Supplementary Material: "Optimal Design for Social Learning"

Yeon-Koo Che and Johannes Hörner

April 28, 2015

1 Endogenous Entry: Proof of Proposition 3

Given the optimality of revealing good news whenever it is received, we can without loss focus on the agents' check in times and their experimentation decisions as the two control variables. The optimal policy is therefore described by a process $(X_t, \alpha_t)_t$ where $X_t \leq t$ is the right-continuous measure of all agents that have checked in by time t, and $\alpha_t \in [0, 1]$ is the probability of consumption (or experimentation) by agents who check in at time t, conditional on having received no news by t. As will be seen below, the set of all such processes is not rich enough to admit an optimal mechanism. We thus consider a richer space of policies.

Specifically, we allow for the possibility that a mass of agents check in instantaneously—that is, without elapse of any real time—but sequentially experiment at rates that depend on the ℓ. Formally, we enrich the policy space so that whenever X_t jumps at t (so a mass of agents check in at t), we allow the designer to run a second "virtual" clock $s \in [0, m_t]$, where $m_t := X_t - X_t^-$, where $X_t^- := \lim_{t' \uparrow t} X_{t'}$. This virtual clock does not take any real time, but can be used to sequence the agents' check in and experimentation decisions.[1]

Formally, we use (X_t^s, α_t^s) to denote the enriched policy, where X_t^s is the mass of agents who check in by real time t and virtual time s, and α_t^s is the experimentation probability for agents who check in at (t, s). For consistency, we require $X_t^s \geq X_{t'}^-$ and $X_t^{m_t} = X_t$. The virtual clock can be used to split the atom into flows of agents checking in sequentially,

[1]The need for the virtual clock, or the enrichment of the strategy space more generally, arises from the peculiarity of the continuous time game that the soonest next time after any time t is not well defined. To avoid nonexistence, therefore, it is sometimes necessary to allow for sequential moves that do not take any real time. See Simon and Stinchcombe (1989) and Ausubel (2004) for adopting similar enrichment of strategies to guarantee existence of an equilibrium. Note also that the need for the virtual clock disappears once we formulate the policy as a function of ℓ, in which case a split mass is captured by a time "state" variable $t(\ell)$ that is constant for an interval of ℓ's.

in which case X_t^s admits density for all $s \in (0, m_t)$. In such a case, we shall without loss assume that $X_t^s = s + X_t^-$. The virtual clock also admits the entire mass of agents checking in simultaneously, in which case $X_t^s = X_t$ for all s. The current framework also allows for the mass to be split in finite number of lumps.

Throughout, we fix an optimal policy (X, α) and derive properties it must satisfy.

Lemma 1. *At the optimal policy, $t^* := \inf\{t \geq 0 \mid X_t > 0\} > 0$. That is, a mass of agents are induced to wait for a strictly positive amount of time before they check in.*

Proof. Let (ℓ_t, α_t) denote the optimal policy. There exists $T > 0$ such that the surplus accruing to agents checking in at T:

$$e^{-rT}(\ell_0 - \ell_T - \alpha_T(k - \ell_T))$$

is strictly positive, or else the policy will be even inferior to the full transparency. To induce agents to check in any $t < T$, we must have

$$e^{-rt}(\ell_0 - \ell_t - \alpha_t(k - \ell_t)) \geq e^{-rT}(\ell_0 - \ell_T - \alpha_T(k - \ell_T)).$$

The LHS of this inequality is no greater than $\ell_0 - \ell_t$, which goes to zero as $t \to 0$. This proves that there exists $t^* > 0$ such that no agents will check in at $t < t^*$. □

Lemma 2. *Suppose $\hat{t} \in supp(X)$ and $\hat{t} > X_{\hat{t}}^- := \lim_{t \uparrow \hat{t}} X_t$. Then, X jumps at $\hat{t}$.*

The proof consists of two steps.

Step 1. *Suppose X does not jump at $\hat{t}$. Then, there exists $\epsilon > 0$ such that positive density of agents check in $t \in (\hat{t}, \hat{t} + \epsilon)$ and experiment along the locus*

$$\breve{\alpha}(\ell_t) := 1 - \frac{r(k - \ell^0)}{r(k - \ell_t) - \lambda \ell \rho},$$

and the belief evolves according to $\dot{\ell} = -\lambda(\rho + \breve{\alpha}(\ell_t))\ell_t$.

Proof. For $\epsilon > 0$ sufficiently small, X_t admits density $x_t > 0$ for $t \in (\hat{t}, \hat{t} + \epsilon)$. Consider the agents who check in during this time interval. We shall consider the implication of the property of the optimal policy that the designer cannot be better off by redistributing within this time interval while maintaining their incentive to comply with the redistribution. Fix any $t \in (\hat{t}, \hat{t} + \epsilon)$, and let $t' = t + 2\mathrm{d}t < t^* + \epsilon$, and the designer reduces the flow of the agents who check in during $[t, t + \mathrm{d}t)$ by $\delta < x_t$ and increases the flow of the agents who check in during $[t + \mathrm{d}t, t + 2\mathrm{d}t)$ by the same δ, while ensuring that the agents are indifferent over the check in times during that interval. This operation is feasible for sufficiently small

$\delta \in (0, x_t)$. For the original policy to be optimal, such an operation should not lower the posterior (in likelihood) $\ell_{t'}$ (since a lower $\ell_{t'}$ means a higher learning benefit which is strictly preferred under commitment).

We thus require that the operation cannot lower $\ell_{t'}$. To study the effect of the operation on $\ell_{t'}$, let $\alpha_1 := \alpha_t, \alpha_2 := \alpha_{t+\mathrm{d}t}, \alpha_3 := \alpha_{t+2\mathrm{d}t}$, and $\ell_1 := \ell_t$, $\ell_2 := \ell_{t+\mathrm{d}t}, \ell_3 := \ell_{t+2\mathrm{d}t}$. A few conditions must be satisfied:

- Agents are indifferent over consuming at time t and $t + \mathrm{d}t$, *i.e.*,

$$\ell_0 - \ell_1 - \alpha_1(k - \ell_1) = e^{-r\mathrm{d}t}(\ell_0 - \ell_2 - \alpha_2(k - \ell_2)).$$

- Since the flow of consumers picking time t is $(x - \delta)$ while the flow at time $t + \mathrm{d}t$ is $(x + \delta)$, the beliefs at ℓ_2 and ℓ_3 must satisfy

$$\ell_2 = \ell_1 e^{-\lambda(\rho+(x-\delta)\alpha_1)\mathrm{d}t},$$

  and

$$\ell_3 = \ell_2 e^{-\lambda(\rho+(x+\delta)\alpha_2)\mathrm{d}t}.$$

We now fix ℓ_1 and α_1 (and hence the utility for a regular consumer to choose the first instant), and solve this system for ℓ_2, ℓ_3, α_2, as a function of δ, α_1, ℓ_1. Differentiate ℓ_3 with respect to δ, and evaluate the derivative at $\delta = 0$ to get:

$$\operatorname{sgn}\left(\left.\frac{\partial \ell_3}{\partial \delta}\right|_{\delta=0}\right) := \operatorname{sgn}\Big((\alpha(\ell_t) - \alpha_t)(\mathrm{d}t)^2 + o((\mathrm{d}t)^2)\Big).$$

This shows that if $\alpha_t > \check{\alpha}(\ell_t)$, there exists $\delta \in (0, x_t)$ and $t' > 0$ sufficiently small such that the redistribution of agents lowers the posterior at t', a contradiction to its optimality. Hence, we conclude that $\alpha_t \leq \check{\alpha}(\ell_t)$ for any $t \in T_0$.

Next suppose $\alpha_t < \check{\alpha}(\ell_t)$. Then, since $X_t < t$ for all $t \in (\hat{t}, \hat{t} + \epsilon)$, a perturbation from the optimal policy considered above, with $\delta < 0$ (meaning shifting forward the check-in time of some flow of agents), is feasible and will also lower the posterior, contradicting the optimality of the original policy.

Given the experimentation policy follows the locus $\check{\alpha}(\cdot)$, the belief must follow the law of motion. □

Step 2. *X must jump at $\hat{t}$.*

Proof. By Step 1, for some $\epsilon > 0$, a positive flow of agents check in at each $t \in (\hat{t}, \hat{t} + \epsilon)$ and experiment along the locus $\check{\alpha}(\ell_t) := 1 - \frac{r(k-\ell^0)}{r(k-\ell_t)-\lambda\ell\rho}$. For the optimal policy to be incentive compatible, the agents must be indifferent along the locus $\breve{\alpha}(\cdot)$. In particular, the payoff of agents who check in at $t \in (\hat{t}, \hat{t}+\epsilon)$,

$$e^{-rt}(\ell_0 - \ell_t - \breve{\alpha}(\ell_t)(k-\ell_t)), \tag{1}$$

must be constant in t, where the belief ℓ_t evolves according to $\dot{\ell} = -\lambda(\rho + \breve{\alpha}(\ell_t))\ell_t$. One can check that this is not the case. In particular, the payoff (1) decreases in t, which contradicts incentive compatibility. □

Lemma 3. *If X jumps at $\hat{t}$, then $X^s_{\hat{t}}$ admits no atom at $s \in [0, m_{\hat{t}}]$. In other words, any atom of $X_{[\cdot]}$ is split into a flow of agents checking in continuously according to the virtual clock.*

Proof. Suppose to the contrary that $X^{[\cdot]}_{\hat{t}}$ jumps at $\hat{s} \in [0, m_{\hat{t}}]$, and let $m := X^{\hat{s}}_{\hat{t}} - \lim_{s\uparrow\hat{s}} X^s_{\hat{t}}$ be the mass of agents who experiment simultaneously with probability $\alpha > 0$ at the virtual clock $s = \hat{s}$. These agents enjoy the expected payoff of

$$e^{-r\hat{t}}(\ell^0 - \ell^{\hat{s}}_{\hat{t}} + \alpha(k - \ell^{\hat{s}}_{\hat{t}})),$$

where $\ell^{\hat{s}}_{\hat{t}} := \lim_{s\uparrow\hat{s}} \ell^s_{\hat{t}}$ is the belief just before $(\hat{t}, \hat{s})$, and $\ell^{\hat{s}+\mathrm{d}s}_{\hat{t}} := \ell^{\hat{s}}_{\hat{t}} e^{-\lambda(\alpha m + \mathcal{O}(\mathrm{d}s))}$ is the belief just after, for small $\mathrm{d}s > 0$. Consider now a deviation in which the designer splits mass $\delta \in (0, m)$ out of mass m and move it by $\mathrm{d}s$ in the virtual clock. After the experimentation by the first mass $m - \delta$, the belief becomes

$$\hat{\ell}^{\hat{s}+\mathrm{d}s}_{\hat{t}} = \ell^{\hat{s}}_{\hat{t}} e^{-\lambda(\alpha(m-\delta) + \mathcal{O}(\mathrm{d}s))},$$

which is smaller than, and bounded away from, $\ell^{\hat{s}}_{\hat{t}}$ for any $\mathrm{d}s > 0$. The second mass δ of agents can be induced to experiment at rate $\hat{\alpha}$ such that

$$e^{-r\hat{t}}(\ell^0 - \hat{\ell}^{\hat{s}+\mathrm{d}s}_{\hat{t}} + \hat{\alpha}(k - \hat{\ell}^{\hat{s}+\mathrm{d}s}_{\hat{t}})) = e^{-r\hat{t}}(\ell^0 - \ell^{\hat{s}}_{\hat{t}} + \alpha(k - \ell^{\hat{s}}_{\hat{t}})).$$

Since $\hat{\ell}^{\hat{s}+\mathrm{d}s}_{\hat{t}}$ is smaller than and is bounded away from $\ell^{\hat{s}}_{\hat{t}}$, it follows that $\hat{\alpha} - \alpha$ is bounded away from zero for any $\mathrm{d}s$. Hence, the deviation results in the belief,

$$\ell^{\hat{s}}_{\hat{t}} e^{-\lambda(\alpha(m-\delta) + \hat{\alpha}\delta + \mathcal{O}(\mathrm{d}s))} < \ell^{\hat{s}}_{\hat{t}} e^{-\lambda(\alpha m + \mathcal{O}(\mathrm{d}s))} = \ell^{\hat{s}+\mathrm{d}s}_{\hat{t}},$$

for $\mathrm{d}s$ small enough. This is a contradiction to the optimality of the original policy. We therefore conclude that $X^s_{\hat{t}}$ is atomless in s. □

Remark 3. *The argument of Lemma 3 also implies that the optimal policy is not well defined without the enriching of the space. Lemmas 1 and 2 imply that a positive of mass of agents is induced to wait before they check in, and check in at some time $t^* > 0$. But it is never* *optimal for them to check in all simultaneously. The argument given in the proof of Lemma 3 suggests that it is optimal to split the mass and move them later, with an arbitrarily small delay.*

We next investigate an atom X. As Lemma 3 suggests, any atom must be split. We derive a necessary condition for the endpoint of the atom split.

Lemma 4. *Let $(\alpha_t^s, \ell_t^s)_{s\in[0,m_t]}$ be the process of experimentation and beliefs associated with a split atom at t. Then, defining*

$$\bar{\alpha}(\ell) := \frac{r\ell^0 - \ell(\rho\lambda + r)}{rk - \ell(\rho\lambda + r)},$$

it must hold that there exists a (unique) (α_t^s, ℓ_t^s) with $s \in [0, m_t]$ such that $\ell_t^s = \hat{\alpha}(\ell_t^s)$.

Proof. Let there be a mass at t, with $X_t - X_t^- > 0$, and Lemma 3, the mass is split with (α^-, ℓ^-) and (α^+, ℓ^+) denoting start and end points of the split mass. We shall consider a variation which moves the a small segment of the mass forward or backward slightly, and ask when such a variation is profitable. Specifically, we move a segment $\{X_t^s\}_{s'}^t$ of agents with sufficiently small mass $m := X_t^t - X_t^{s'}$ forward in time by a small interval $\mathrm{d}t > 0$ subject to the constraint that the moved agents enjoy the same discounted payoffs, and that the belief at time $t + \mathrm{d}t$ remain the same as before, namely $\ell^+ - \lambda\rho\ell^+\mathrm{d}t$. We then derive the belief at (t, s'), denoted ℓ', that would allow for such a move to be feasible. If $\ell' > \ell_t^{s'}$, then this means that starting from $\ell_t^{s'}$, it is indeed possible to move the segment that would result in the belief at $t + \mathrm{d}t$ being strictly lower than ℓ^+, a welfare improvement. So, the optimality of the original policy will require that $\ell' \leq \ell_t^{s'}$. We shall show that this requirement produces a condition: $\alpha^+ \leq \bar{\alpha}(\ell^+)$.

To begin, consider the variation. First, we require the agents involved in the moved segment to be indifferent to the move. In particular, the agents at the end point of the moved split must enjoy the utility equal to:

$$\mathcal{U}_t = e^{-rt}\left(\ell^0 - \ell^+ - \alpha^+(k - \ell^+)\right).$$

Upon differentiating, this means that

$$(k - \ell^+)\mathrm{d}\alpha^+ = ((1 - \alpha^+)\rho\lambda\ell^+ - r(\ell^0 - \ell^+ - \alpha^+(k - \ell^+)))\mathrm{d}t, \tag{2}$$

using that $\mathrm{d}\ell^+ = -\lambda\rho\ell^+\mathrm{d}t$, which follows from the requirement that the belief at $t + \mathrm{d}t$ must be equal to $\ell^+ - \lambda\rho\ell^+\mathrm{d}t$, the level that would prevail at $t + \mathrm{d}t$ had there been no variation.

Meanwhile, the agents involved in the split atom must be indifferent along the process $(\alpha^\sigma, \ell^\sigma)_{\sigma\in[0,m]}$, where σ is the new virtual clock that is run to sequence the agents who are moved. Let $(\alpha_1, \ell_1) := (\alpha^0, \ell^0)$ be the start point of the moved agents. The end point is $(\alpha^m, \ell^m) = (\alpha^+, \ell^+)$. The indifference means that

$$\ell_0 - \ell^\sigma - \alpha^\sigma (k - \ell^\sigma) = \ell_0 - \ell^+ - \alpha^+ (k - \ell^+),$$

for all $\sigma \in [0, m]$. Hence,

$$\alpha^\sigma = \alpha(\ell^\sigma) := \frac{h - \ell^\sigma}{k - \ell^\sigma}, \tag{3}$$

where $h := \ell^+ + \alpha^+ (k - \ell^+)$.

Now, the equation

$$\frac{\mathrm{d}\ell^\sigma}{ds} = -\lambda \ell^\sigma \alpha(\ell^\sigma)$$

can be solved for the virtual time $s(\ell)$ that it takes to reach a given belief:

$$s'(\ell) = -\frac{k - \ell}{\lambda \ell (h - \ell)},$$

which gives, along the curve,

$$s(\ell) = C - \frac{k \ln \ell + (h - k) \ln(h - \ell)}{\lambda h}, \quad C \in \mathbf{R}.$$

Now recall that $s(\ell^+) - s(\ell_1) = m$. We therefore obtain:

$$m = \frac{k \ln \ell_1 + (h - k) \ln(h - \ell_1)}{\lambda h} - \frac{k \ln \ell^+ + (h - k) \ln(h - \ell^+)}{\lambda h}. \tag{4}$$

We may derive $\mathrm{d}\ell_1$ from this expression, by totally differentiating with respect to α^+ and ℓ^+, using (2). This yields (as a change for a given $\mathrm{d}t$)

$$\mathrm{d}\ell_1 = \frac{\ell_1 (h - \ell_1) \left(\frac{k\left(h - \ell^0\right) r \left(\ln\left(\frac{h - \ell^+}{h - \ell_1}\right) + \ln\left(\frac{\ell_1}{\ell^+}\right) \right)}{h} - \frac{h \lambda \rho (h - \ell_1) \left(k - \ell^+\right) - \left(h - \ell^0\right)\left(\ell^+ - \ell_1\right) r (h - k)}{(h - \ell^+)(h - \ell_1)} \right)}{h (k - \ell_1)} \mathrm{d}t + o(\mathrm{d}t).$$

We now consider doing this for a small change m. That is, we are considering delaying by $\mathrm{d}t$ the experimentation performed by a small mass m (that is, picking a small measure of those agents supposed to check in, and delaying this checking-in –making sure they are willing to wait). To be clear, the change in m is small (so that Taylor expansions apply to (4)), but given this m, we take $\mathrm{d}t$ to be small (so that the previous differential holds approximately). Expanding ℓ_1 from (4) in m gives that (in terms of m)

$$\ell_1 = \ell^+ + \lambda \alpha^+ \ell^+ m + o(m).$$

This is the impact of the mass m on the belief at the end of the splitting. We now delay their experimentation by $\mathrm{d}t$, and use the expression for $\mathrm{d}\ell_1$. Because some background learning would have occurred after the original splitting during the $\mathrm{d}t$ interval, to evaluate the new belief at t, we must account for the learning. To obtain the effect on the new belief at t consistent with the move, we are moving backward in time, so we must add $\rho\lambda\ell_1\mathrm{d}t$ (we had already subtract $\lambda\rho\ell^+\mathrm{d}t$).[2] In sum, the effect on the belief that must prevail at t for the above variation to be possible must equal:

$$\mathrm{d}\ell_1 + \lambda\rho\ell_1\mathrm{d}t = \frac{\ell^+}{k-\ell^+}\left((\alpha^+ k - \ell^0)r + (1-\alpha^+)\ell^+(r+\lambda\rho)\right)\lambda m + o(m).$$

If this expression were positive, this means that the new belief at t consistent with this move is higher than the original belief at (t, s'), meaning that it would be possible to move the agents in a way that keeps the incentives of all agents intact and yields a lower belief at $t + \mathrm{d}t$. Since the latter move would be strictly profitable for the designer, the optimality of the original policy requires the expression above to be nonpositive, or

$$\alpha^+ \leq \frac{r\ell^0 - \ell^+(\rho\lambda + r)}{rk - \ell^+(\rho\lambda + r)}.$$

Similar reasoning applies at the start of the atom splitting, pushing backward in time by $\mathrm{d}t$ a small mass m of experimenters. For this not be profitable, we then get

$$\alpha^- \geq \frac{r\ell^0 - \ell^-(\rho\lambda + r)}{rk - \ell^-(\rho\lambda + r)}.$$

Combining the inequalities, we conclude that the atom splitting must cross the locus $(\ell, \hat{\alpha}(\ell))$. It is readily checked from (3) that the slope of the locus $\alpha_t^s(\ell)$ is larger than the slope $\mathrm{d}\alpha/\mathrm{d}\ell$ along the locus $(\alpha_t^s, \ell_t^s)_{s\in[0,m_t]}$, so that they cross only once. □

Lemma 5. *Suppose $\alpha_t > 0$ at $t > 0$.*[3] *Then, the discounted utility of the agents who check in before cannot be strictly higher.*

Proof. We will prove that the discounted utility of an agent at time t is not boundedly lower than those who check in immediately before. The argument can be extended, as mentioned below, to show that the agents' discounted utility does not decline over time. Suppose that the utility of an agent at time t is boundedly lower than the utility of an agent who has not checked in at time $t - \epsilon$, for all $\epsilon > 0$. The reasoning below assumes a gradual check in over some interval $[t-\Delta, t+\Delta]$, but can be adjusted (in case there is a mass point). Let α denote

[2] The change of experimentation due to agents arriving during the interval $[t, t + \mathrm{d}t)$ and possibly asked to experiment is of order $\mathrm{d}t \cdot \mathrm{d}m$ and so ignored.

[3] In case X has a mass point at t, we select $\alpha_t := \sup_s \alpha_t^s$.

the probability of recommendation at times $[t-\Delta, t)$. More precisely, let $\alpha = \lim_{\tau \uparrow t} \alpha_\tau$ and let $\Delta > 0$ be such that $|\alpha_\tau - \alpha| < \varepsilon$ for all $\tau \in [t-\Delta, t)$, for some fixed $\varepsilon > 0$ arbitrarily small.)

Similarly, let α' denote the probability of recommendation at times $[t, t+\Delta)$ (again, take limits in the obvious way.) Because the utility is boundedly lower at times in $[t, t+\Delta)$, we may increase the probability α by some $\delta > 0$ and decrease α' by the same amount, so that (i) the total experimentation (and hence belief) at time $t+\Delta$ is the same before and after the change, (ii) agents supposed to check in at times $[t-\Delta, t)$ prefer to do so than to check at times $[t, t+\Delta)$. Clearly, agents checking in at time $[t, t+\Delta)$ gain from this change and so will still check in.[4]

We ask, when does such a change increase welfare? By construction, it does not affect what happens before (agents before certainly don't want to wait now) nor after (the belief at $t+\Delta$ hasn't changed). Hence, the change in payoff is

$$(\ell_0 - \ell - (\alpha+\delta)(k-\ell) + e^{-r\Delta}(\ell_0 - \ell_\Delta - (\alpha'-\delta)(k-\ell_\Delta)) + o(\Delta),$$

where ℓ, ℓ_Δ are the beliefs at the beginning of each subinterval. Taking derivatives with respect to δ and then a Taylor expansion with respect to Δ, evaluated at $\delta = 0$, gives that the derivative equals

$$(\ell(1+\rho+r) - rk)\Delta + o(\Delta),$$

where we normalize λ to 1 (Alternatively, replace all occurrences of r by r/λ). Hence, such a change is profitable if $\ell > \frac{rc}{1+\rho+r}$. Thus, we may assume otherwise.

If the change in utility is not bounded below by some constant, the same reasoning applies, but $\delta = \delta(\tau)$ must be chosen so that $\delta(t) = 0$, $\delta(\tau) > 0$ for $\tau < t$, and $\delta(\tau) > 0$ for $\tau > t$, such that agents do not wish to change their check-in time over these intervals. If there is an atom at t, then there must be an black-out immediately before t, and a similar reasoning applies for moving a small mass m' of split atom backward in time and raise their experimentation by small δ. Both extensions are omitted.

Let us now recall that the total continuation payoff is given by

$$J = \int_{s \geq t} e^{-rs}(\ell^0 - \ell_s - \alpha_s(k-\ell_s))\mathrm{d}s.$$

Because $\dot{\ell} = -(\rho+\alpha)\ell_s$, we can substitute to obtain an expression that only depends on $\ell, \dot{\ell}$, integrate by parts to eliminate the terms involving $\dot{\ell}$, and ignoring constants, obtain that,

[4] Adding small $\delta > 0$ to α does not violate the incentive constraint for consumption since the constraint is not binding.

up to a constant, J is equal to

$$\begin{aligned}
J &= \int_{s\geq t} e^{-rs}(\ell^0 - \ell_s - \alpha_s(k-\ell_s))\mathrm{d}s \\
&= \int_{s\geq t} e^{-rs}(\ell^0 - \ell_s + (\rho + \frac{\dot{\ell}}{\lambda \ell_s})(k-\ell_s))\mathrm{d}s \\
&= \int_{s\geq t} e^{-rs}\left(\ell^0 - (1+\rho)\ell_s - \frac{\dot{\ell}_s}{\lambda} + \frac{\dot{\ell}}{\lambda \ell_s}k\right)\mathrm{d}s \\
&= \int_{s\geq t} e^{-rs}(\ell^0 - (1+\rho)\ell_s)\mathrm{d}s - e^{-rt}\left(\frac{\ell_s}{\lambda} - \frac{\ln \ell_s}{\lambda}k\right)_t^\infty - \int_{s\geq t} e^{-rs} r\left(\frac{\ell_s}{\lambda} - \frac{\ln \ell_s}{\lambda}k\right)\mathrm{d}s \\
&= -\frac{1}{\lambda}\int_{s\geq t} e^{-rs}\left(\lambda(1+\rho)+r)\ell_s - rk\ln \ell_s\right)\mathrm{d}s + \text{Const.}
\end{aligned}$$

The derivative of the integrand with respect to ℓ is

$$rk - (\lambda(1+\rho)+r)\ell > 0.$$

Hence, given t as defined above, we note that this derivative is positive: if we replace the trajectory $\{\ell_\tau : \tau \geq t\}$ by a trajectory $\{\hat{\ell}_\tau : \tau \geq t\}$, with $\hat{\ell}_\tau \geq \ell_\tau$ (with a strict inequality for some non-zero measure interval of times), the payoff increases. Hence, decrease α at time t (or rather, fix a time higher than, but arbitrarily close to t and decrease α at that time), and adjust α_τ for all later $\tau \in T_0$ so that incentives to check in do not change. This requires non-positive changes in α_τ (which can be expressed in terms of a differential equation), and results in a higher trajectory $\hat{\ell}$, and hence an increase in payoff. We thus obtain a contradiction. □

Lemma 6. *$X_t = 0$ for $t < t^*$ and $X_t = t$ for $t \geq t^*$.*

Proof. By Lemma 1, $t^* := \inf\{t \geq 0 | X_t > 0\} > 0$. Since $X^-_{t^*} = 0 < t^*$, by Lemma 2, X has a mass point at t^*. Further, by Lemma 3, the atom at t^* must be split. Let (α^+, ℓ^+) be the end point of the split atom at t^*. Then, by Lemma 4, $\alpha^+ \leq \hat{\alpha}(\ell^+)$. One can show hat $\hat{\alpha}(\cdot)$ is steeper than $\alpha(\cdot)$.

We next show that for any $t > t^*$, $X_t = t$. Suppose to the contrary that there exists $t' > t$ such that $X_{t'} < t'$. Then, by Lemmas 2–4, X has an atom at t', and it is split, and the splitting crosses the locus $\hat{\alpha}(\cdot)$. But this is impossible, since by Lemma 5, the agents' utility never strictly decrease for $t \in [t^*, t']$, and this means that during that interval, α can never rise at a faster rate in ℓ than $\alpha(\ell)$ (the locus followed in the first split atom) does. Since $X_{t'} = t'$ for all $t' > t$, and since X is required to be right continuous, the claim follows. □

Armed with the lemmas, we now complete our characterization of the optimal policy.

Proof of Proposition 3 and the solution algorithm. Lemmas 4-6 pin down the structure of the optimal policy: there exist times $t^* > 0$ and $T > t^*$ such that agents who arrive before t^* wait until t^*; the accumulated mass is split at t^*; and then the agents who arrive after t^* check in upon arrival and experiment at a rate that falls to zero at T. All agents' discounted payoff is constant for all argents arriving prior to T, and the agents arriving after T enjoys payoff according to *full transparency regime*. This structure, along with further necessary conditions, enables us to derive an one-dimensional family of optimal policies indexed by the belief $\bar{\ell}$ at which agents' experimentation stops fully.

Initially, we fix both $\bar{\ell}$ and t^*. The variables to be determined are $(\ell^-, \alpha^+, \ell^+, T)$, where T is such that $\ell_T = \bar{\ell}$. Several conditions are derived to determine these variables. First, since only background learning occurs during the blackout, the (designer's) belief just prior to the splitting must satisfy

$$\ell^- = \ell^0 e^{-\lambda\rho t^*}. \tag{5}$$

Second, the agents associated with split mass must be indifferent across the locus $(\alpha_t^s, \ell_t^s)_{s\in[0,t^*]}$. This requires (3). This, together with the fact that the entire mass of t^* is split, gives rise to an equation (4) with $m = t^*$, or

$$t^* = \frac{k \ln \ell^- + (h-k)\ln(h-\ell^-)}{\lambda h} - \frac{k \ln \ell^+ + (h-k)\ln(h-\ell^+)}{\lambda h}, \tag{6}$$

where $h := \ell^+ + \alpha^+(k - \ell^+)$.

Third, by Lemma 6, agents check in as they arrive during the the smooth tapering phase, and by Lemma 5, they must experiment at levels that make them all indifferent. Hence, during the $t \in (t^*, T)$, (α_t, ℓ_t) must satisfy the indifference condition:

$$e^{-rt}(\ell^0 - \ell_t - \alpha_t(k - \ell_t)) = e^{-rT}(\ell^0 - \bar{\ell}), \tag{7}$$

and the belief evolution condition:

$$\ell_t = \bar{\ell} e^{\lambda \int_t^T (\rho + \alpha_s)\mathrm{d}s}. \tag{8}$$

Given $(t^*, \bar{\ell})$, conditions (7) and (8) uniquely determine (α^+, ℓ^+) as a function of T.[5]

In sum, the conditions (5)–(8) uniquely pin down $T, \ell^-, \alpha^+, \ell^+$ as functions of $(\bar{\ell}, t^*)$. Next, we hold $\bar{\ell}$ fixed and characterize the condition for optimal choice of t^*. In particular, we use the fact that for a fixed $\bar{\ell}$, T must be minimized. This follows from the fact that the earlier time T at which a given belief $\bar{\ell}$ is reached without affecting the payoff of the

[5]Essentially, given any $T > t^*$, the two conditions give rise to a differential equation that runs from T backward to t^*, with the initial value $(\alpha_T, \ell_T) = (0, \bar{\ell})$, which admits a unique solution (α_t, ℓ_t), the end point of which is (α^+, ℓ^+).

agents who arrive before, the higher the payoff is for all agents arriving afterwards, and thus the higher the overall welfare is. This means that, as t^* is varied slightly by $\mathrm{d}t$, subject to the constraint that all agents who arrive before T should enjoy the same payoff, the optimal choice of T remains constant. Since the variation should not alter T at the optimum, as t^* is raised by $\mathrm{d}t$, the last phase is shortened by length $\mathrm{d}t$. Hence, the variation affects ℓ^+, the belief at which (going backward) the differential equation ends, by:

$$\mathrm{d}\ell^+ = -\lambda(\rho + \alpha^+)\ell^+ \mathrm{d}t. \tag{9}$$

Next, the variation keeps the payoff of early agents constant. Totally differentiating the indifference condition, we obtain the change in α^+ arising from the variation:

$$r(\ell_0 - \ell^+ - \alpha^+(k - \ell^+))\mathrm{d}t = -(1 - \alpha^+)\mathrm{d}\ell^+ - (k - \ell^+)\mathrm{d}\alpha^+. \tag{10}$$

Now, since the initial phase has been lengthened by $\mathrm{d}t$, we should have

$$\ell^- + \mathrm{d}\ell^- = \ell_0 e^{-\lambda\rho(t^* + \mathrm{d}t)},$$

or

$$\frac{\mathrm{d}\ell^-}{\ell^-} = -\lambda\rho \mathrm{d}t. \tag{11}$$

Finally, substituting from (5) into (6) to eliminate t^*, we get

$$\frac{\ell^+ + \alpha^+(k - \ell^+)}{\rho} \ln \frac{\ell^-}{\ell_0} = k \ln \frac{\ell^+}{\ell^-} + (1 - \alpha^+)(k - \ell^+) \ln\left(1 + \frac{\ell^+ - \ell^-}{\alpha^+(k - \ell^+)}\right), \tag{12}$$

The effect of the variation on the end points of split mass can be obtained by totally differentiating (12). The resulting equation, after (9)–(11) are substituted into it, must hold for any small $\mathrm{d}t$. This gives rise to another condition, which is too long and cumbersome to include here. This condition, together with the earlier observations, pins down $(t^*, T, \ell^-, \alpha^+, \ell^+)$ as functions of only one variable, $\bar{\ell}$. (Incidentally, one can vary $\bar{\ell}$ and trace out the locus of (ℓ^+, α^+) at which the atom splitting terminates under the optimal policy, yielding a dashed locus in Figure 2.)

Let $t^*(\bar{\ell}), \alpha^+(\bar{\ell}), \ell^+(\bar{\ell}), T(\bar{\ell})$ be the key variables of the optimal policy as functions of $\bar{\ell}$. The resulting welfare is:

$$\mathcal{U}(\bar{\ell}) := T(\bar{\ell})e^{-rt^*(\bar{\ell})}(\ell^0 - \ell^+(\bar{\ell}) - \alpha^+(t^*)(k - \ell^+(\bar{\ell}))) + \int_{T(\bar{\ell})}^{\infty} e^{-rt}(\ell^0 - \ell_t)\mathrm{d}t.$$

where

$$\ell_t := \bar{\ell}e^{-\lambda\rho(t - T(\bar{\ell}))}, \forall t \geq T(\bar{\ell}).$$

To characterize the optimal policy, it remains to choose $\bar{\ell}$ to maximize $\mathcal{U}(\bar{\ell})$. A closed form solution on the optimal $\bar{\ell}$, or a simple characterization, is difficult to come by.[6] But a numerical solution for the optimal $\bar{\ell}$, and thus the optimal policy, can be obtained. Figure 2 in the paper describes the optimal policy for the case $(r, \lambda, \rho, k, \ell^0) = (1/2, 2, 1/5, 15, 9)$.

The optimal policy employs a "blackout" until $t^* \simeq 1.45$, and the mass accumulated by then is then split at time t^* into (sequential) check in; this mass experiments along the locus that starts from $(\alpha^-, \ell^-) \simeq (0.017, 5.04)$ and ends at $(\alpha^+, \ell^+) \simeq (.066, 4.52)$. After time t^*, the agents check in as they arrive. They experiment along the locus of (α, ℓ)'s that begins at (α^+, ℓ^+) and tapers gradually down to $(\alpha, \ell) \simeq (0, 2.72)$. □

2 General Signal Structure: Proof from Section 7.1

Here, we extend our model to allow for both good news and bad news. Specifically, if a flow of size μ consumes the good over some time interval $[t, t+\mathrm{d}t)$, then the designer learns during this time interval that the movie is "good" with probability $\lambda_g(\rho+\mu)\mathrm{d}t$, that it is "bad" with probability $\lambda_b(\rho+\mu)\mathrm{d}t$, where $\lambda_g, \lambda_b \geq 0$, and ρ is the rate of background learning.

The designer commits to the following policy: At time t, she recommends the movie to a fraction $\gamma_t \in [0,1]$ of agents if she learns the movie to be good, a fraction $\beta_t \in [0,1]$ if she learns it to be bad, and she recommends to fraction $\alpha_t \in [0,1]$ if no news has arrived by t. Clearly,

$$\mu_t = \rho + \alpha_t.$$

The designer's belief evolves according to

$$\dot{p}_t = -(\lambda_g - \lambda_b)\mu_t p_t(1-p_t), \tag{13}$$

with the initial value $p_0 = p^0$. It is worth noting that the evolution of the posterior depends on the relative arrival rates of the good news and the bad news. If $\lambda_g > \lambda_b$ (so the good news arrive faster than the bad news), then "no news" leads the designer to form a pessimistic inference on the quality of the movie, with the posterior falling. By contrast, if $\lambda_g < \lambda_b$, then "no news" leads to on optimistic inference, with the posterior rising. We label the former case **good news** case and the latter **bad news** case. Recall that main body of the paper treats the special case of $\lambda_b = 0$, a pure good news case.

Let g_t and b_t denote the probability that the designer's belief is 1 and 0, respectively.

[6] In particular, we have no proof that this constrained maximization admits a unique solution $\bar{\ell}$. There might be (presumably non-generic) parameter configurations for which this is the case, in which case there would be multiple optimal policies.

Given the experimentation rate μ_t, these probabilities evolve according to

$$\dot{g}_t = (1 - g_t - b_t)\lambda_g \mu_t p_t, \tag{14}$$

with the initial value $g_0 = 0$, and

$$\dot{b}_t = (1 - g_t - b_t)\lambda_b \mu_t (1 - p_t), \tag{15}$$

with the initial value $b_0 = 0$.[7] Further, these beliefs must form a martingale:

$$p_0 = g_t \cdot 1 + b_t \cdot 0 + (1 - g_t - b_t)p_t. \tag{16}$$

The designer chooses the policy (α, β, γ), measurable, to maximize social welfare, namely

$$\mathcal{W}(\alpha, \beta, \chi) := \int_{t \geq 0} e^{-rt} g_t \gamma_t (1 - c)\mathrm{d}t + \int_{t \geq 0} e^{-rt} b_t \beta_t (-c)\mathrm{d}t + \int_{t \geq 0} e^{-rt} (1 - g_t - b_t)\alpha_t (p_t - c)\mathrm{d}t,$$

where (p_t, g_t, b_t) must follow the required laws of motion: (13), (14), (15), and (16), where $\mu_t = \rho + \alpha_t$ is the total experimentation rate and r is the discount rate of the designer.[8]

Given policy (α, β, γ), conditional on being recommended to watch the movie, the agent will have the incentive to watch the movie, if and only if the expected quality of the movie—the posterior that it is good—is no less than the cost, or

$$\frac{g_t \gamma_t + (1 - g_t - b_t)\alpha_t p_t}{g_t \gamma_t + b_t \beta_t + (1 - g_t - b_t)\alpha_t} \geq c. \tag{17}$$

The following is immediate:

Lemma 7. *It is optimal for the designer to disclose the breakthrough (both good and bad) news immediately. That is, an optimal policy has $\gamma_t \equiv 1, \beta_t \equiv 0$.*

Proof. If one raises γ_t and lowers β_t, it can only raise the value of objective $\mathcal{W}$ and relax (17) (and do not affect other constraints). ☐

[7] These formulae are derived as follows. Suppose the probability that the designer has seen the good news by time t and the probability that she has seen the bad news by t are respectively g_t and b_t. Then, the probability of the good news arriving by time $t + \mathrm{d}t$ and the probability of the bad news arriving by time $t + \mathrm{d}t$ are, respectively, and to the first-order,

$$g_{t+\mathrm{d}t} = g_t + \lambda_g \mu_t p_t \mathrm{d}t(1 - g_t - b_t) \text{ and } b_{t+\mathrm{d}t} = b_t + \lambda_b \mu_t (1 - p_t)\mathrm{d}t(1 - g_t - b_t).$$

Dividing these equations by $\mathrm{d}t$ and taking the limit as $\mathrm{d}t \to 0$ yields (14) and (15).

[8] More precisely, the designer is allowed to randomize over the choice of policy (α, β, γ) (using a relaxed control, as such randomization is defined in optimal control). A corollary of our results is that there is no gain for him from doing so.

Using $\ell_t = \frac{p_t}{1-p_t}$, (13) can be restated as:

$$\dot{\ell}_t = -\ell_t \Delta \lambda_g \mu_t, \quad \ell_0 := \frac{p_0}{1-p_0}, \tag{18}$$

where $\Delta := \frac{\lambda_g - \lambda_b}{\lambda_g}$, assuming for now $\lambda_g > 0$.

The two other state variables, namely the posteriors g_t and b_t on the designer's belief, are pinned down by ℓ_t (and thus by p_t) at least when $\lambda_g \neq \lambda_b$ (*i.e.*, when no news is not informationally neutral.) (We shall remark on the case of the neutrality case $\Delta = 0$.)

Lemma 8. *If $\Delta \neq 0$, then*

$$g_t = p_0 \left(1 - \left(\frac{\ell_t}{\ell_0}\right)^{\frac{1}{\Delta}}\right) \text{ and } b_t = (1-p_0)\left(1 - \left(\frac{\ell_t}{\ell_0}\right)^{\frac{1}{\Delta}-1}\right).$$

Proof. Let $\kappa_t := p_0/(p_0 - g_t)$. Note that $\kappa_0 = 1$. Then, it follows from (14) and (16) that

$$\dot{\kappa}_t = \lambda_g \kappa_t \mu_t, \quad \kappa_0 = 1. \tag{19}$$

Dividing both sides of (19) by the respective sides of (18), we get,

$$\frac{\dot{\kappa}_t}{\dot{\ell}_t} = -\frac{\kappa_t}{\ell_t \Delta},$$

or

$$\frac{\dot{\kappa}_t}{\kappa_t} = -\frac{1}{\Delta}\frac{\dot{\ell}_t}{\ell_t}.$$

It follows that, given the initial condition,

$$\kappa_t = \left(\frac{\ell_t}{\ell_0}\right)^{-\frac{1}{\Delta}}.$$

We can then unpack κ_t to recover g_t, and from this we can obtain b_t via (16). □

This result is remarkable. A priori, there is no reason to expect that the designer's belief p_t serves as a "sufficient statistic" for the posteriors that the agents attach to the arrival of news, since different histories for instance involving even different experimentation over time could in principle lead to the same p.

Next, substitute g_t and b_t into (17) to obtain:

$$\alpha_t \leq \bar{\alpha}(\ell_t) := \min\left\{1, \frac{\left(\frac{\ell_t}{\ell_0}\right)^{-\frac{1}{\Delta}} - 1}{k - \ell_t}\ell_t\right\}, \tag{20}$$

if the normalized cost $k := c/(1-c)$ exceeds ℓ_t and $\bar{\alpha}(\ell_t) := 1$ otherwise.

The next lemma will figure prominently in our characterization of the second-best policy later.

Lemma 9. *If $\ell_0 < k$ and $\Delta \neq 0$, then $\bar{\alpha}(\ell_t)$ is zero at $t = 0$, and increasing in t, strictly so whenever $\bar{\alpha}(\ell_t) \in [0, 1)$.*[9]

Proof. We shall focus on

$$\tilde{\alpha}(\ell) := \frac{\left(\frac{\ell}{\ell_0}\right)^{-\frac{1}{\Delta}} - 1}{k - \ell}\ell.$$

Recall $\bar{\alpha}(\ell) = \min\{1, \tilde{\alpha}(\ell)\}$. Since ℓ_t falls over t when $\Delta > 0$ and rises over t when $\Delta < 0$. It suffices to show that $\tilde{\alpha}(\cdot)$ is decreasing when $\Delta > 0$ and increasing when $\Delta < 0$.

We make several preliminary observations. First, $\tilde{\alpha}(\ell) \in [0, 1)$ if and only if

$$1 - (\ell/\ell_0)^{\frac{1}{\Delta}} \geq 0 \text{ and } k\ell^{\frac{1}{\Delta}-1}\ell_0^{-\frac{1}{\Delta}} > 1. \tag{21}$$

Second,

$$\tilde{\alpha}'(\ell) = \frac{(\ell_0/\ell)^{\frac{1}{\Delta}} h(\ell, k)}{\Delta(k - \ell)^2}, \tag{22}$$

where

$$h(\ell, k) := \ell - k(1 - \Delta) - k\Delta(\ell/\ell_0)^{\frac{1}{\Delta}}.$$

Third, (21) implies that

$$\frac{\mathrm{d}h(\ell, k)}{\mathrm{d}\ell} = 1 - k\ell^{\frac{1}{\Delta}-1}\ell_0^{-\frac{1}{\Delta}} < 0, \tag{23}$$

on any range of ℓ over which $\tilde{\alpha} \leq 1$. Note

$$h(0, k) = -k(1 - \Delta) = -k\frac{\lambda_b}{\lambda_g} \leq 0. \tag{24}$$

It follows from (23) and (24) that $h(\ell, k) < 0$ for any $\ell \in (0, k)$ and $\tilde{\alpha}(\ell) \in [0, 1)$. By (22), this last fact implies that $\tilde{\alpha}'(\ell) < 0$ if $\Delta > 0$ and $\tilde{\alpha}'(\ell) > 0$ if $\Delta < 0$, as was to be shown. □

[9] The case $\Delta = 0$ is similar: the same conclusion holds but $\bar{\alpha}$ need to be defined separately.

Substituting the posteriors from Lemma 8 into the objective function and using $\mu_t = \rho + \alpha_t$, and with normalization of the objective function, the second-best program is restated as follows:

$$[SB] \qquad \sup_{\alpha} \int_{t \geq 0} e^{-rt} \ell_t^{\frac{1}{\Delta}} \left(\alpha_t \left(1 - \frac{k}{\ell_t} \right) - 1 \right) \mathrm{d}t$$

subject to

$$\dot{\ell}_t = -\Delta \lambda_g (\rho + \alpha_t) \ell_t, \tag{25}$$

proposition

$$0 \leq \alpha_t \leq \bar{\alpha}(\ell_t). \tag{26}$$

Obviously, the first-best program, labeled $[FB]$, is the same as $[SB]$, except that the upper bound for $\bar{\alpha}(\ell_t)$ is replaced by 1. We next characterize the optimal recommendation policy. The precise characterization depends on the sign of Δ, *i.e.*, whether the environment is that of predominantly good news or bad news.

2.1 "Good news" environment: $\Delta > 0$

The analysis is similar to that for Proposition 1 in the paper. As in the paper, we first switch the roles of variables so that we treat ℓ as a "time" variable and $t(\ell) := \inf\{t | \ell_t \leq \ell\}$ as the state variable, interpreted as the time it takes for a posterior ℓ to be reached. Up to constant (additive and multiplicative) terms, the designer's problem is written as: For problem $i = SB, FB$,

$$\sup_{\alpha(\ell)} \int_0^{\ell^0} e^{-rt(\ell)} \ell^{\frac{1}{\Delta} - 1} \left(1 - \frac{k}{\ell} - \frac{\rho \left(1 - \frac{k}{\ell}\right) + 1}{\rho + \alpha(\ell)} \right) \mathrm{d}\ell.$$

s.t. $t(\ell^0) = 0$,

$$t'(\ell) = -\frac{1}{\Delta \lambda_g (\rho + \alpha(\ell)) \ell},$$

$$\alpha(\ell) \in \mathcal{A}^i(\ell),$$

where $\mathcal{A}^{SB}(\ell) := [0, \bar{\alpha}(\ell)]$, and $\mathcal{A}^{FB} := [0, 1]$.

This transformation enables us to focus on the optimal recommendation policy directly as a function of the posterior ℓ. Given the transformation, the admissible set no longer depends on the state variable (since ℓ is no longer a state variable), thus conforming to the standard specification of the optimal control problem.

Next, we focus on $u(\ell) := \frac{1}{\rho+\alpha(\ell)}$ as the control variable. With this change of variable, the designer's problem (both second-best and first-best) is restated, up to constant (additive and multiplicative) terms: For $i = SB, FB$,

$$\sup_{u(\ell)} \int_0^{\ell^0} e^{-rt(\ell)} \ell^{\frac{1}{\Delta}-1} \left(1 - \frac{k}{\ell} - \left(\rho\left(1 - \frac{k}{\ell}\right) + 1\right) u(\ell)\right) \mathrm{d}\ell, \tag{27}$$

s.t. $t(\ell^0) = 0$,

$$t'(\ell) = -\frac{u(\ell)}{\Delta\lambda_g \ell},$$

$$u(\ell) \in \mathcal{U}^i(\ell),$$

where the admissible set for the control is $\mathcal{U}^{SB}(\ell) := [\frac{1}{\rho+\bar{\alpha}(\ell)}, \frac{1}{\rho}]$ for the second-best problem and $\mathcal{U}^{FB}(\ell) := [\frac{1}{\rho+1}, \frac{1}{\rho}]$. With this transformation, the problem becomes a standard linear optimal control problem (with state t and control α). A solution exists by the Filippov-Cesari theorem (Cesari, 1983).

The characterization of the solution is summarized in the proposition which extends Proposition 1 of the paper for the general good news case.

Proposition 1. *The second-best policy prescribes, absent any news, the maximal experimentation at $\alpha(p) = \bar{\alpha}(\frac{p}{1-p})$ until the posterior falls to p_g^*, and no experimentation $\alpha(p) = 0$ thereafter for $p < p_g^*$, where*

$$p_g^* := c\left(1 - \frac{rv}{\rho + r(v + \frac{1}{\lambda_g})}\right),$$

where $v := \frac{1-c}{r}$ is the continuation payoff upon the arrival of good news. The first-best policy has the same structure with the same threshold posterior, except that $\bar{\alpha}(p)$ is replaced by 1. If $p_0 \geq c$, then the second-best policy implements the first-best, where neither No Social Learning nor Full Transparency can. If $p_0 < c$, then the second-best induces a slower experimentation/learning than the first-best.

Proof. We first focus on the necessary condition for optimality to characterize the optimal recommendation policy. To this end, we write the Hamiltonian:

$$\mathcal{H}(t, u, \ell, \nu) = e^{-rt(\ell)} \ell^{\frac{1}{\Delta}-1} \left(1 - \frac{k}{\ell} - \left(\rho\left(1 - \frac{k}{\ell}\right) + 1\right) u(\ell)\right) - \nu \frac{u(\ell)}{\Delta\lambda_g \ell}. \tag{28}$$

The necessary optimality conditions state that there exists an absolutely continuous function

$\nu : [0, \ell^0]$ such that, for all ℓ, either

$$\phi(\ell) := \Delta\lambda_g e^{-rt(\ell)} \ell^{\frac{1}{\Delta}} \left(\rho\left(1 - \frac{k}{\ell}\right) + 1\right) + \nu(\ell) = 0, \tag{29}$$

or else $u(\ell) = \frac{1}{\rho+\bar{\alpha}(\ell)}$ if $\phi(\ell) > 0$ and $u(\ell) = \frac{1}{\rho}$ if $\phi(\ell) < 0$.

Furthermore,

$$\nu'(\ell) = -\frac{\partial\mathcal{H}(t, u, \ell, \nu)}{\partial t} = re^{-rt(\ell)} \ell^{\frac{1}{\Delta}-1} \left(\left(1 - \frac{k}{\ell}\right)(1 - \rho u(\ell)) - u(\ell)\right) \ (\ell - \text{a.e.}). \tag{30}$$

Finally, transversality at $\ell = 0$ implies that $\nu(0) = 0$ (since $t(\ell)$ is free) .

Note that

$$\begin{aligned} \phi'(\ell) &= -rt'(\ell)\Delta\lambda_g e^{-rt(\ell)} \ell^{\frac{1}{\Delta}} \left(\rho\left(1 - \frac{k}{\ell}\right) + 1\right) \\ &+ \lambda_g e^{-rt(\ell)} \ell^{\frac{1}{\Delta}-1} \left(\rho\left(1 - \frac{k}{\ell}\right) + 1\right) + \rho k \Delta\lambda_g e^{-rt(\ell)} \ell^{\frac{1}{\Delta}-2} + \nu'(\ell), \end{aligned}$$

or using the formulas for t' and ν',

$$\phi'(\ell) = e^{-rt(\ell)} \ell^{\frac{1}{\Delta}-2} \left(r\left(\ell - k\right) + \rho\Delta\lambda_g k + \lambda_g\left(\rho\left(\ell - k\right) + \ell\right)\right), \tag{31}$$

so ϕ cannot be identically zero over some interval, as there is at most one value of ℓ for which $\phi'(\ell) = 0$. Every solution must be "bang-bang." Specifically,

$$\phi'(\ell) \gtreqless 0 \Leftrightarrow \ell \gtreqless \tilde{\ell} := \left(1 - \frac{\lambda_g(1 + \rho\Delta)}{r + \lambda_g(1 + \rho)}\right) k > 0.$$

Also, $\phi(0) \leq 0$ (specifically, $\phi(0) = 0$ for $\Delta < 1$ and $\phi(0) = -\Delta\lambda_g e^{-rt(\ell)}\rho k$ for $\Delta = 1$). So $\phi(\ell) < 0$ for all $0 < \ell < \ell_g^*$, for some threshold $\ell_g^* > 0$, and $\phi(\ell) > 0$ for $\ell > \ell_g^*$. The constraint $u(\ell) \in \mathcal{U}^i(\ell)$ must bind for all $\ell \in [0, \ell^*)$ (a.e.), and every optimal policy must switch from $u(\ell) = 1/\rho$ for $\ell < \ell_g^*$ to $1/(\rho + \bar{\alpha}(\ell))$ in the second-best problem and to $1/(\rho + 1)$ in the first-best problem for $\ell > \ell_g^*$. It remains to determine the switching point ℓ_g^* (and establish uniqueness in the process).

For $\ell < \ell_g^*$,

$$\nu'(\ell) = -\frac{r}{\rho} e^{-rt(\ell)} \ell^{\frac{1}{\Delta}-1}, \quad t'(\ell) = -\frac{1}{\rho\Delta\lambda_g\ell}$$

so that

$$t(\ell) = C_0 - \frac{1}{\rho\Delta\lambda_g} \ln \ell, \text{ or } e^{-rt(\ell)} = C_1 \ell^{\frac{r}{\rho\Delta\lambda_g}}$$

for some constants $C_1, C_0 = -\frac{1}{r}\ln C_1$. Note that $C_1 > 0$; or else $C_1 = 0$ and $t(\ell) = \infty$ for every $\ell \in (0, \ell_g^*)$, which is inconsistent with $t(\ell_g^*) < \infty$. Hence,

$$\nu'(\ell) = -\frac{r}{\rho} C_1 \ell^{\frac{r}{\rho\Delta\lambda_g} + \frac{1}{\Delta} - 1},$$

and so (using $\nu(0) = 0$),

$$\nu(\ell) = -\frac{r\Delta\lambda_g}{r + \rho\lambda_g} C_1 \ell^{\frac{r}{\rho\Delta\lambda_g} + \frac{1}{\Delta}},$$

for $\ell < \ell_g^*$. We now substitute ν into ϕ, for $\ell < \ell_g^*$, to obtain

$$\phi(\ell) = \Delta\lambda_g C_1 \ell^{\frac{r}{\rho\Delta\lambda_g}} \ell^{\frac{1}{\Delta}} \left(\rho\left(1 - \frac{k}{\ell}\right) + 1\right) - \frac{r\Delta\lambda_g}{r + \rho\lambda_g} C_1 \ell^{\frac{r}{\rho\Delta\lambda_g} + \frac{1}{\Delta}}.$$

We now see that the switching point is uniquely determined by $\phi(\ell) = 0$, as ϕ is continuous and C_1 cancels. Simplifying,

$$\frac{k}{\ell_g^*} = 1 + \frac{\lambda_g}{r + \rho\lambda_g},$$

which leads to the formula for p_g^* in the Proposition (via $\ell = p/(1-p)$ and $k = c/(1-c)$). We have identified the unique solution to the program for both first- and second-best, and shown in the process that the optimal threshold p^* applies to both problems.

The second-best implements the first-best if $p_0 \geq c$, since then $\bar{\alpha}(\ell) = 1$ for all $\ell \leq \ell_0$. If not, then $\bar{\alpha}(\ell) < 1$ for a positive measure of $\ell \leq \ell_0$. Hence, the second-best implements a lower and thus a slower experimentation than does the first-best.

As for sufficiency, we use Arrow sufficiency theorem (Seierstad and Sydsæter, 1987, Theorem 5, p.107). This amounts to showing that the maximized Hamiltonian $\hat{\mathcal{H}}(t, \ell, \nu(\ell)) = \max_{u \in \mathcal{U}^i(\ell)} \mathcal{H}(t, u, \ell, \nu(\ell))$ is concave in t (the state variable), for all ℓ. To this end, it suffices to show that the terms inside the big parentheses in (28) are negative for all $u \in \mathcal{U}^i$, $i = FB, SB$. This is indeed the case:

$$\begin{aligned}
&1 - \frac{k}{\ell} - \left(\rho\left(1 - \frac{k}{\ell}\right) + 1\right) u(\ell) \\
\leq &1 - \frac{k}{\ell} - \min\left\{\left(\rho\left(1 - \frac{k}{\ell}\right) + 1\right)\frac{1}{1+\rho}, \left(\rho\left(1 - \frac{k}{\ell}\right) + 1\right)\frac{1}{\rho}\right\} \\
= &-\min\left\{\frac{k}{(1+\rho)\ell}, \frac{1}{\rho}\right\} < 0,
\end{aligned}$$

where the inequality follows from the linearity of the expression in $u(\ell)$ and the fact that $u(\ell) \in \mathcal{U}^i \subset [\frac{1}{\rho+1}, \frac{1}{\rho}]$, for $i = FB, SB$. The concavity of maximized Hamiltonian in t, and thus sufficiency of our candidate optimal solution, then follows. $\square$

2.2 "Bad news" environment: $\Delta < 0$

The analysis is qualitatively the same for the general bad news case. The same change of variable produces the following program for the designer: For problem $i = SB, FB$,

$$\sup_u \int_{\ell^0}^{\infty} e^{-rt(\ell)} \ell^{\frac{1}{\Delta}-1} \left(\left(1 - \frac{k}{\ell}\right) (1 - \rho u(\ell)) - u(\ell) \right) \mathrm{d}\ell,$$

s.t. $t(\ell^0) = 0$,

$$t'(\ell) = -\frac{u(\ell)}{\Delta \lambda_g \ell},$$

$$u(\ell) \in \mathcal{U}^i(\ell),$$

where as before $\mathcal{U}^{SB}(\ell) := [\frac{1}{\rho+\alpha(\ell)}, \frac{1}{\rho}]$ and $\mathcal{U}^{FB}(\ell) := [\frac{1}{\rho+1}, \frac{1}{\rho}]$. Again, a solution exists Filippov-Cesari theorem (Cesari, 1983).

Proposition 2. *The first-best policy (absent any news) prescribes no experimentation until the posterior p rises to p_b^{**}, and then full experimentation at the rate of $\alpha(p) = 1$ thereafter, for $p > p_b^{**}$, where*

$$p_b^{**} := c\left(1 - \frac{rv}{\rho + r(v + \frac{1}{\lambda_b})}\right).$$

The second-best policy implements the first-best if $p_0 \geq c$ or if $p_0 \leq \hat{p}_0$ for some $\hat{p}_0 < p_b^{**}$. If $p_0 \in (\hat{p}_0, c)$, then the second-best policy prescribes no experimentation until the posterior p rises to p_b^*, and then maximal experimentation at the rate of $\bar{\alpha}(\frac{p}{1-p})$ thereafter for any $p > p_b^*$, where $p_b^* > p_b^{**}$. In other words, the second-best policy triggers experimentation at a later date and at a lower rate than does the first-best.

Proof. As before, the necessary conditions for the second-best policy now state that there exists an absolutely continuous function $\nu : [0, \ell^0]$ such that, for all ℓ, either

$$\psi(\ell) := -\phi(\ell) = \Delta\lambda_g e^{-rt(\ell)} \ell^{\frac{1}{\Delta}} \left(\rho\left(1 - \frac{k}{\ell}\right) + 1\right) - \nu(\ell) = 0, \tag{32}$$

or else $u(\ell) = \frac{1}{\rho+\alpha(\ell)}$ if $\psi(\ell) > 0$ and $u(\ell) = \frac{1}{\rho}$ if $\psi(\ell) < 0$. The formula for $\nu'(\ell)$ is the same as before, given by (30). Finally, transversality at $\ell = \infty$ ($t(\ell)$ is free) implies that $\lim_{\ell\to\infty} \nu(\ell) = 0$.

Since $\psi(\ell) = -\phi(\ell)$, we get from (32) that

$$\psi'(\ell) = -e^{-rt(\ell)} \ell^{\frac{1}{\Delta}-2} \left(r\left(\ell - k\right) + \rho\Delta\lambda_g k + \lambda_g\left(\rho\left(\ell - k\right) + \ell\right)\right).$$

Letting $\tilde{\ell} := \left(1 - \frac{\lambda_g(1+\rho\Delta)}{r+\lambda_g(1+\rho)}\right)k$, namely the solution to $\psi(\ell) = 0$. Then, ψ is maximized at

$\tilde{\ell}$, and is strictly quasi-concave. Since $\lim_{\ell\to\infty} h(\ell) = 0$, this means that there must be a cutoff $\ell_b^* < \tilde{\ell}$ such that $\psi(\ell) < 0$ for $\ell < \ell_b^*$ and $\psi(\ell) > 0$ for $\ell > \ell_b^*$. Hence, the solution is bang-bang, with $u(\ell) = 1/\rho$ if $\ell < \ell_b^*$, and $u(\ell) = 1/(\rho + \alpha(\ell))$ if $\ell > \ell_b^*$.

The first-best policy has the same cutoff structure, except that the cutoff may be different from ℓ_b^*. Let ℓ_b^{**} denote the first-best cutoff.

First-best policy: We shall first consider the first best policy. In that case, for $\ell > \ell_b^{**}$,

$$t'(\ell) = -\frac{1}{\Delta\lambda_g(1+\rho)\ell}$$

gives

$$e^{-rt(\ell)} = C_2 \ell^{\frac{r}{(1+\rho)\Delta\lambda_g}},$$

for some non-zero constant C_2. Then

$$\nu'(\ell) = -\frac{rk}{1+\rho} C_2 \ell^{\frac{r}{(1+\rho)\Delta\lambda_g}+\frac{1}{\Delta}-2}$$

and $\lim_{\ell\to\infty}\nu(\ell) = 0$ give

$$\nu(\ell) = -\frac{rk\Delta\lambda_g}{r+(1+\rho)(1-\Delta)\lambda_g} C_2 \ell^{\frac{r}{(1+\rho)\Delta\lambda_g}+\frac{1}{\Delta}-1}.$$

So we get, for $\ell > \ell_b^{**}$,

$$\psi(\ell) = -\Delta\lambda_g C_2 \ell^{\frac{r}{(1+\rho)\Delta\lambda_g}} \ell^{\frac{1}{\Delta}-1} \left(\ell(1+\rho) - k\rho\right) + \frac{rk\Delta\lambda_g}{r+(1+\rho)(1-\Delta)\lambda_g} C_2 \ell^{\frac{r}{(1+\rho)\Delta\lambda_g}+\frac{1}{\Delta}-1}.$$

Setting $\psi(\ell_b^{**}) = 0$ gives

$$\frac{k}{\ell_b^{**}} = \frac{r+(1+\rho)(1-\Delta)\lambda_g}{r+\rho(1-\Delta)\lambda_g} = \frac{r+(1+\rho)\lambda_b}{r+\rho\lambda_b} = 1 + \frac{\lambda_b}{r+\rho\lambda_b},$$

or

$$p_b^{**} = c\left(1 - \frac{rv}{\rho + r(v + \frac{1}{(1-\Delta)\lambda_g})}\right) = c\left(1 - \frac{rv}{\rho + r(v+\frac{1}{\lambda_b})}\right).$$

Second-best policy. We now characterize the second-best cutoff. There are two cases, depending upon whether $\alpha(\ell) = 1$ is incentive-feasible at the threshold ℓ_b^{**} that characterizes the first-best policy. In other words, for the first-best to be implementable, we should have

$\bar{\alpha}(\ell^{**}) = 1$, which requires

$$\ell_0 \geq k \left(\frac{r + \rho\lambda_b}{r + (1+\rho)\lambda_b} \right)^{1-\Delta} =: \hat{\ell}_0.$$

Observe that since $\Delta < 0$, $\hat{\ell}_0 < \ell^{**}$. If $\ell_0 \leq \hat{\ell}_0$, then the designer begins with no experimentation and waits until the posterior belief improves sufficiently to reach ℓ^{**}, at which point the agents will be asked to experiment with full force, *i.e.*, with $\bar{\alpha}(\ell) = 1$, that is, given that no news has arrived by that time. This first-best policy is implementable since, given the sufficiently favorable prior, the designer will have built sufficient "credibility" by that time. Hence, unlike the case of $\Delta > 0$, the first best can be implementable even when $\ell_0 < k$.

Suppose $\ell_0 < \hat{\ell}_0$. Then, the first-best is not implementable. That is, $\bar{\alpha}(\ell_b^{**}) < 1$. Let ℓ_b^* denote the threshold at which the constrained designer switches to $\bar{\alpha}(\ell)$. We now prove that $\ell_b^* > \ell_b^{**}$.

For the sake of contradiction, suppose that $\ell_b^* \leq \ell_b^{**}$. Note that $\psi(x) = \lim_{\ell\to\infty} \phi(\ell) = 0$. This means that

$$\int_{\ell_b^*}^{\infty} \psi'(\ell)\mathrm{d}\ell = \int_{\ell_b^*}^{\infty} e^{-rt(\ell)} \ell^{\frac{1}{\Delta}-2} \left((r + \lambda_b\rho)k - (r + \lambda_g(\rho+1))\ell \right) \mathrm{d}\ell = 0,$$

where $\psi'(\ell) = -\phi'(\ell)$ is derived using the formula in (32).

Let t^{**} denote the time at which ℓ_b^{**} is reached along the first-best path. Let

$$f(\ell) := \ell^{\frac{1}{\Delta}-2} \left((r + \lambda_b\rho)k - (r + \lambda_g(\rho+1))\ell \right).$$

We then have

$$\int_{\ell_b^*}^{\infty} e^{-rt^{**}(\ell)} f(\ell)\mathrm{d}\ell \geq 0, \tag{33}$$

(because $\ell_b^* \leq \ell_b^{**}$; note that $f(\ell) \leq 0$ if and only if $\ell > \tilde{\ell}$, so h must tend to 0 as $\ell \to \infty$ from above), yet

$$\int_{\ell_b^*}^{\infty} e^{-rt(\ell)} f(\ell)\mathrm{d}\ell = 0. \tag{34}$$

Multiplying $e^{rt^{**}(\tilde{\ell})}$ on both sides of (33) gives

$$\int_{\ell_b^*} e^{-r(t^{**}(\ell) - t^{**}(\tilde{\ell}))} f(\ell)\mathrm{d}\ell \geq 0. \tag{35}$$

Likewise, multiplying $e^{rt(\tilde{\ell})}$ on both sides of (34) gives

$$\int_{\ell_b^*}^{\infty} e^{-r(t(\ell)-t(\tilde{\ell}))} f(\ell)\mathrm{d}\ell = 0. \tag{36}$$

Subtracting (35) from (36) gives

$$\int_{\ell_b^*} \left(e^{-r(t(\ell)-t(\tilde{\ell}))} - e^{-r(t^{**}(\ell)-t^{**}(\tilde{\ell}))}\right) f(\ell)\mathrm{d}\ell \leq 0. \tag{37}$$

Note $t'(\ell) \geq (t^{**})'(\ell) > 0$ for all ℓ, with strict inequality for a positive measure of ℓ. This means that $e^{-r(t(\ell)-t(\tilde{\ell}))} \leq e^{-r(t^{**}(\ell)-t^{**}(\tilde{\ell}))}$ if $\ell > \tilde{\ell}$, and $e^{-r(t(\ell)-t(\tilde{\ell}))} \geq e^{-r(t^{**}(\ell)-t^{**}(\tilde{\ell}))}$ if $\ell < \tilde{\ell}$, again with strict inequality for a positive measure of ℓ for $\ell \geq \ell_b^{**}$ (due to the fact that the first best is not implementable; *i.e.*, $\bar{\alpha}(\ell_b^{**}) < 1$). Since $f(\ell) < 0$ if $\ell > \tilde{\ell}$ and $f(\ell) > 0$ if $\ell < \tilde{\ell}$, we have a contradiction to (37).

For sufficiency, the same argument as with Proposition 1 establishes that the maximized Hamiltonian will necessarily be concave in t, which implies optimality of our candidate solution, by Arrow's sufficiency theorem. □

2.3 "Neutral news" environment: $\Delta = 0$

In this case, the designer's posterior on the quality of the good remains unchanged in the absence of breakthrough news. Experimentation could be still desirable for the designer. If $p_0 \geq c$, then the agents will voluntarily consume the good, so experimentation is clearly self-enforcing. If $p_0 < c$, then the agents will not voluntarily consume, so spamming is needed to incentivize experimentation. As before the optimal policy has the familiar cutoff structure.

Proposition 3. *The second-best policy prescribes, absent any news, the maximal experimentation at $\bar{\alpha}_t$ if $p_0 \geq p_0^*$, and no experimentation if $p_0 < p_0^*$, where $p_0^* := p_g^*(= p_b^{**})$ and $\bar{\alpha}_t$ (given in the Appendix) is increasing and convex in t and reaches 1 when $t^* = \frac{k-\ell_0}{\lambda_b(\ell_0-k\rho)} \ln \frac{\ell_0}{k\rho}$. The first-best policy has the same structure with the same threshold posterior, except that $\bar{\alpha}_t$ is replaced by 1. The first-best is implementable if and only if $p_0 \geq c$ or $p_0 < p_0^*$.*

Proof. In that case, $\ell = \ell_0$. The objective rewrites

$$\begin{aligned}
\mathcal{W} &= \int_{t\geq 0} e^{-rt} \left(g_t(1-c) + \frac{p_0 - c}{p_0}\alpha_t(p_0 - g_t)\right) \mathrm{d}t \\
&= \int_{t\geq 0} e^{-rt} \left(g_t(1-c) + \frac{p_0 - c}{p_0}\left(\frac{\dot{g}_t}{\lambda_g} - (p_0 - g_t)\rho\right)\right) \mathrm{d}t
\end{aligned}$$

$$
\begin{aligned}
&= \int_{t\geq 0} e^{-rt}\left(g_t(1-c)+\frac{p_0-c}{p_0}\left(r\frac{g_t}{\lambda_g}-(p_0-g_t)\rho\right)\right)\mathrm{d}t+\text{Const.} \quad \text{(Integr. by parts)}\\
&= \int_{t\geq 0} e^{-rt}g_t\left(1-c+\frac{p_0-c}{p_0}\left(\frac{r}{\lambda_g}+\rho\right)\right)\mathrm{d}t+\text{Const.}\\
&= \text{Const.}\times\int_{t\geq 0} e^{-rt}g_t\left((\ell_0-k)(r+\lambda_g\rho)+\lambda_g\ell_0\right)\mathrm{d}t+\text{Const.},
\end{aligned}
$$

and so we see that it is best to set g_t to its maximum or minimum value depending on the sign of $(\ell_0-k)(r+\lambda_g\rho)+\lambda_g\ell_0$, specifically, depending on

$$
\frac{k}{\ell_0}\lesseqgtr 1+\frac{\lambda_g}{r+\lambda_g\rho},
$$

which is the relationship that defines $\ell_g^*=\ell_b^{**}$. Now, g_t is maximized by setting $\alpha_\tau=\bar{\alpha}_\tau$ and minimized by setting $\alpha_\tau=0$ (for all $\tau<t$).

We can solve for $\bar{\alpha}_t$ from the incentive compatibility constraint, plug back into the differential equation for g_t and get, by solving the ordinary differential equation,

$$
g_t=\frac{\left(e^{\frac{\lambda_g(\ell_0-k\rho)t}{k-\ell_0}}-1\right)\ell_0(k-\ell_0)\rho}{(1+\ell_0)(\ell_0-k\rho)},
$$

and finally

$$
\bar{\alpha}_t=\frac{\ell_0}{\frac{\rho k-\ell_0}{\rho\left(1-e^{\frac{\lambda_g(\ell_0-\rho k)t}{k-\ell_0}}\right)}-(k-\ell_0)},
$$

which is increasing in t and convex in t (for $\gamma>l^0$) and equal to 1 when

$$
\lambda_g t^*=\frac{k-\ell_0}{\ell_0-k\rho}\ln\frac{\ell_0}{k\rho}.
$$

The optimal policy in that case is fairly obvious: experiment at maximum rate until t^*, at rate 1 from that point on (conditional on no feedback). ☐

3 Heterogeneous Costs: Proofs from Section 7.2

3.1 Proof of Proposition 4

The objective function reads

$$\int_{t\geq 0} e^{-rt}\left(g_t(1-\bar{c}) + (1-g_t-b_t)(q_H\alpha_H(p_t-c_L) + q_L\alpha_L(p_t-c_L)\right)\mathrm{d}t,$$

where $\bar{c} := q_H c_H + q_L c_L$. Substituting for g_t, b_t and re-arranging, this gives

$$\int_{t\geq 0} e^{-rt}\ell(t)\left(\alpha_H(t)q_H\left(1-c_H\left(1+\frac{1}{\ell(t)}\right)\right) + \alpha_L(t)q_L\left(1-c_L\left(1+\frac{1}{\ell(t)}\right)\right) - (1-\bar{c})\right)\mathrm{d}t.$$

As before, it is more convenient to work with $t(\ell)$ as the state variable, and doing the change of variables gives

$$\int_0^{\ell_0} e^{-rt(\ell)}\left(x_H(\ell)u_H(\ell) + x_L(\ell)u_L(\ell) - \frac{1-\bar{c}}{\rho}\right)\mathrm{d}\ell,$$

where for $j = L, H$, $x_j(\ell) := 1 - c_j\left(1+\frac{1}{\ell}\right) + \frac{1-\bar{c}}{\rho}$, and $u_j(\ell) := \frac{q_j\alpha_j(t(\ell))}{\rho + q_L\alpha_L(t(\ell)) + q_H\alpha_H(t(\ell))}$ are the control variables that take values in the sets $\mathcal{U}^j(\ell) = [\underline{u}_k, \bar{u}_k]$ (whose definition depends on first- vs. second-best). This is to be maximized subject to

$$t'(\ell) = \frac{u_H(\ell) + u_L(\ell) - 1}{\rho\lambda\ell}.$$

As before, we invoke Pontryagin's principle. There exists an absolutely continuous function $\eta : [0, \ell_0] \to \mathbf{R}$, such that, a.e.,

$$\eta'(\ell) = re^{-rt(\ell)}\left(x_H(\ell)u_H(\ell) + x_L(\ell)u_L(\ell) - \frac{1-\bar{c}}{\rho}\right),$$

and u_j is maximum or minimum, depending on the sign of

$$\phi_j(\ell) := \rho\lambda\ell e^{-rt(\ell)}x_j(\ell) + \eta(\ell).$$

This is because this expression cannot be zero except for a specific value of $\ell = \ell_j$. Namely, note first that, because $x_H(\ell) < x_L(\ell)$ for all ℓ, at least one of $u_L(\ell), u_H(\ell)$ must be extremal, for all ℓ. Second, upon differentiation,

$$\phi_H'(\ell) = e^{-rt(\ell)}\left(\left(\lambda - \frac{r}{\rho}\right)(1-\bar{c}) + \rho\lambda(1-c_H) + ru_L(\ell)(c_H-c_L)\left(1+\frac{1}{\ell}\right)\right)$$

implies that, if $\phi_H(\ell) = 0$ were identically zero over some interval, then $u_L(\ell)$ would be extremal over this range, yielding a contradiction, as the right-hand side cannot be zero identically, for $u_L(\ell) = \bar{u}_L(\ell)$. Similar reasoning applies to $u_L(\ell)$, considering $\phi'_L(\ell)$. Hence, the optimal policy is characterized by two thresholds, ℓ_H, ℓ_L, with $\ell_0 \geq \ell_H \geq \ell_L \geq 0$, such that both types of regular consumers are asked to experiment whenever $\ell \in [\ell_H, \ell_0]$, low-cost consumers are asked to do so whenever $\ell \in [\ell_L, \ell_0]$, and neither is asked to otherwise.

We now characterize the threshold beliefs under first-best and second-best policies. Throughout, we shall use superscript $**$ to denote the first-best and superscript $*$ to denote the second-best policy. By the principle of optimality, the threshold ℓ_L must coincide with $\ell^* = \ell^{**}$ in the case of only one type of regular consumers (with cost c_L). To compare ℓ^*_H and ℓ^{**}_H, we proceed as in the bad news case, by noting that, in either case,

$$\phi_H(\ell_H) = 0,$$

and

$$\phi_H(\ell_L) = \phi_L(\ell_L) + \rho\lambda\ell_L e^{-rt(\ell_L)}(x_H(\ell_L) - x_L(\ell_L)) = -\rho\lambda e^{-rt(\ell_L)}(c_H - c_L)\,(1 + \ell_L)\,.$$

Hence,

$$\int_{\ell_L}^{\ell_H} e^{rt(\ell_L)}\phi'_H(\ell)\mathrm{d}\ell = \rho\lambda(c_H - c_L)\,(1 + \ell_L)$$

holds both for the first- and second-best. Note now that, in the range $[\ell_L, \ell_H]$,

$$e^{rt(\ell_L)}\phi'_H(\ell) = e^{-r\int_{\ell_H}^{\ell}\frac{u_L(l)+u_H(l)-1}{\rho\lambda l}\mathrm{d}l}\left(\left(\lambda - \frac{r}{\rho}\right)(1-\bar{c}) + \rho\lambda(1-c_H) + ru_L(\ell)(c_H - c_L)\left(1 + \frac{1}{\ell}\right)\right).$$

Because $\bar{\alpha}_L(\ell) > \bar{\alpha}_H(\ell)$, $\bar{u}^*_L(\ell) > \bar{u}^{**}_L(\ell)$, and also $\bar{u}^{**}_L(\ell) + \bar{u}^{**}_H(\ell) \geq \bar{u}^*_L(\ell) + \bar{u}^*_H(\ell)$, so that, for all ℓ in the relevant range,

$$e^{rt(\ell_L)}\frac{\mathrm{d}\phi^{**}_H(\ell)}{\mathrm{d}\ell} < e^{rt(\ell_L)}\frac{\mathrm{d}\phi^*_H(\ell)}{\mathrm{d}\ell},$$

and it then follows that $\ell^*_H < \ell^{**}_H$. □

3.2 Uniform Cost: Derivation of the Optimum

We characterize the recommendation policy as $r \to 0$. To derive this policy, let us first describe the designer's payoff. This is his payoff *in expectation*. Her objective is

$$\int_0^{t_1} e^{-rt}\left[\int_0^{\frac{\ell_0}{1+\ell_0}} \frac{\ell_0-\ell_t}{1+\ell_0}(1-c)\,\mathrm{d}c + \int_{\frac{k_t}{1+k_t}}^{\bar{c}} \frac{\ell_0-\ell_t}{1+\ell_0}(1-c)\,\mathrm{d}c + \int_0^{\frac{\ell_0}{1+\ell_0}} \frac{1+\ell_t}{1+\ell_0}\left(\frac{\ell_t}{1+\ell_t}-c\right)\mathrm{d}c\right]\mathrm{d}t$$
$$+\int_{t_1}^{\infty} e^{-rt}\left[\int_0^{\bar{c}} \frac{\ell_0-\ell_t}{1+\ell_0}(1-c)\,\mathrm{d}c + \int_0^{\frac{k_t}{1+k_t}} \frac{1+\ell_t}{1+\ell_0}\left(\frac{\ell_t}{1+\ell_t}-c\right)\mathrm{d}c\right]\mathrm{d}t.$$

To understand this expression, consider $t < t_1$. Types in $t \in (\ell_0, k_t)$ derive no surplus, because they are indifferent between buying or not (what they gain from being recommended to buy when the good has turned out to be good is exactly offset by the cost of doing so when this is myopically suboptimal). Hence, their contribution to the expected payoff cancels out (but it does not mean that they are disregarded, because their behavior affects the amount of experimentation.) Types above k_t get recommended to buy only if the good has turned out to be good, in which case they get a flow surplus of $\lambda \cdot 1 - c = 1 - c$. Types below ℓ_0 have to purchase for both possible posterior beliefs, and while the flow revenue is 1 in one case, it is only $p_t = \ell_t/(1+\ell_t)$ in the other case.

The payoff in case $t \geq t_1$ can be understood similarly. There are no longer indifferent types. In case of an earlier success, all types enjoy their flow payoff $1-c$, while in case of no success, types below γ_t still get their flow $p_t - c$.

This expression can be simplified to

$$\begin{aligned}
J(k) &= \int_0^{\infty} e^{-rt}\left[\int_0^{\frac{\ell_0}{1+\ell_0}\wedge\frac{k_t}{1+\gamma_t}} \left(\frac{\ell_0}{1+\ell_0}-c\right)\mathrm{d}c + \int_{\frac{\gamma_t}{1+\gamma_t}}^{\frac{\bar{k}}{1+\bar{k}}} \frac{\ell_0-\ell_t}{1+\ell_0}(1-c)\,\mathrm{d}c\right]\mathrm{d}t \\
&= \int_0^{\infty} e^{-rt}\left[\frac{\ell_0}{1+\ell_0}\left(\frac{\ell_0}{1+\ell_0}\wedge\frac{\gamma_t}{1+\gamma_t}\right) - \frac{1}{2}\left(\frac{\ell_0}{1+\ell_0}\wedge\frac{\bar{k}_t}{1+\gamma_t}\right)^2\right]\mathrm{d}t \\
&+ \int_0^{\infty} e^{-rt}\left[\frac{\ell_0-\ell_t}{1+\ell_0}\left(\frac{\bar{k}}{1+\bar{k}} - \frac{k_t}{1+k_t} - \frac{1}{2}\left(\left(\frac{\bar{k}}{1+\bar{k}}\right)^2 - \left(\frac{\gamma_t}{1+\gamma_t}\right)^2\right)\right)\right]\mathrm{d}t,
\end{aligned}$$

with the obvious interpretation. For $t \geq t_1$,

$$\dot{\ell}_t = -\ell_t \int_0^{\frac{k_t}{1+k_t}} \frac{\mathrm{d}c}{\bar{c}} = -\frac{\ell_t}{\bar{c}}\frac{k_t}{1+k_t},$$

while for $t \leq t_1$, it holds that

$$\begin{aligned}\dot{\ell}_t &= -\ell_t\left(\frac{p_0}{\bar{c}} + \int_{p_0}^{\frac{\gamma_t}{1+\gamma_t}} \alpha_t(k)\frac{\mathrm{d}c}{\bar{c}}\right) = -\frac{\ell_t}{\bar{c}}\left(p_0 + \int_{p_0}^{\frac{\gamma_t}{1+\gamma_t}} \frac{\ell_0 - \ell_t}{k(c) - \ell_t}\mathrm{d}c\right) \\ &= -\frac{\ell_t}{\bar{c}}\left[\frac{k_t\ell_t + \ell_0}{(1+\gamma_t)(1+\ell_0)} - \frac{\ell_0 - \ell_t}{(1+\ell_t)^2}\ln\frac{(1+k_t)(\ell_0-\ell_t)}{(1+\ell_0)(k_t-\ell_t)}\right].\end{aligned}$$

Finally, note that the value of k_0 is free.

To solve this problem, we apply Pontryagin's maximum principle. Consider first the case $t \geq t_1$. The Hamiltonian is then

$$\mathcal{H}(\ell,\gamma,\mu,t) = \frac{e^{-rt}}{2(1+k_t)^2}\left(2k_t(1+k_t)\frac{\ell_0}{1+\ell_0} - k_t^2 + \frac{(\bar{k}-k_t)(2+\gamma_t+\bar{k})(\ell_0-\ell_t)}{(1+\bar{k})^2(1+\ell_0)}\right) - \mu_t\ell_t\frac{\gamma_t(1+\bar{k})}{(1+\gamma_t)\bar{k}},$$

where μ is the co-state variable. The maximum principle gives, taking derivatives with respect to the control γ_t,

$$\mu_t = -e^{-rt}\bar{k}\frac{k_t - \ell_t}{(1+\bar{k})(1+k_t)(1+\ell_0)\ell_t}.$$

The adjoint equation states that

$$\dot{\mu} = -\frac{\partial\mathcal{H}}{\partial\ell} = \frac{e^{-rt}}{2(1+\bar{k})^2(1+\ell_0)(1+k_t)^2\ell_t}\left(k_t^2(2(1+\bar{k})^2+\ell_t) - \bar{k}(2+\bar{k})(2k_t+1)\ell_t\right),$$

after inserting the value for μ_t. Differentiate the formula for μ, combine to get a differential equation for k_t. Letting $r \to 0$, and changing variables to $k(\ell)$, we finally obtain

$$2(1+\bar{k})^2\frac{(1+\ell)\gamma(\ell)}{1+\gamma(\ell)}\gamma'(\ell) = \bar{k}(2+\bar{k})(1+2\gamma(\ell)) - \gamma(\ell)^2.$$

Along with $k(0) = 0$, $k > 0$ we get

$$\gamma(\ell) = \frac{\bar{k}(2+\bar{k})\ell + (1+\bar{k})\sqrt{\bar{k}(2+\bar{k})\ell(1+\ell)}}{(1+\bar{k})^2 + \ell}.$$

This gives us, in particular, $\gamma(l_0)$. Note that, in terms of cost c, this gives

$$c(\ell) = \frac{\sqrt{\bar{k}(2+\bar{k})\ell/(1+\ell)}}{1+\bar{k}},$$

We now turn to the Hamiltonian for the case $t \leq t_1$, or $\gamma_t \geq \ell_0$. It might be that the solution is a "corner" solution, that is, all agents experiment ($\gamma_t = \bar{k}$). Hence, we abuse notation, and solve for the unconstrained solution γ: the actual solution should be set at $\min\{\bar{k}, \gamma_t\}$. Proceeding in the same fashion, we get again

$$\mu_t = -e^{-rt}\bar{k}\frac{\gamma_t - \ell_t}{(1+\bar{k})(1+\gamma_t)(1+\ell_0)\ell_t},$$

and continuity of μ (which follows from the maximum principle) is thus equivalent to the values of $\gamma(\ell)$ obtained from both cases matching at $\ell = \ell_0$. The resulting differential equation for $\gamma(\ell)$ admits no closed-form solution. It is given by

$$\begin{aligned} & \left(4k_0\left(k_0+2\right)+\ell(\ell+2)+5\right)k(\ell)^2 - k_0\left(k_0+2\right)\left((\ell+1)^2+4\ell_0\right) - 4\ell_0 \\ - \; & 2\left(k_0\left(k_0+2\right)\left(\ell(\ell+2)+2\ell_0-1\right)+2\left(\ell_0-1\right)\right)k(\ell) \\ = \; & 2\frac{\left(k_0+1\right)^2}{1+\ell}\left(k(\ell)+1\right)\left(\left(\ell-2\ell_0-1\right)k(\ell)-\ell_0+\ell\left(\ell_0+2\right)\right)\log\left(\frac{\left(\ell-\ell_0\right)\left(k(\ell)+1\right)}{\left(\ell_0+1\right)\left(\ell-k(\ell)\right)}\right) \\ + \; & 2\left(k_0+1\right)^2(\ell+1)k'(\ell)\left(\left(\ell_0-\ell\right)\log\left(\frac{\left(\ell-\ell_0\right)\left(k(\ell)+1\right)}{\left(\ell_0+1\right)\left(l-k(\ell)\right)}\right) - \frac{\left(\ell+1\right)\left(\ell k(\ell)+\ell_0\right)}{k(\ell)+1}\right). \end{aligned}$$

This suffices to represent the solution, as we have done in Figure 5 in the main body. □

3.3 Proof of Proposition 5

We allow the designer to randomize over finitely many paths of experimentation, so there are finitely many possible posterior beliefs, 1, p^j, $j = 1, \dots, J$. We allow then for multiple (finitely many) recommendations R. So a policy is now a collection $(\alpha_j^R, \gamma_j^R)_j$, depending on the path j that is followed. Along the path j, conditional on the posterior being 1, a recommendation R is given by probability γ_j^R, and conditional on the posterior being p_j, the probabilities α_j^R are used. One last parameter is the probability with which each path j is being used, μ_j.

Correspondingly, there are as many thresholds γ^R as recommendations; namely, given recommendation R, a consumer buys if his cost is no larger than

$$c^R = \frac{\sum_j \mu_j \left(\frac{p_0-p_j}{1-p_j}\beta_j^R + \frac{1-p_0}{1-p_j}p_j\alpha_j^R\right)}{\sum_j \mu_j \left(\frac{p_0-p_j}{1-p_j}\beta_j^R + \frac{1-p_0}{1-p_j}\alpha_j^R\right)},$$

Hence we set

$$\gamma^R = \frac{\sum_j \mu_j \left(\alpha_j^R \ell_j + \beta_j^R(\ell_0 - \ell_j)\right)}{\sum_j \mu_j \alpha_j^R}.$$

We remark for future reference that

$$
\begin{aligned}
\sum_R \gamma^R \sum_j \mu_j \alpha_j^R &= \sum_R \sum_j \mu_j \left(\alpha_j^R \ell_j + \beta_k^R (\ell_0 - \ell_j) \right) \\
&= \sum_j \mu_j \left(\left(\sum_R \alpha_j^R \right) \ell_j + \left(\sum_R \beta_j^R \right) (\ell_0 - \ell_j) \right) \\
&= \sum_j \mu_j \ell_0 = \ell_0.
\end{aligned}
$$

We now turn to the value function. We have that

$$
\begin{aligned}
rV(\ell_1, \ldots, \ell_J) &= \sum_j \mu_j \left(\frac{1+\ell_j}{1+\ell_0} \sum_R \alpha_j^R \int_0^{c^R} (p_j - x) \mathrm{d}x + \frac{\ell_0 - \ell_j}{1+\ell_0} \sum_R \beta_j^R \int_0^{c^R} (1-x) \mathrm{d}x \right) \\
&- \sum_j \ell_j \mu_j \left(\sum_R \alpha_j^R \int_0^{c^R} \mathrm{d}x \right) \frac{\partial V(\ell_1, \ldots, \ell_J)}{\partial \ell_j}.
\end{aligned}
$$

We shall do a few manipulations. First, we work on the flow payoff. From the first to the second equation, we gather terms involving the revenue ("p_j" and 1) on one hand, and cost ("x") on the other. From the second to the third, we use the definition of γ^R (in particular, note that the term in the numerator of γ^R appears in the expressions). The last line uses the remark above.

$$
\begin{aligned}
& \sum_j \mu_j \left(\frac{1+\ell_j}{1+\ell_0} \sum_R \alpha_j^R \int_0^{c^R} \left(\frac{\ell_j}{1+\ell_j} - x \right) \mathrm{d}x + \frac{\ell_0 - \ell_j}{1+l_0} \sum_R \beta_j^R \int_0^{c^R} (1-x) \mathrm{d}x \right) \\
= \; & \frac{1}{1+\ell_0} \sum_R c^R \sum_j \mu_k \left(\ell_j \alpha_j^R + (\ell_0 - \ell_j) \beta_j^R \right) - \frac{1}{2(1+\ell_0)} \sum_R (c^R)^2 \sum_j \mu_j \left((1+\ell_j) \alpha_j^R + (\ell_0 - \ell_j) \beta_j^R \right) \\
= \; & \frac{1}{1+\ell_0} \sum_R \frac{\gamma^R}{1+\gamma^R} \left(\gamma^R \sum_j \mu_j \alpha_j^R \right) - \frac{1}{2(1+\ell_0)} \sum_R \left(\frac{\gamma^R}{1+\gamma^R} \right)^2 \left((1+\gamma_j^R) \sum_j \mu_j \alpha_j^R \right) \\
= \; & \frac{1}{2(1+\ell_0)} \sum_R \frac{(\gamma^R)^2}{1+\gamma^R} \left(\sum_j \mu_j \alpha_j^R \right) \\
= \; & \frac{1}{2(1+\ell_0)} \sum_R \left(\gamma^R - \frac{\gamma^R}{1+\gamma^R} \right) \left(\sum_j \mu_j \alpha_j^R \right) \\
= \; & \frac{1}{2(1+\ell_0)} \sum_R \gamma^R \sum_j \mu_j \alpha_j^R - \frac{1}{2(1+\ell_0)} \sum_R \frac{\gamma^R}{1+\gamma^R} \left(\sum_j \mu_j \alpha_j^R \right) \\
= \; & \frac{\ell_0 - \sum_j \mu_j x_j}{2(1+\ell_0)},
\end{aligned}
$$

where we define

$$x_j := \sum_R \frac{\gamma^R}{1+\gamma^R}\alpha_j^R.$$

Let us now simplify the coefficient of the partial derivative

$$\mu_j \left(\sum_R \alpha_j^R \int_0^{c^R} \mathrm{d}x \right) = \mu_j \sum_R \alpha_j^R \frac{\gamma^R}{1+\gamma^R} = \mu_j x_j.$$

To conclude, given (μ_j) (ultimately, a choice variable as well), the optimality equality simplifies to

$$rV(\ell_1,\ldots,\ell_J) \quad = \quad \frac{\ell_0}{2(1+\ell_0)} - \sum_j \max_{x_j} \mu_j x_j \left\{ \frac{1}{2(1+\ell_0)} + \ell_j \frac{\partial V(\ell_1,\ldots,\ell_J)}{\partial \ell_j} \right\},$$

or letting $W = 2(1+\ell_0)V - \frac{\ell_0}{r}$,

$$rW(\ell_1,\ldots,\ell_J) + \sum_j \mu_j \max_{x_j} x_j \left\{ 1 + \ell_j \frac{\partial W(\ell_1,\ldots,\ell_J)}{\partial \ell_j} \right\} = 0.$$

where $(x_j)_j$ must be feasible, *i.e.*, appropriate values for (α,γ) must exist. This is a tricky restriction, and the resulting set of (x_j) is convex, but not necessarily a polytope. In particular, it is not the product of the possible quantities of experimentation that would obtain if the agents knew which path were followed, $\times_j \left[\frac{\ell_j}{1+\ell_j}, \frac{\ell_0}{1+\ell_0}\right]$. It is a strictly larger set: by blurring recommendation policies, he can obtain pairs of amounts of experimentation outside this set, although not more or less in all dimensions simultaneously.

Let us refer to this set as B_J. This set is of independent interest, as it is the relevant set of possible experimentation schemes independently of the designer's objective function. This set is difficult to compute, as for a given J, we must determine what values of x can be obtained for *some* number of recommendations. Even in the case $J = 2$, this requires substantial effort, and it is not an obvious result that assuming without loss that $\ell_1 \geq \ell_2$, B_2 is the convex hull of the three points

$$x^P := \left(\frac{\sum_j \mu_j \ell_j}{1+\sum_j \mu_j \ell_j}, \frac{\sum_j \mu_j \ell_j}{1+\sum_j \mu_j \ell_j} \right), x^S := \left(\frac{\ell_1}{1+\ell_1}, \frac{\ell_2}{1+\ell_2} \right), x^A := \left(\frac{\ell_0 - \mu_2 \ell_2}{1+\ell_0-\mu_2(1+\ell_2)}, \frac{\ell_2}{1+\ell_2} \right),$$

and the two curves

$$S^U := \left(x_1, 1 + \frac{\mu_2(1-x_1)}{\mu_1 - (1+\ell_0)(1-x_1)} \right),$$

for $x_1 \in \left[\frac{\ell_1}{1+\ell_1}, \frac{\ell_0-\mu_2\ell_2}{1+\ell_0-\mu_2(1+\ell_2)}\right]$, and

$$S^L := \left(x_1, x_1 + \frac{(x_1-(1-x_1)\ell_0)(x_1-(1-x_1)(\mu_1\ell_1+\mu_2\ell_2))}{\mu_2(\mu_1\ell_1+\mu_2\ell_2+\ell_0\ell_2-(1+\ell_0)(1+\ell_2)x_1)}\right),$$

for $x_1 \in \left[\frac{\sum_j \mu_j\ell_j}{1+\sum_j \mu_j\ell_j}, \frac{\ell_1}{1+\ell_1}\right]$, that intersect at the point

$$\left(\frac{\ell_1}{1+\ell_1}, \frac{\ell_0-\mu_1\ell_1}{1+\ell_0-\mu_1(1+\ell_1)}\right).$$

It is worth noting that the point $\left(\frac{\ell_0}{1+\ell_0}, \frac{\ell_0}{1+\ell_0}\right)$ lies on the first (upper) curve, and that the slope of the boundary at this point is $-\mu_1/\mu_2$: hence, this is the point within B_2 that maximizes $\sum_j \mu_j x_j$. See Figure 1 below. To achieve all extreme points, more than two messages are necessary (for instance, achieving x^S requires three messages, corresponding to the three possible posterior beliefs at time t), but it turns out that three suffice.

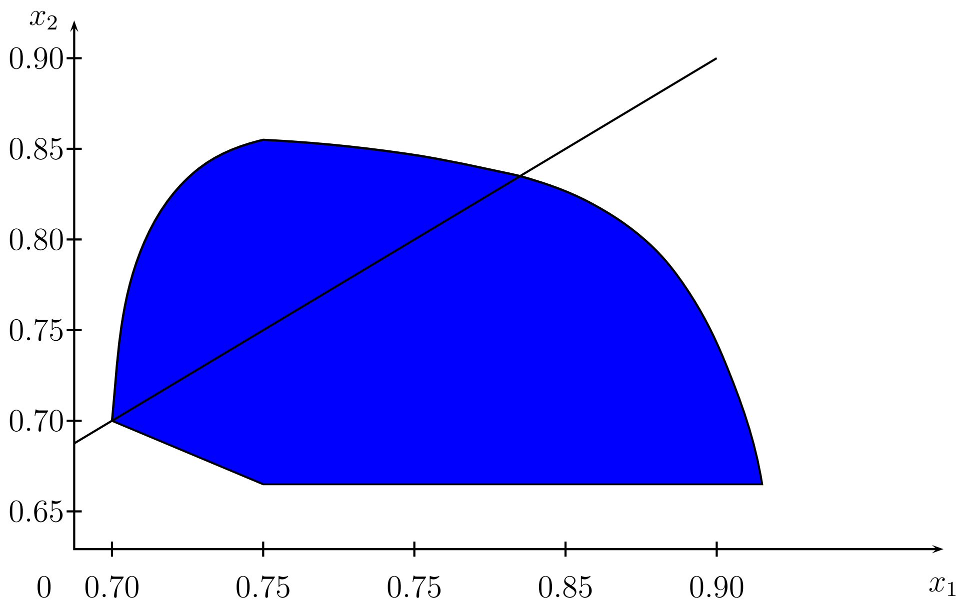

Figure 1: Region B_2 of feasible (x_1, x_2) in the case $J = 2$ (here, for $\ell_0 = 5, \ell_1 = 3, \ell_2 = 2, \mu_1 = 2/3$.

In terms of our notation, the optimum value of a non-randomized strategy is

$$W^S(\ell) = -e^{\frac{r}{\ell}} E_{1+r}\left(\frac{r}{\ell}\right).$$

We claim that the solution to the optimal control problem is given by the "separating" strategy, given μ and $\mathbf{l} = (\ell_1, \dots, \ell_K)$, *for the case* $J = 2$ to begin with. That is,

$$W(\mathbf{l}) = W^S(\mathbf{l}) := -\sum_j \mu_j W^S(\ell_j).$$

To prove this claim, we invoke a verification theorem (see, for instance, Theorem 5.1 in Fleming and Soner, 2005). Clearly, this function is continuously differentiable and satisfies the desired transversality conditions on the boundaries (when $\ell_j = 0$). We must prove that it achieves the maximum. Given the structure of B_2, we have to ensure that for every state ℓ and feasible variation $(\partial x_1, \partial x_2)$, starting from the policy $x = x^S$, the cost increases. That is, we must show that

$$\sum_j \mu_j \left(1 + \ell_j \frac{\mathrm{d}W^S(\ell_j)}{\mathrm{d}\ell_j}\right) \partial x_j \geq 0,$$

for every ∂x such that (i) $\partial x_2 \geq 0$, (ii) $\partial x_2 \geq -\frac{\mu_1}{\mu_2}\frac{1+\ell_1}{1+\ell_2}\partial x_1$. (The first requirement comes from the fact that x^S minimizes x_2 over B_2; the second comes from the other boundary line of B_2 at x^S.) Given that the result is already known for $J = 1$, we already know that this is true for the special cases $\partial x_j = 0, \partial x_{-j} \geq 0$. It remains to verify that this holds when

$$\partial x_2 = -\frac{\mu_1}{\mu_2}\frac{1+\ell_1}{1+\ell_2}\partial x_1,$$

i.e., we must verify that, for all $\ell_1 \geq \ell_2$,

$$(1+\ell_1)\ell_2 \frac{\mathrm{d}W^S(\ell_j)}{\mathrm{d}\ell_2} - (1+\ell_2)\ell_1 \frac{\mathrm{d}W^S(\ell_j)}{\mathrm{d}\ell_1} \geq \ell_2 - \ell_1,$$

or rearranging,

$$\frac{\ell_2}{1+\ell_2}\left(\frac{\mathrm{d}W^S(\ell_j)}{\mathrm{d}\ell_2} - 1\right) - \frac{\ell_1}{1+\ell_1}\left(\frac{\mathrm{d}W^S(\ell_j)}{\mathrm{d}\ell_1} - 1\right) \geq 0,$$

which follows from the fact that the function $\ell \mapsto \frac{\ell}{1+\ell}\left(\frac{\mathrm{d}\left[re^{\frac{r}{\ell}}E_{1+r}\left(\frac{r}{\ell}\right)\right]}{\mathrm{d}\ell} - 1\right)$ is decreasing.

To conclude, starting from $\ell_1 = \ell_2 = \ell_0$, the value of μ is irrelevant: the optimal strategy ensures that the posterior beliefs satisfy $\ell_1 = \ell_2$. Hence, the principal does not randomize.

The argument for a general J is similar. Fix $\ell_0 \geq \ell_1 \geq \cdots \geq \ell_J$. We argue below below that, at x^S, all possible variations must satisfy, for all $j' = 1, \dots, J$,

$$\sum_{j=j'}^{J} \mu_j (1+\ell_j)\partial x_j \geq 0,$$

It follows that we have

$$\sum_j \mu_j \left(1+\ell_j \frac{\mathrm{d}W^S(\ell_j)}{\mathrm{d}\ell_j}\right)\partial x_j = \frac{\ell_1}{1+\ell_1}\left(\frac{\mathrm{d}W^S(\ell_1)}{\mathrm{d}\ell_1}-1\right)\sum_{j'=1}^{J}\mu_{j'}(1+\ell_{j'})\partial x_{j'} +$$
$$\sum_{j=1}^{J-1}\left(\frac{\ell_{j+1}}{1+\ell_{j+1}}\left(\frac{\mathrm{d}W^S(\ell_{j+1})}{\mathrm{d}\ell_{j+1}}-1\right)-\frac{\ell_j}{1+\ell_j}\left(\frac{\mathrm{d}W^S(\ell_j)}{\mathrm{d}\ell_j}-1\right)\right)\sum_{j'=j+1}^{J}\mu_{j'}(1+\ell_{j'})\partial x_{j'} \geq 0,$$

by monotonicity of the map $\frac{\ell}{\ell+1}\left(\frac{\partial W^S(\ell)}{\partial \ell}-1\right)$, as in the case $J=2$.

To conclude, we argue that, from x^S, all variations in B_J must satisfy, for all j',

$$\sum_{j=j'}^{J}\mu_j(1+\ell_j)\partial x_j \geq 0.$$

In fact, we show that all elements of B satisfy

$$\sum_{j=j'}^{J}\mu_k\left((1+\ell_j)x_j-\ell_j\right) \geq 0,$$

and the result will follow from the fact that all these inequalities trivially bind at x^S. Consider the case $j'=1$, the modification for the general case is indicated below. To minimize

$$\sum_{j=1}^{J}\mu_j(1+\ell_j)x_j,$$

over B_J, it is best, from the formula for x_j (or rather, γ^R that are involved), to set $\gamma^{R'}=1$ for some R' for which $\alpha_j^{R'}=0$, all j. (To put it differently, to minimize the amount of experimentation conditional on the low posterior, it is best to disclose when the posterior belief is one.) It follows that

$$\begin{aligned}
&\sum_j \mu_j\left[(1+\ell_j)x_j-\ell_j\right] \\
=\ &\sum_j \mu_j\left[(1+\ell_j)\sum_R \alpha_j^R \frac{\sum_{k'}\mu_{j'}\ell_{j'}\alpha_{j'}^R}{\sum_{j'}\mu_{k'}(1+\ell_{j'})\alpha_{j'}^R}-\ell_j\right] \\
=\ &\sum_R \mu_j(1+\ell_j)\alpha_j^R\frac{\sum_{j'}\mu_{j'}\ell_{j'}\alpha_{j'}^R}{\sum_{j'}\mu_{j'}(1+\ell_{j'})\alpha_{j'}^R}-\sum_j\mu_j\ell_j \\
=\ &\sum_R\sum_{j'}\mu_{j'}\ell_{j'}\alpha_{j'}^R-\sum_j\mu_j\ell_j=\sum_{j'}\mu_{j'}\ell_{j'}\sum_R\alpha_{j'}^R-\sum_j\mu_j\ell_j=0.
\end{aligned}$$

The same argument generalizes to other values of j'. To minimize the corresponding sum, it is best to disclose the posterior beliefs that are above (*i.e.*, reveal if the movie is good, or if the chosen j is below j'), and the same argument applies, with the sum running over the relevant subset of states. □

References

[1] Ausubel, L., 2004. "An Efficient Ascending-Bid Auction for Multiple Objects," *American Economic Review*, **94**, 1452–1475.

[2] Cesari, L., 1983. *Optimization-Theory and Applications. Problems with ordinary differential equations, Applications of Mathematics* 17, Berlin-Heidelberg-New York: Springer-Verlag.

[3] Fleming, W.H. and H.N. Soner, 2005. *Controlled Markov Processes and Viscosity Solutions*, Heidelberg: Springer-Verlag.

[4] Seierstad, A. and K. Sydsæter, 1987. *Optimal Control Theory with Economic Applications*, Amsterdam: North-Holland.

[5] Simon, L.K. and M. Stinchcombe, 1989. "Extensive Form Games in Continuous Time: Pure Strategies," *Econometrica*, **57**, 1171–1214.