Finally, let $V^\epsilon_x(l)$ 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

$$V^\epsilon_{m(l)}(l) \quad = z_{m(l)} V^\epsilon_{m(l),G}(l) + (1 - z_{m(l)}) V^\epsilon_{m(l),G}(l)$$

By Proposition 6, there exist parameters such that there exists an equilibrium in which the organization experiments forever. Note that $V^0_x(l)$ denotes 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 $V^0_{m(l)}(l)$, which we will call $l^*$. Because, by Corollary 2, $V^0_m(l) = V^0_{z_m(l)}(0)$ 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 $z_m(l)$ over $l$ is a singleton. This follows from Lemma 19.[12]

Note that $V^0_{m(l)}(l)$ does not depend on $r$. Thus we can choose $r$ such that $V^0_{m(l^*)}(l^*) = r$. 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 then aim to show that, if we change the equilibrium to require that experimentation stop 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 value functions can be written recursively, in the following sense:

**Claim 7.6.** For any strategy profile $\sigma$, and any two values $l, l' \in \mathbb{R}$, let

$$\begin{aligned}
T_{x,l'}(l) &= \int_0^{+\infty} \gamma e^{-\gamma t} P(\exists s \in [0,t] : l_s = l' | l_0 = l) dt \\
\hat{V}_{x,l'}(l) &= \int_0^{+\infty} \gamma e^{-\gamma t} E(u_x(h^t) \mathbb{1}_{\exists s \in [0,t] : l_s(h^t) = l'} | l_0 = l) dt \\
\tilde{V}_{x,l'}(l) &= \frac{\hat{V}_{x,l'}(l)}{1 - T_{x,l'}(l)},
\end{aligned}$$

where $u_x(h^t)$ is agent $x$'s flow payoff at time $t$ and history $h^t$. Then

$$V_x(l) = (1 - T_{x,l'}(l)) \tilde{V}_{x,l'}(l) + T_{x,l'}(l) V_x(l').$$

Intuitively, $T_{x,l'}(l)$ is the weighted discounted probability that $(l_s)_s$ has hit the value $l'$ by time $t$; $\hat{V}_{x,l'}(l)$ is the expected utility of agent $x$ starting with $l_0 = l$, but setting the continuation value to zero whenever $l_t$ hits $l'$; and $\tilde{V}_{x,l'}(l)$ is a normalization.

[12] Technically we also need the condition $V^0_{m(l^*)}(l^*) < \lim_{l \to +\infty} V^0_{m(l)}(l)$, but this is also satisfied by the example in Lemma 19.

Next, note that, for any $0 \leq \epsilon \leq 1$,

$$V_x^\epsilon(l) = (1 - T_{x,l^*}(l))\tilde{V}_{x,l^*}(l) + T_{x,l^*}(l)V_x^\epsilon(l^*),$$

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

Let $W_x^\epsilon = \lim_{l \searrow l^*} V_x^\epsilon(l)$. $W_x^\epsilon$ 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 $V_x^\epsilon(l^*)$, which is the expected continuation value taken before $m(l^*)$ has decided whether to stop experimenting or not. Specifically

$$\begin{aligned} V_x^\epsilon(l^*) = \epsilon r + (1-\epsilon)W_x^\epsilon = \epsilon r + (1-\epsilon)\left((1 - T_{x,l^*}(l^{*+}))\tilde{W}_x^\epsilon + T_{x,l^*}(l^{*+})V_x^\epsilon(l^*)\right) \\ V_x^\epsilon(l^*) = \frac{\epsilon r + (1-\epsilon)(1 - T_{x,l^*}(l^{*+}))\tilde{W}_x^\epsilon}{1 - (1-\epsilon)(1 - T_{x,l^*}(l^{*+}))} \\ = V_x^0(l^*) + \epsilon \frac{r - V_x^0(l^*)}{1 - (1-\epsilon)(1 - T_{x,l^*}(l^{*+}))}. \end{aligned}$$

Hence

$$\begin{aligned} V_x^\epsilon(l) &= (1 - T_{x,l^*}(l))\tilde{V}_{x,l^*}(l) + T_{x,l^*}(l)\left(V_x^0(l^*) + \epsilon \frac{r - V_x^0(l^*)}{1 - (1-\epsilon)(1 - T_{x,l^*}(l^{*+}))}\right) \\ &= T_{x,l^*}(l)\epsilon \frac{r - V_x^0(l^*)}{1 - (1-\epsilon)(1 - T_{x,l^*}(l^{*+}))} + V_x^0(l). \end{aligned}$$

At the same time, by Lemma 12 and Corollary 2 we have that, for some $\delta > 0$ small enough and all $l \in (l^* - \delta, l^* + \delta)$,

$$\begin{aligned} V_{m(l)}^0(l) = V_{z_{m(l)}(l)}^0(0) = V_{z_{m(l^*)}(l^*)}^0(0) + (z_{m(l)}(l) - z_{m(l^*)}(l^*))\frac{\partial}{\partial x}V_{\tilde{x}}^0(0) \geq \\ \geq V_{z_{m(l^*)}(l^*)}^0(0) + K|l - l^*| = r + K|l - l^*| \end{aligned}$$

for some $K > 0$, as a result of the facts that $\frac{\partial}{\partial x}V_{x,G}^0(0) > 0$ and that $l \mapsto z_{m(l)}(l)$ has a kink at $l^*$. On the other hand, for $l \notin (l^* - \delta, l^* + \delta)$, we have

$$V_{m(l)}^0(l) = V_{z_{m(l)}(l)}^0(0) \geq V_{z_{m(l^*)}(l^*)}^0(0) + K' = r + K'$$

due to the fact that $z_{m(l)}(l) - z_{m(l^*)}(l^*)$ is bounded away from zero in this case.

Let $\epsilon_1$ be small enough that $\frac{T_{x,l^*}(l)}{T_{x,l^*}(l^{*+})}\epsilon_1 \limsup_{l\to l^*}\left|\frac{V^0_{m(l)}(l^*)-V^0_{m(l^*)}(l^*)}{l-l^*}\right| < K$. Let $\epsilon_2$ be small enough that $K' \geq \frac{T_{x,l^*}(l)}{T_{x,l^*}(l^{*+})}\epsilon_2 \max(b-r, r-c)$. Take $\epsilon = \min(\epsilon_1, \epsilon_2)$. Then $V^{\epsilon}_{m(l)}(l) \geq r$ for all $l$, as desired.

(For these arguments to be valid, we need to verify that $T_{x,l^*}(l^{*+}) > 0$ and that $\limsup_{l\to l^*}\left|\frac{V^0_{m(l)}(l^*)-V^0_{m(l^*)}(l^*)}{l-l^*}\right|$ is finite. The former follows from the fact that, taking $t = \frac{\ln(b)-\ln(c)+l^*-l}{b-c}$, we have $P(l_t = l^*|l_0 = l) = P(k(t) = 1) > 0$. The latter follows from the fact that $\frac{\partial}{\partial x}V^0_x(l^*)$ and $m'$ are bounded.)

■

**Proof of Corollary 1.**

Take the example constructed in Proposition 7, and assume that $l(0) = l_0 > l^*$.[13] Let $P_G(l_0)$ be the probability that the organization stops experimenting for some finite $t \geq 0$ when the state of the world is good, and $P_B(l_0)$ the corresponding probability when the state of the world is bad.

We will show that $P_G(l_0) > P_B(l_0)$ for $l_0$ large enough. In fact, we prove a stronger result: we show that there is $C > 0$ such that $P_G(l_0) \geq C > 0$ for all $l_0 > l^*$, but $\lim_{l_0\to+\infty} P_B(l_0) = 0$.

To do this, we define two auxiliary objects. Let $Q_G(l_0)$ be the probability that $l_t \in \left(l^* - \ln(\frac{b}{c}), l^*\right]$ for some $t$ when the state of the world is good, and $Q_B(l_0)$ be the corresponding probability when the state is bad. In words, $Q_\theta(l_0)$ is the probability that $l_t$ ever crosses over to the left of $l^*$.

It can be shown that $Q_G(l_0) = 1$ for all $l_0 > l^*$, but $\lim_{l_0\to+\infty} Q_B(l_0) = 0$. This follows from the fact that, when $\theta = G$, $(l_t)_t$ is a type of random walk with negative drift and hence, in fact, $l_t \to -\infty$ as $t \to +\infty$ with probability one, while when $\theta = G$, $(l_t)_t$ has positive drift and converges to $+\infty$.

From there it follows that $P_B(l_0) \leq Q_B(l_0) \to 0$ as $l_0 \to +\infty$. On the other

[13] Technically, our definition of $l(t)$ forces $l(0) = 0$, but we can relax this assumption by considering a continuation of the game starting at some $t_0 > 0$, where, by assumption, $k(t_0)$ is such that $l(k(t_0), t_0) = l_0$. This example can be fit into our original framework by redefining the density of prior beliefs $\tilde{f}$ to be the density of the posteriors held by agents when $l = l_0$ and $f$ is as in Proposition 7. With this relabeling, $l(0)$ would equal zero and $l^*$ would shift to some negative value, but it is easier to think in terms of shifting $l(0)$ and leaving $f$ unchanged.

hand, $P_G(l_0) \geq Q_G(l_0) \min_{l \in \left(l^* - \ln(\frac{b}{c}), l^*\right]} P_G(l) > 0$. ■

Let $V(\mu, L)$ denote the value function of an agent with prior $\mu$ for a given value of $L$. Then

$$V(\mu, L) = \mu V(\mu, L, G) + (1-\mu) V(\mu, L, B)$$

where $V(\mu, L, \theta)$ is the expected utility of an agent with prior $\mu$ given $L$ and conditional on the true state being $\theta$.

**Lemma 13.** *$V(\mu, L, G)$ and $V(\mu, L, B)$ satisfy the following equations for $L \in \mathbb{R}_{\geq 0} \backslash \mathcal{L}$:*

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

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

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

**Proof of lemma 13.**

Because, by lemma 11, an agent with belief $\mu$ at time $t$ is in the organization at time $t$ if and only if $L \leq \frac{\mu(b-s)}{(1-\mu)(s-c)}$, and, provided that the risky technology is good, an agent's flow payoff during the time period of length $\epsilon$ is

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

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

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 $L\frac{c}{b}e^{(b-c)\epsilon}$. With probability approximately equal to $e^{-b\epsilon}$, a success does not arrive within this time period, which changes the state from $L$ to $Le^{(b-c)\epsilon}$.

Then we have

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

Subtracting $V\left(\mu, Le^{(b-c)\epsilon}, G\right)$ from both sides, we obtain

$$\begin{aligned} V(\mu, L, G) - V\left(\mu, Le^{(b-c)\epsilon}, G\right) \approx \quad & \left(1 - e^{-\gamma\epsilon}\right) \left( \mathbb{1}_{L \leq \frac{\mu(b-s)}{(1-\mu)(s-c)}} (b-s) + s \right) + \\ & + \left(e^{-(\gamma+b)\epsilon} - 1\right) V\left(\mu, Le^{(b-c)\epsilon}, G\right) \\ & + e^{-\gamma\epsilon} \left(1 - e^{-b\epsilon}\right) V\left(\mu, L\frac{c}{b} e^{(b-c)\epsilon}, G\right) \end{aligned}$$

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

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

The boundary conditions $V(\mu, 0, G) = \frac{b}{\gamma}$ and $V(\mu, 0, B) = \frac{c}{\gamma}$ 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 $\int_0^\infty e^{-\gamma\tau} b d\tau = \frac{b}{\gamma}$ if the technology is good and a payoff of $\int_0^\infty e^{-\gamma\tau} c d\tau = \frac{c}{\gamma}$ 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 $\overline{\mathcal{L}} = \{l = \ln(L) : L \in \mathcal{L}\}$ denote the set of the values of $l$ for which the organization stops experimentation. For convenience, we rewrite the equations in lemma 13 with $l = \ln L$ as our state variable. We let $W(\mu, l, G)$ and $W(\mu, l, B)$ denote the resulting value functions.

**Lemma 14.** *$W(\mu, l, G)$ and $W(\mu, l, B)$ satisfy the following equations for $l \in \mathbb{R} \setminus \overline{\mathcal{L}}$:*

$$(b-c)\frac{\partial W(\mu, l, G)}{\partial l} = \gamma \left( \mathbb{1}_{l \leq \ln \frac{\mu(b-s)}{(1-\mu)(s-c)}} (s-b) - s \right) + (\gamma+b) W(\mu, l, G) - bW\left(\mu, l + \ln\frac{c}{b}, G\right)$$

$$(b-c)\frac{\partial W(\mu, l, B)}{\partial l} = \gamma \left( \mathbb{1}_{l \leq \ln \frac{\mu(b-s)}{(1-\mu)(s-c)}} (s-c) - s \right) + (\gamma+b) W(\mu, l, B) - cW\left(\mu, l + \ln\frac{c}{b}, B\right)$$

*$W(\mu, l, \theta) = \frac{r}{\gamma}$ for $l \in \overline{\mathcal{L}}$ and $\theta \in \{B, G\}$. Moreover, the boundary conditions $\lim_{l \to -\infty} W(\mu, l, G) = \frac{b}{\gamma}$ and $\lim_{l \to -\infty} W(\mu, l, B) = \frac{c}{\gamma}$ are satisfied.*

**Proof of lemma 14.**

Note that $\ln\left(L\frac{c}{b}\right) = \ln\left(e^l e^{\ln\frac{c}{b}}\right) = \ln\left(e^{l+\ln\frac{c}{b}}\right) = l + \ln\frac{c}{b}$ and that

$$\frac{\partial W(\mu, l, G)}{\partial l} = \frac{\partial V(\mu, L, G)}{\partial L}\frac{\partial L}{\partial l} = \frac{\partial V(\mu, L, G)}{\partial L}\frac{\partial}{\partial l}\left(e^l\right) = \frac{\partial V(\mu, L, G)}{\partial L}e^l = \frac{\partial V(\mu, L, G)}{\partial L}L$$

which implies that $\frac{\partial V(\mu,L,G)}{\partial L} = \frac{1}{L}\frac{\partial W(\mu,l,G)}{\partial l}$. Note also that $V(\mu, L, G) = W(\mu, l, G)$.

Substituting the formulas for $\frac{\partial V(\mu,L,G)}{\partial L}$ and $V(\mu, L, G)$ into the equations from lemma 13, we obtain the desired equation for $W(\mu, l, G)$ in the statement of lemma 14. The proof for $W(\mu, l, B)$ is similar. ■

We make several definitions that we find convenient to use in the proofs below. We let $d = \ln\frac{\mu(b-s)}{(1-\mu)(s-c)}$ denote the threshold value of $l$ such that an agent is in the organization if and only if $l$ is below this threshold. We let $a = -\ln\frac{c}{b}$ denote the amount by which $l$ decreases after the technology experiences a success. We let $W_0(l) = W(\mu, l, G)$ denote the value function of an agent with prior $\mu$ given that $l \leq d$ and that the technology is good. Finally, we let $W_n(l) = W(\mu, l, G)$ denote the value function an agent with prior $\mu$ given that $l \in (d + (n-1)a, d + na]$ for $n \geq 1$ and given that the technology is good.

**Lemma 15.**

$$W_0(l) = De^{\omega_0 l} + C_0$$

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

**Proof of lemma 15.**

Note that if $W_0(l) = De^{\omega_0 l} + C_0$, then $W_0'(l) = D\omega_0 e^{\omega_0 l}$.

Suppose that $l \leq d$. Then the equation from lemma 14 can be written as

$$(b - c)W_0'(l) = (\gamma + b)W_0(l) - bW_0(l - a) - \gamma b$$

Substituting in the conjectured formula for $W_0(l)$, 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

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

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

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

which pins down $\omega_0$. ■

**Lemma 16.** *If $W_0(l) = De^{\omega_0 l} + b$ for $\omega_0$ satisfying $(b-c)\omega_0 = \gamma+b-be^{-\omega_0 a}$ and some constant $D$, then*

$$W_1(l) = De^{\omega_0 l} + a_0 e^{\omega_1 l} + C_1$$

*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 16.**

Note that if $W_1(l) = D_1 e^{\omega_0 l} + a_0 e^{\omega_1 l} + C_1$ for some constant $D_1$, then $W_1'(l) = \omega_0 D_1 e^{\omega_0 l} + \omega_1 a_0 e^{\omega_1 l}$.

Suppose that $l \in (d, d+a]$, so that $l-a \in (d-a, d]$. Then the equation from lemma 14 can be written as

$$(b-c)W_1'(l) = (\gamma+b)W_1(l) - bW_0(l-a) - \gamma s$$

Substituting in the formulas for $W_1(l)$ and $W_1'(l)$, this is equivalent to

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

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

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

Then the equation simplifies to

$$(b-c)\left(\omega_0 D_1 e^{\omega_0 l} + \omega_1 a_0 e^{\omega_1 l}\right) = (\gamma+b)\left(D_1 e^{\omega_0 l} + a_0 e^{\omega_1 l}\right) - bD_0 e^{\omega_0(l-a)}$$

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

Then the equation simplifies to

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

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

**Lemma 17.**

$$W_n(l) = P_n(l)e^{\omega_1 l} + D_n e^{\omega_0 l} + C_n$$

*where $P_n$ is a polynomial of degree $n-1$, $\omega_0$ satisfies $(b-c)\omega_0 = \gamma + b - be^{\omega_0 a}$ for $a = -\ln\frac{c}{b}$, $\omega_1 = \frac{\gamma+b}{b-c}$ and $D_n = D$ for some constant $D > 0$ for all $n \geq 1$.*

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

$$P_n'(l) = -\frac{b}{(b-c)e^{\omega_1 a}} P_{n-1}(l-a)$$

*for all $n \geq 1$.*

**Proof of lemma 17.**

We will prove the lemma by induction.

Lemma 16 shows that the statement is true for $n = 1$. Suppose as an inductive hypothesis that the statement is true for $n = k$, and consider $W_{k+1}$.

We have

$$W_{k+1}'(l) = P_{k+1}'(l)e^{\omega_1 l} + P_{k+1}(l)\omega_1 e^{\omega_1 l} + D_{k+1}\omega_0 e^{\omega_0 l}$$

and we want to show that

$$(b-c)W'_{k+1}(l) = (\gamma+b)W_{k+1}(l) - bW_k(l-a) - \gamma s$$

Substituting in the formulas for $W'_{k+1}(l)$, $W_{k+1}(l)$ and $W_k(l-a)$, we want to show that

$$\begin{aligned}
&(b-c)\left(P'_{k+1}(l)e^{\omega_1 l} + P_{k+1}(l)\omega_1 e^{\omega_1 l} + D_{k+1}\omega_0 e^{\omega_0 l}\right) = \\
&(\gamma+b)(P_{k+1}(l)e^{\omega_1 l} + D_{k+1}e^{\omega_0 l} + C_{k+1}) - b(P_k(l-a)e^{\omega_1(l-a)} + De^{\omega_0(l-a)} + C_k) - \gamma s
\end{aligned}$$

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

$$\begin{aligned}
&(b-c)\left(P'_{k+1}(l)e^{\omega_1 l} + P_{k+1}(l)\omega_1 e^{\omega_1 l} + D_{k+1}\omega_0 e^{\omega_0 l}\right) = \\
&= (\gamma+b)(P_{k+1}(l)e^{\omega_1 l} + D_{k+1}e^{\omega_0 l}) - b(P_k(l-a)e^{\omega_1(l-a)} + De^{\omega_0(l-a)})
\end{aligned}$$

As in lemma 16, 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

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

■

**Lemma 18.** *If the distribution of priors is power law, then $z_m(L)$ is decreasing in $L$. Moreover, if $L_0 m'(L_0) < m(L_0)(1-m(L_0))$, then $L \mapsto z(L)$ is strictly decreasing at $L = L_0$, and if $L_0 m'(L_0) > m(L_0)(1-m(L_0))$, then $L \mapsto z(L)$ is strictly increasing at $L = L_0$.*

**Proof of lemma 18.**

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

Let $m(L)$ and $y(L)$ denote the median and the marginal members of the organization respectively. 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 $z_m(L) = \frac{1}{1+\left(\frac{1}{m(L)}-1\right)L}$. Then $z'(L) \propto -\frac{\partial}{\partial L}\left(1+\left(\frac{1}{m(L)}-1\right)L\right)$ and $\frac{\partial}{\partial L}\left(1+\left(\frac{1}{m(L)}-1\right)L\right) = \frac{\partial}{\partial L}\left(\left(\frac{1}{m(L)}-1\right)L\right) = \frac{1}{m(L)} - 1 - \frac{L}{(m(L))^2}m'(L)$.

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

After some algebra, using the fact that $y(L) = \frac{s-c}{s-c+(b-c)\frac{1}{L}}$, we get that this is equivalent to $0 < (1-\kappa)(1-\zeta)$, where $\zeta = \frac{s-c}{b-c}$. Since $\kappa$ and $\zeta$ are between 0 and 1, this always holds. ■

**Lemma 19.** *There exist distributions for which $z_m(L)$ is increasing for some values of $L$.*

**Proof of lemma 19.**

We will show that there exist distributions for which $Lm'(L) = m(L)(1-m(L))$ in some interval. Because, by lemma 18, if $L_0 m'(L_0) < m(L_0)(1-m(L_0))$, then $L \mapsto z(L)$ is strictly decreasing at $L = L_0$, this would imply that $L \mapsto z(L)$ is strictly decreasing in some interval.

Let $\tilde{m}(l) = m(e^l)$. Then $Lm'(L) = m(L)(1-m(L))$ is equivalent to $\tilde{m}'(l) = \tilde{m}(l) - \tilde{m}^2(l)$. The general solution to this equation is $\tilde{m}(l) = \frac{e^l}{e^l+C}$, or $m(L) = \frac{L}{L+C}$.

We rewrite $m(L)$ in terms of $y(L)$ as follows. The formula for $y(L)$ implies that $L = \frac{y(L)\frac{b-s}{s-c}}{1-y(L)}$. Substituting this into $m(L) < \frac{L}{L+C}$, we obtain $m(L) < \frac{y(L)(b-s)}{y(L)(b-s-C(s-c))+C(s-c)}$.

Equivalently, we have $m(L) < \frac{(\lambda+\beta)y(L)}{\lambda y(L)+\beta}$ for some $\lambda > 0$ and $\beta > 0$ satisfying $\lambda + \beta = b - s$ (in particular, we have $\beta = C(s-c)$, $\lambda = b - s - C(s-c)$).

For $L \mapsto z_m(L)$ to be increasing, we need that $m(L) < \frac{(\lambda+\beta)y(L)}{\lambda y(L)+\beta}$.

Consider a distribution with a density $f(x) = a_1$ for $x \in [0, b_1]$ and $f(x) = a_2$ for $x \in [b_1, 1]$. We must then have $a_1 b_1 + a_2(1 - b_1) = 1$.

Denote $b - c = \overline{b}$, $s - c = \overline{s}$, $y(L) = y$, $m(L) = m$, $z_{m(L)}(L) = z$.

Let $L_1$ be such that $m(L_1) = b_1$ and $L_2$ be such that $y(L_2) = b_1$. Clearly $0 < L_1 < L_2$. For $L > L_2$, $m(L)$ and $z_m(L)$ are the same as in the uniform case; in particular, $z_m(L) = \frac{2L\overline{s}+\overline{b}-\overline{s}}{L(\overline{s}+\overline{b})+\overline{b}-\overline{s}}$, which is decreasing in $L$.

In general $y = \frac{L\overline{s}}{L\overline{s}+\overline{b}-\overline{s}}$.

For $L \in (L_1, L_2)$, we must have $a_1(b_1 - y) + a_2(m - b_1) = a_2(1 - m)$, i.e., $m = \frac{1+b_1}{2} - \frac{a_1 b_1}{2a_2} + \frac{a_1}{2a_2} y$. Equivalently $m = \left(1 - \frac{1}{2a_2}\right) + \frac{a_1}{2a_2} y = \left(1 - \frac{1}{2a_2}\right) + \frac{a_1}{2a_2} \frac{L\overline{s}}{L\overline{s}+\overline{b}-\overline{s}}$. Then

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

For $L < L_1$, we have $a_1(m - y) = a_1(b_1 - m) + a_2(1 - b_1)$, i.e., $2a_1 m = a_1 b_1 + a_2(1 - b_1) + a_1 y = 1 + a_1 y$, so $m = \frac{1}{2a_1} + \frac{1}{2} y$, and

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

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 $L \mapsto \frac{1}{z} - 1$ is increasing in $(0, L_1)$ and decreasing in $(L_1, L_2)$. In other words, $L \mapsto z_{m(L)}(L)$ is decreasing in $(0, L_1)$ and $(L_2, +\infty)$ but increasing in $(L_1, L_2)$, so $L_1$ is a local minimizer for $z_m(L)$.

Moreover, we can verify that under some extra conditions $L_1$ is a global minimizer: $\lim_{L\to\infty} \frac{1}{z_{m(L)}} - 1 = \frac{\bar{b}-\bar{s}}{2\bar{s}}$, while $\frac{1}{z_{m(L_1)}(L_1)} - 1 = \frac{L_1(1-a_1)\bar{s}+\bar{b}-\bar{s}}{a_1\bar{s}}$. Since $m(L_1) = b_1$,

$$\frac{1}{z_{m(L_1)}(L_1)} - 1 = \frac{L_1}{m(L_1)} - L_1 = \frac{L_1}{b_1} - L_1 = \frac{L_1(1-a_1)\bar{s} + \bar{b} - \bar{s}}{a_1\bar{s}}$$

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

$$\frac{1}{z_{m(L_1)}(L_1)} - 1 = \frac{L_1}{b_1} - L_1 = \frac{\bar{b} - \bar{s}}{\bar{s}} \frac{\frac{1}{b_1} - 1}{\frac{a_1}{b_1} - 1} = \frac{\bar{b} - \bar{s}}{\bar{s}} \frac{1 - b_1}{a_1 - b_1} = \frac{\bar{b} - \bar{s}}{\bar{s}} \frac{2a_1 - 2}{2a_1^2 - a_1 - 1} = \frac{\bar{b} - \bar{s}}{\bar{s}} \frac{1}{a_1 + \frac{1}{2}}$$

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

■

# References


Acemoglu, Daron, Georgy Egorov, and Kostantin Sonin. 2008. "Coalition Formation in Non-Democracies." *Review of Economic Studies* 75(4): 987–1009.

Acemoglu, Daron, Georgy Egorov, and Kostantin Sonin. 2012. "Dynamics and Stability of Constitutions, Coalitions and Clubs." *American Economic Review* 102(4): 1446–1476.

Acemoglu, Daron, Georgy Egorov, and Kostantin Sonin. 2015. "Political Economy in a Changing World." *Journal of Political Economy* 123(5): 1038–1086.

Bai, Jinhui, and Roger Lagunoff. 2011. "On the Faustian Dynamics of Policy and Political Power." *Review of Economic Studies* 78(1): 17–48.

Gieczewski, Germán. 2017. "Policy Persistence and Drift in Organizations." Working Paper.

Keller, Godfrey, and Sven Rady. 2010. "Strategic experimentation with Poisson bandits." *Theoretical Economics* 5(2): 275–311.

Keller, Godfrey, and Sven Rady. 2015. "Breakdowns." *Theoretical Economics* 10(1): 175–202.

Keller, Godfrey, Sven Rady, and Martin Cripps. 2005. "Strategic Experimentation with Exponential Bandits." *Econometrica* 73(1): 39–68.

Strulovici, Bruno. 2010. "Learning While Voting: Determinants of Collective Experimentation." *Econometrica* 78(3): 933–971.