Online Appendix for Dynamic Persuasion and Strategic Search

1 Equivalent Representations of the Signal

The following proposition formalizes the claim that we can work with mean-preserving posterior beliefs rather than the specific signal structure in section 2.1.

Proposition 1. *The following are equivalent:*

1. *There exists a binary signal $s \in \Delta(\{0,1\})$ such that $\mathbb{P}[s=1|g] > \mathbb{P}[s=1|b]$.*

2. *There exists a binary-valued posterior belief whose expectation is equal to the prior.*

Proof of Proposition 1. $1 \Rightarrow 2$: Given a binary signal such that $\mathbb{P}[s=1|g] > \mathbb{P}[s=1|b]$, law of iterated expectation implies that $\mathbb{E}[\mathbb{P}[g|s]] = \mathbb{E}[\mathbb{E}[\mathbf{1}_{[g]}|s]] = \mathbb{E}[\mathbf{1}_{[g]}] = \mathbb{P}[g]$. So, the expectation of the posterior belief is equal to the prior. Note that $\mathbb{P}[s=0|g] < \mathbb{P}[s=01|b]$. By Bayes' rule, $\mathbb{P}[g|s=1] = \frac{\mathbb{P}[s=1|g]\mathbb{P}[g]}{\mathbb{P}[s=1|g]\mathbb{P}[g]+\mathbb{P}[s=1|b]\mathbb{P}[b]} > \frac{\mathbb{P}[s=1|g]\mathbb{P}[g]}{\mathbb{P}[s=1|g]\mathbb{P}[g]+\mathbb{P}[s=1|g]\mathbb{P}[b]} = \mathbb{P}[g] = \frac{\mathbb{P}[s=0|g]\mathbb{P}[g]}{\mathbb{P}[s=0|g]\mathbb{P}[g]+\mathbb{P}[s=0|g]\mathbb{P}[b]} > \frac{\mathbb{P}[s=0|g]\mathbb{P}[g]}{\mathbb{P}[s=0|g]\mathbb{P}[g]+\mathbb{P}[s=0|b]\mathbb{P}[b]} = \mathbb{P}[g|s=0]$. So, the posterior belief is binary-valued.

$2 \Rightarrow 1$: Given a binary-valued posterior belief whose expectation is equal to the prior, μ_0. Denote the distribution of the belief by $\mu = \begin{cases} \bar{\mu}_1 > \mu_0 & w.p.\ \lambda_1 \\ \underline{\mu}_1 < \mu_0 & w.p.\ 1-\lambda_1 \end{cases}$. We now construct a binary signal $s \in \Delta(\{0,1\})$. Define $\mathbb{P}[s=1|g] = \frac{\bar{\mu}_1\lambda_1}{\mu_0}$ and $\mathbb{P}[s=1|b] = \frac{(1-\bar{\mu}_1)\lambda_1}{(1-\mu_0)}$. One can verify by Bayes' rule that this signal s induces exactly the same posterior belief, using the assumption that $\mu_0 = \lambda_1\bar{\mu}_1 + (1-\lambda_1)\underline{\mu}_1$. We just need to show that $\mathbb{P}[s=1|g] > \mathbb{P}[s=1|b]$, which follows from the fact that $\bar{\mu}_1 > \mu_0$. □

2 Proof of the Benchmark Models

Proof of Proposition 2. We first show that the sender never provides information in both periods. For any feasible two-period signals, we will show that there exists a feasible single-period signal which gives the sender a strictly higher payoff. There are three cases.

1. A pair of one-shot signals (RHS of Figure 3).

   The sender's payoff is $\Pi = -k + p\lambda_0 + (1-\lambda_0)(-k + p\lambda_1)$.Consider a one-period strategy $(\lambda_0', \bar{\mu}_0') = (\lambda_0 + (1-\lambda_0)\lambda_1, \frac{\lambda_0\bar{\mu}_0 + (1-\lambda_0)\lambda_1\bar{\mu}_1}{\lambda_0+(1-\lambda_0)\lambda_1})$. One can verify that it satisfies (IR_0) and (F_0). The sender's payoff from it is

$$\begin{aligned}\Pi' &= -k + p[\lambda_0 + (1-\lambda_0)\lambda_1] \\ &= -k + p\lambda_0 + (1-\lambda_0)p\lambda_1 \\ &> -k + p\lambda_0 + (1-\lambda_0)p\lambda_1 - (1-\lambda_0)k \\ &= \Pi\end{aligned}$$

2. A pair of iterative signals (LHS of Figure 4). The receiver searches (takes action B) if the signal is positive (negative) in the first period, and takes action G (B) if the signal is positive (negative) in the second period.

   The sender's payoff is $\Pi = -k + \lambda_0(-k + p\lambda_1)$. Consider a one-period strategy $(\lambda'_0, \bar{\mu}'_0) = (\lambda_0\lambda_1, \bar{\mu}_1)$. One can verify that it satisfies (IR_0) and (F_0). The sender's payoff from it is

$$\begin{aligned}\Pi' &= -k + p\lambda_0\lambda_1 \\ &> -k + p\lambda_0\lambda_1 - \lambda_0 k \\ &= \Pi\end{aligned}$$

3. A pair of iterative signals (RHS of Figure 4). The receiver searches regardless of the signal realization in the first period, and takes action G (B) after observing a positive (negative) signal in the second period.

   Denote the information structure in the first period by $(\lambda_0, \bar{\mu}_0, \underline{\mu}_0)$. Denote the information structure in the second period by $(\lambda_1^p, \bar{\mu}_1^p, \underline{\mu}_1^p)$ if the receiver observes a positive signal in the first period, and by $(\lambda_1^n, \bar{\mu}_1^n, \underline{\mu}_1^n)$ if the receiver observes a negative signal in the first period. The sender's payoff is $\Pi = -k+\lambda_0(-k+p\lambda_1^p)+(1-\lambda_0)(-k+p\lambda_1^n)$. Consder a one-period strategy $(\lambda'_0, \bar{\mu}'_0, \underline{\mu}'_0) = (\lambda_0\lambda_1^p+(1-\lambda_0)\lambda_1^n, \frac{\lambda_0\lambda_1^p}{\lambda_0\lambda_1^p+(1-\lambda_0)\lambda_1^n}\bar{\mu}_1^p + \frac{(1-\lambda_0)\lambda_1^n}{\lambda_0\lambda_1^p+(1-\lambda_0)\lambda_1^n}\bar{\mu}_1^n, \frac{\mu_0-\lambda_0\lambda_1^p\bar{\mu}_1^p-(1-\lambda_0)\lambda_1^n\bar{\mu}_1^n}{1-\lambda_0\lambda_1^p-(1-\lambda_0)\lambda_1^n})$. One can verify that it satisfies (IR_0) and (F_0). The sender's payoff from it is

$$\begin{aligned}\Pi' &= -k + p\lambda'_0 \\ &= -k + p[\lambda_0\lambda_1^p + (1-\lambda_0)\lambda_1^n] \\ &= -k + \lambda_0 p\lambda_1^p + (1-\lambda_0)p\lambda_1^n \\ &> -k + \lambda_0(-k + p\lambda_1^p) + (1-\lambda_0)(-k + p\lambda_1^n) \\ &= \Pi\end{aligned}$$

We have shown that for any feasible two-period signals, there exists a feasible single-period signal which gives the sender a strictly higher payoff. Therefore, the sender never provides information in both periods. The sender either does not provide information and receives zero payoff, or provide information in one period:

$$\begin{aligned}&\max_{\lambda_0,\bar{\mu}_0} -k + p\lambda_0 \\ \text{s.t. } &\lambda_0(\bar{\mu}_0 + v_b) \geq c \qquad (IR'_0) \\ &(F_0), \lambda_0 \in [0,1], \underline{\mu}_0 \in [0, \mu_0)\end{aligned}$$

Lemma 2 implies that the constraint is equivalent to: $\lambda_0 \in \left[\frac{c}{v_g}, \frac{\mu_0 - c}{-v_b}\right], \mu_0 \geq c/v_g$. One can see that the optimal λ_0^* must be $\frac{\mu_0-c}{-v_b}$ if the sender provides information in one period. The corresponding $\bar{\mu}_0^*$ is $\frac{-v_b\mu_0}{\mu_0-c}$.

We only need to determine whether or not the sender provides information. She will provide information if and only if it is feasible, $\mu_0 \geq c/v_g$, and gives her a positive payoff, $-k + p\lambda_0^* \geq 0 \Leftrightarrow$

$\mu_0 \geq c - v_b k/p$.

$\square$

Proof of Proposition 3. One can see that if the expected surplus of the receiver in the second period is 0 in the optimal solution to (P_{dc}), then $(\lambda_1^*, \bar{\mu}_1^*)$ solves (P_1). Otherwise, the sender can strictly increase the payoff by using the same $(\lambda_0^*, \bar{\mu}_0^*)$ and replacing $(\lambda_1^*, \bar{\mu}_1^*)$ by the optimal solution to (P_1), holding the same μ_1. Hence, if dynamic commitment power strictly increases the sender surplus, the solution to (P_{dc}) must satisfy: $\mathbb{E}[\text{receiver surplus at } t = 1] = (1 - \lambda_0)[\lambda_1(\bar{\mu}_1 + v_b) - c] > 0$. Denote the optimal sender surplus when the sender does not have dynamic commitment power by Π_{wo}. The corresponding optimal strategy of the sender is $(\lambda_t^*, \bar{\mu_t}^*) = (\frac{c}{v_g}, 1)$ according to Proposition 6 in the main text. Consider the following strategy: $(\lambda_0, \bar{\mu_0}, \lambda_1, \bar{\mu_1}) = (\frac{c-\delta v_g}{(1-\delta)v_g}, 1, \frac{c}{v_g} + \delta, 1)$. Denote the corresponding sender surplus by $\Pi(\delta)$. One can verify that it is feasible when $\delta > 0$ is small and the sender has dynamic commitment power, and leads to a payoff no larger than the optimal sender surplus with dynamic commitment (denote it by Π_w).

$$
\begin{aligned}
\Pi(\delta) =& - K\left[\frac{c-\delta v_g}{(1-\delta)v_g}\right] + p\frac{c-\delta v_g}{(1-\delta)v_g} + \left[1 - \frac{c-\delta v_g}{(1-\delta)v_g}\right]\left[-K(\frac{c}{v_g}+\delta) + \frac{cp}{v_g} + \delta p\right] \\
\frac{d\Pi(\delta)}{d\delta} =& \frac{v_g - c}{(1-\delta)^2 v_g}\left[I_1(\delta) + I_2(\delta)\right] \\
&\text{, where } I_1(\delta) = K'\left[\frac{c-\delta v_g}{(1-\delta)v_g}\right] - (1-\delta)K'(\frac{cp}{v_g} + \delta) \\
&I_2(\delta) = \frac{cp}{v_g} - K(\frac{c}{v_g} + \delta)
\end{aligned}
$$

$I_1(\delta) \to 0,\ I_2(\delta) \to \frac{cp}{v_g} - K(\frac{c}{v_g}) > 0$ as $\delta \to 0$. Therefore, $\exists \widehat{\delta} > 0\ s.t.\ \frac{d\Pi(\delta)}{d\delta} > 0,\ \forall \delta \in (0, \widehat{\delta}]$. Since $\Pi(\delta)$ is continuous in δ and $\Pi(0) = \Pi_{wo}$, we have $\Pi_w \geq \Pi(\delta) > \Pi_{wo},\ \forall \delta \in (0, \widehat{\delta}]$.

We now show that the benefit of dynamic commitment power vanishes as the search cost approaches zero in several steps. Denote the optimal sender surplus when the sender has (does not have) dynamic commitment power and the search cost is c by $\Pi_{w,c}$ ($\Pi_{wo,c}$). Note that $\Pi_{w,c} \geq \Pi_{wo,c}, \forall c \geq 0$.

Lemma 11. *When the search cost is zero, the optimal sender surpluses with and without dynamic commitment power are the same,* $\Pi_{w,0} = \Pi_{wo,0}$.

Proof. Suppose the sender provides information in both periods. When the search cost is 0 and the sender uses one-shot signals, one can see that (IR_0) and (IR_1) are equivalent to $\bar{\mu}_0 \geq -v_b$ and $\bar{\mu}_1 \geq -v_b$. Therefore, even if the receiver obtains strictly positive surplus in the second period, the first-period participation constraint is not relaxed. Hence, the second-period strategy of the sender maximizes her second-period payoff in the solution to the program with dynamic commitment, (P_{dc}), and thus satisfies the constraints of the program without dynamic commitment, (P_2). To show that dynamic commitment power does not improve the sender's payoff when $c = 0$, we just need to show that iterative signals are not optimal when the sender has dynamic commitment power.[17] When $c = 0 < v_g\lambda_1^{**}$ and $\mu_1 > c - v_b\lambda_1^{**} = -v_b\lambda_1^{**}$, the optimal strategy with and without dynamic commitment power coincides, as $(\lambda_1^*, \bar{\mu}_1^*) = (\lambda_1^{**}, \frac{\mu_1}{\lambda_1^{**}} \wedge 1)$. In the proof of Proposition 1 in the main text, we have shown that iterative signals are not optimal. So, we just need to show that iterative signals are not optimal when $\mu_1 \leq -v_b\lambda_1^{**}$.

If $\mu_0 \geq -v_b\lambda_1^{**}$, then $\mu_1 = \bar{\mu}_0 > \mu_0 \geq -v_b\lambda_1^{**}$. So, iterative signals are not optimal.

If $\mu_0 < -v_b\lambda_1^{**}$, we have:

$$\begin{aligned}
\Pi_1(\mu_0) =& -K(\frac{\mu_0}{-v_b}) + \frac{\mu_0 p}{-v_b} \\
\Pi_{iter}(\mu_0) =& \max_{\lambda_0,\mu_1} -K(\lambda_0) + \lambda_0\left[-K(\frac{\mu_1}{-v_b}) + \frac{\mu_1 p}{-v_b}\right] \\
& \text{s.t. } (IR_{0,iter}), (IR_{1,iter}), (F_0), (F_1), \mu_1 = \bar{\mu}_0
\end{aligned}$$

Denote the optimal soution when the sender uses iterative signals by $(\tilde{\lambda_0}, \tilde{\mu_1})$. $\Pi_{iter}(\mu_0) \leq \Pi_1(\mu_0) \Leftrightarrow -K(\tilde{\lambda_0}) + \tilde{\lambda_0}\left[-K(\frac{\tilde{\mu_1}}{-v_b}) + \frac{\tilde{\mu_1}p}{-v_b}\right] \leq -K(\frac{\mu_0}{-v_b}) + \frac{\mu_0 p}{-v_b}$. To show that iterative signals are not optimal, it suffices to show that $-\tilde{\lambda_0}K(\frac{\tilde{\mu_1}}{-v_b}) + \frac{\tilde{\lambda_0}\tilde{\mu_1}p}{-v_b} \leq -K(\frac{\mu_0}{-v_b}) + \frac{\mu_0 p}{-v_b}$. Strict convexity of $K(\cdot) \Rightarrow \tilde{\lambda_0}K(\frac{\tilde{\mu_1}}{-v_b}) = \tilde{\lambda_0}K(\frac{\tilde{\mu_1}}{-v_b}) + (1-\tilde{\lambda_0})K(0) \geq K(\frac{\tilde{\lambda_0}\tilde{\mu_1}}{-v_b})$. Thus, $-\tilde{\lambda_0}K(\frac{\tilde{\mu_1}}{-v_b}) \leq -K(\frac{\tilde{\lambda_0}\tilde{\mu_1}}{-v_b})$. It suffices to show that $-K(\frac{\tilde{\lambda_0}\tilde{\mu_1}}{-v_b}) + \frac{\tilde{\lambda_0}\tilde{\mu_1}p}{-v_b} \leq -K(\frac{\mu_0}{-v_b}) + \frac{\mu_0 p}{-v_b}$, which hold because $(F_0) \Rightarrow \tilde{\lambda_0}\tilde{\mu_1} \leq \mu_0$ and we know from the F.O.C. that $-K(\lambda) + p\lambda$ strictly increases in λ when $\lambda < \lambda_1^{**}$ (here, $\lambda \leq \frac{\mu_0}{-v_b} < \lambda_1^{**}$). $\square$

We now make the following observation. Since $\Pi_{w,c} \geq \Pi_{wo,c}, \forall c \geq 0$ and $\Pi_{w,0} = \Pi_{wo,0}$, we must have $\Pi_{wo,c} \to \Pi_{w,c}$ as $c \to 0$ if $\Pi_{wo,c} \to \Pi_{wo,0}$ as $c \to 0$.[18] The next result confirms that it is indeed the case and thus finishes the proof.

Lemma 12. *$\Pi_{wo,c} \to \Pi_{wo,0}$ as $c \to 0$.*

Proof. Proposition 11 shows that the optimal strategy is the S_0 strategy when the search cost is low. So, according to Lemma 6 and 7, for any c small enough, the sender's problem is:

$$\Pi_{wo,c} = \max_{\mu_1} -K(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}) + p\,\frac{\mu_0-\mu_1-c}{-v_b-\mu_1} + (1-\frac{\mu_0-\mu_1-c}{-v_b-\mu_1})\left[-K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{-v_b}\right] \tag{P'''_{2S_0}}$$

$$\text{s.t. } \mu_1 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0 - c}{v_g - c}\right]$$

Denote the solution when the search cost is 0 by $\mu_{1,0}^*$. Define $\mu_{1,c}$ to be the closest value to $\mu_{1,0}^*$ among $\left[\frac{c}{v_g}, \frac{v_g\mu_0-c}{v_g-c}\right]$ and denote the corresponding sender surplus by $\underline{\Pi}_{wo,c}$. One can see that

[17] We have shown in Proposition 1 that iterative signals are not optimal when the sender does not have dynamic commitment power.

[18] One can easily show this observation formally by the triangle inequality.

$\underline{\Pi}_{wo,c} \leq \Pi_{wo,c}$. Since $\mu_{1,0}^* \in \left[0, \frac{v_g \mu_0}{v_g}\right]$, we have $\mu_{1,c} \to \mu_{1,0}^*$ as $c \to 0$. Therefore, $\underline{\Pi}_{wo,c} \to \Pi_{wo,0}$ as $c \to 0$. Since $\underline{\Pi}_{wo,c} \leq \Pi_{wo,c} \leq \Pi_{wo,0}$, we also have $\Pi_{wo,c} \to \Pi_{wo,0}$ as $c \to 0$. □

□

3 Comparative Statics

3.1 Comparative Statics With Regard to the Prior Belief When the Search Cost Is Intermediate

Proposition 2. *When the search cost is intermediate, $v_g\lambda_0^{**} < c < v_g\lambda_1^{**}$, and the sender provides information in both periods. There exists $\mu_{2,+} \in [\frac{2v_g - c}{(v_g)^2}c, \widehat{\mu_0})$ and $\mu_{2,0} \in [\frac{2v_g - c}{(v_g)^2}c, \mu_{2,+}]$. The probability of a positive signal in the first period, λ_0^*, remains the same when $\mu_0 < \mu_{2,0}$ and strictly increases in the prior when $\mu_0 > \mu_{2,+}$. The probability of a positive signal in the second period, λ_1^*, strictly increases in the prior when $\mu_0 < \mu_{2,0}$ and remains the same when $\mu_0 > \mu_{2,+}$. The belief after observing a positive signal in the second period, $\bar{\mu}_1^*$, strictly decreases in the prior when $\mu_0 < \mu_{2,0}$ or $\mu_0 > \mu_{2,+}$. A positive signal always fully reveals the state in the first period, $\bar{\mu}_0^* \equiv 1$.*

When the expected receiver surplus in the second period is zero (the S_0 strategy), the minimum amount of persuasion for the receiver to search in the first period is already too high. Under the S_+ strategy, the receiver anticipates that the sender will give him a high benefit from searching in the second period, which relaxes the first-period participation constraint. Therefore, the receiver is willing to search even if the sender reduces the amount of persuasion in the first period. This benefits the sender. However, the S_+ strategy has the disadvantage of inducing higher expected receiver search costs.

When the prior is low, the disadvantage of the S_+ strategy dominates the advantage. The sender prefers the S_0 strategy to the S_+ strategy. The likelihood of a positive signal is higher than its unconstrained optimum in the first period and lower than its unconstrained optimum in the second period. In the first period, she provides the minimum amount of persuasion for the receiver to search. So, the probability of a positive signal in the first period, λ_0^*, does not depend on the prior. In the second period, she provides the maximum amount of feasible persuasion. More frequent positive signals are feasible when the prior is higher. Even if the receiver becomes less certain about the state being good after observing a positive signal, he will still search as long as the likelihood of receiving a positive signal and earning a strictly positive surplus increases. In equilibrium, the sender trades off the precision of a positive signal for frequency as the prior increases.

When the prior is high, the advantage of the S_+ strategy dominates the disadvantage. The sender prefers the S_+ strategy to the S_0 strategy. The likelihood of a positive signal is lower than its unconstrained optimum in the first period and there is no distortion in the second period. In the first period, she provides the maximum amount of feasible persuasion, which strictly increases in the prior. In the second period, she provides the optimal amount of persuasion, which does not depend on the prior. When the prior increases, the receiver's participation constraint in the first period is relaxed. Even if the receiver anticipates a lower surplus in the second period, he will be willing to search in the first period. Therefore, the sender trades off the precision of a positive signal in the second period for the frequency of positive signals in the first period, and the belief after observing a positive signal, $\bar{\mu}_1^*$, strictly decreases in the prior. The optimal strategy is discontinuous when the sender switches the types of strategy, as illustrated in Figure 13.[19]

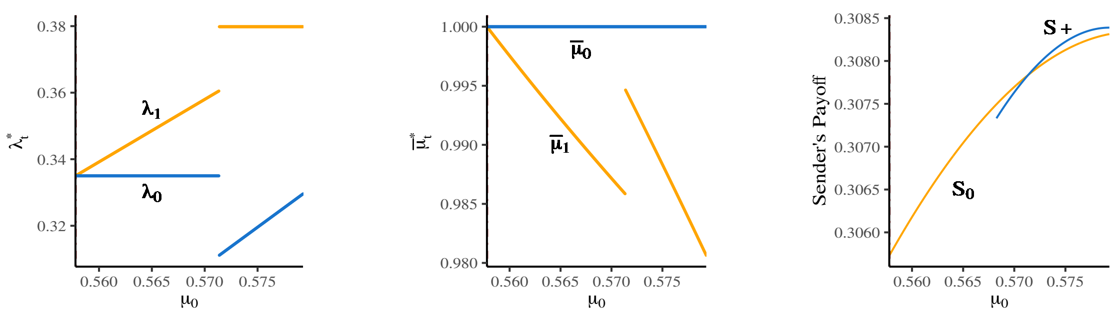

Figure 13: The sender's optimal strategy and payoff when $c = 0.067, p = 0.8, v_g = 0.2, v_b = -0.8, K(\lambda) = k\lambda^2/(1-\lambda), k = 0.5$

Proof of Proposition 2. We compare the sender surplus between the solution to (P_{2S_+}) and the solution to (P_{2S_0}). The following result shows that the optimal S_+ strategy generates a strictly higher sender surplus than the optimal S_0 strategy when the prior is high.

Lemma 13. *If $v_g\lambda_0^{**} < c < v_g\lambda_1^{**}$, the receiver gets strictly positive surplus in the second period when μ_0 is close to $\widehat{\mu_0}$.*

Proof. According to Proposition 9 in the main text, the sender achieves the benchmark payoff $-K(\lambda_0^{**}) + p\lambda_0^{**} + (1-\lambda_0^{**})\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right]$ when $\mu_0 \to \widehat{\mu_0}$ by the S_+ strategy. According to Lemma 8 in the main text, the sender surplus of the S_0 strategy is $\leq -K(\frac{c}{v_g}) + \frac{cp}{v_g} + (1-\frac{c}{v_g})\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] < -K(\lambda_0^{**}) + p\lambda_0^{**} + (1-\lambda_0^{**})\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right]$ as $\lambda_0^{**} < \frac{c}{v_g} = \lambda_0$. The difference of the benchmark sender surplus and the sender surplus of the S_0 strategy is larger than a strictly positve constant. So, when $\mu_0 \to \widehat{\mu_0}$, the S_+ strategy gives the sender strictly higher payoff. $v_g\lambda_0^{**} < c \Leftrightarrow \widehat{\mu_1}(\widehat{\mu_0}) > c - v_b\lambda_1^{**} \Rightarrow$ the receiver gets strictly positive surplus from the S_+ strategy when $\mu_0 \to \widehat{\mu_0}$. Thus, the receiver gets strictly positive surplus in the second period in equilibrium when $\mu_0 \to \widehat{\mu_0}$. Continuity of the optimal strategy and sender surplus then implies that there exists a neighborhood $[\widehat{\mu_0} - \delta, \widehat{\mu_0}]$ for some $\delta > 0$ such that the receiver gets strictly positive surplus in the second period in equilibrium when $\mu_0 \in [\widehat{\mu_0} - \delta, \widehat{\mu_0}]$. □

[19] When the search cost is intermediate, the sender always provides information in both periods provided that it is feasible. So, we do not plot the sender's payoff of providing information in only one period.

When $\frac{c}{v_g} \leq \mu_0 < \mu_{1,2}$, the sender provides information in one period. When $\mu_{1,2} \leq \mu_0 < \widehat{\mu_0}$, the sender provides information in both periods. Proposition 9, and Lemma 8 in the main text imply the explicit form of the optimal strategy. When $\mu_{1,2} \leq \mu_0 < \mu_{2,+}$, $(\lambda_0^*, \bar{\mu}_0^*) = (\frac{c}{v_g}, 1)$, $(\lambda_1^*, \bar{\mu}_1^*) = (\frac{v_g(\mu_0-c)-c(1-c)}{p(v_g-c)}, \frac{[v_g\mu_0-c]p}{v_g(\mu_0-c)-c(1-c)})$, $\mu_1^* = \frac{v_g\mu_0-c}{v_g-c} < c - v_b\lambda_1^{**}$. The receiver gets zero surplus in each period. When $\mu_{2,+} \leq \mu_0 < \widehat{\mu_0}$, $(\lambda_0^*, \bar{\mu}_0^*) = (\frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}, 1)$, $(\lambda_1^*, \bar{\mu}_1^*) = (\lambda_1^{**}, \frac{\mu_1}{\lambda_1^{**}})$, $\mu_1^* = \widehat{\mu_1}(\mu_0) > c - v_b\lambda_1^{**}$, where $\widehat{\mu_1}(\mu_0) = \frac{2c-v_b\lambda_1^{**}-(1+c-v_b\lambda_1^{**}+v_b)\mu_0}{c-v_b-\mu_0}$. The receiver gets strictly positive surplus in the second period and zero total surplus. $\square$

3.2 Comparative Statics With Regard to the Sender's Costs

Proof of Propostion 7. When the search cost is high, $v_g\lambda_1^{**} \leq c < \hat{c}$, the optimal strategy of the sender is $(\lambda_t^*, \bar{\mu_t}^*) = (\frac{c}{v_g}, 1)$ according to Proposition 6 in the main text, which does not depend on η.

When the search cost is low, $c < \tilde{c} < v_g\lambda_0^{**}$. The boundary solution does not depend on η. Consider the interior solution to (P_{2S_0}), we have: $\eta\tilde{K}(\frac{\mu_1-c}{-v_b}) + \frac{p-\mu_1}{p}\eta\tilde{K}'(\frac{\mu_1-c}{-v_b}) - \eta\tilde{K}'(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}) + c = 0$. The LHS strictly increases in μ_1 and strictly decreases in η. So, $\eta \uparrow \Rightarrow \mu_1^* \uparrow \Rightarrow \lambda_1^* = \frac{\mu_1^*-c}{p} \uparrow, \lambda_0^* = \frac{\mu_0-\mu_1^*-c}{p-\mu_1^*} \downarrow$. $\square$

4 Efficient Information Structure

4.1 Efficient Strategy in the Second Period

Proposition 3. *At the second period, the social planner does not provide information when $\mu_1 < \mu_{0,1}$. When $\mu_1 \geq \mu_{0,1}$, the efficient signal fully reveals the state when a positive signal arrives, $\bar{\mu}_{1,e} = 1$; the probability of a positive signal, $\lambda_{1,e}$, depends on the search cost:*

1. *if $c \geq v_g\lambda_1^{**}$, then there exists a unique $\hat{\hat{c}} \in (v_g\lambda_1^{**}, \mu_1 v_g]$ such that the social planner does not provide information if $c > \hat{\hat{c}}$ and $\lambda_{1,e} = \frac{c}{v_g} \vee (\tilde{\lambda_1} \wedge \mu_1)$ if $c \leq \hat{\hat{c}}$.*
2. *if $c \in [\mu_1 + v_b\lambda_1^{**}, v_g\lambda_1^{**})$, then $\lambda_{1,e} = \mu_1$.*
3. *if $c < \mu_1 + v_b\lambda_1^{**} \wedge v_g\lambda_1^{**}$, then $\lambda_{1,e} = \tilde{\lambda_1} \wedge \mu_1$.*

Proof of Proposition 3. We first introduce a relaxed problem, in which the receiver is forced to participate and the social planner can generate any signal that fully reveals the state when a positive signal arrives. The social planner chooses the information structure to maximize total welfare. We will use the solution throughout the remaining section.

$$\max_{\lambda_1} -K(\lambda_1) + \lambda_1 \tag{E_r}$$

Lemma 14. *The optimal solution to* (E_r) *exists and is unique. Denote it by* $\tilde{\lambda_1}$. *The objective function under* $\tilde{\lambda_1}$ *is strictly positive.* $\tilde{\lambda_1}$ *does not depend on the search cost* c *and* $\tilde{\lambda_1} > \lambda_1^{**}$, *the solution to the payoff-maximizing relaxed problem* (P_r).

Proof. All the results follow from the same argument as the proof of Lemma 1 in the main text except $\tilde{\lambda_1} > \lambda_1^{**}$. The F.O.C.'s imply $K'(\tilde{\lambda_1}) = 1 > p = K'(\lambda_1^{**}) \Rightarrow \tilde{\lambda_1} > \lambda_1^{**}$. $\square$

As the social planner wants to maximize the total welfare, she always wants to make the precision of the signal, $\bar{\mu_1}$, as high as possible (subject to the feasibility constraint) given μ_1 and λ_1, to increase receiver surplus while holding sender surplus fixed. So, if the social planner provides information, then $\bar{\mu_1} = \frac{\mu_1}{\lambda_1} \wedge 1$.

1. if $c \geq v_g\lambda_1^{**}$ (*i.e.* $\lambda_1^{**} \leq \frac{c}{v_g}$), then $TS = \begin{cases} -K(\lambda_1) + \mu_1 - c, & if\ \lambda_1 \geq \mu_1 \\ -K(\lambda_1) + \lambda_1 - c, & if\ \lambda_1 < \mu_1 \end{cases} \Rightarrow \lambda_{1,e} = \frac{c}{v_g} \vee (\tilde{\lambda_1} \wedge \mu_1)$ when the social planner provides information. Since $\mu_1 \geq \mu_{0,1}$, we have $\lambda_{1,e} \leq \mu_1$ and therefore $\bar{\mu}_{1,e} = 1$. The social planner provides information if the total surplus is non-negative. Similar to the proof of Proposition 4 in the main text, one can show that there exists a unique $\widehat{\widehat{c}} \in (v_g\lambda_1^{**}, \mu_1 v_g]$ such that the social planner does not provide information iff $c > \widehat{\widehat{c}}$. Moreover, if $\widehat{\widehat{c}} \geq v_g\tilde{\lambda_1}$, $\lambda_{1,e} = \frac{c}{v_g} \Rightarrow \bar{\mu}_{1,e} = 1 \Rightarrow TS = -K(\lambda_{1,e}) + p\lambda_{1,e} \Rightarrow \widehat{\widehat{c}} = \widehat{c}$.

2. if $c \in [\mu_1 + v_b\lambda_1^{**}, v_g\lambda_1^{**})$ (*i.e.* $\lambda_1^{**} \geq \frac{\mu_1 - c}{-v_b} > \frac{c}{v_g}$), then by the previous argument, $\lambda_{1,e} = \frac{c}{v_g} \vee (\tilde{\lambda_1} \wedge \mu_1) = \mu_1 \Rightarrow \bar{\mu}_{1,e} = 1$.

3. if $c < \mu_1 + v_b\lambda_1^{**} \wedge v_g\lambda_1^{**}$ (*i.e.* $\lambda_1^{**} \in \left(\frac{c}{v_g}, \frac{\mu_1 - c}{-v_b}\right)$), then the total welfare

$$TS = \begin{cases} -K(\lambda_1) + \mu_1 - c, & if\ \lambda_1 \geq \mu_1 \\ -K(\lambda_1) + \lambda_1 - c, & if\ \lambda_1 < \mu_1 \end{cases} \Rightarrow \lambda_{1,e} = \tilde{\lambda_1} \wedge \mu_1 \Rightarrow \bar{\mu}_{1,e} = 1.$$

$\square$

4.2 Efficient Strategy for the Entire Game

Proof of Proposition 8. When $v_g\tilde{\lambda_1} \leq c \leq \widehat{\widehat{c}}$, the constrained program of the social planner is:

$$\max -K(\lambda_0) + \lambda_0\bar{\mu_0} - c + (1 - \lambda_0)\left[-K(\frac{c}{v_g}) + \frac{cp}{v_g}\right] \tag{E_{2H}}$$
$$\text{s.t. } (IR_0), (F_0), \mu_1 \geq \frac{c}{v_g}$$

Using similar methods of finding the optimal sender strategy in the main text, one can show that the solution to (E_{2H}) is $(\lambda_{0,e}, \lambda_{1,e}) = (\frac{c}{v_g}, \frac{c}{v_g})$. Therefore, there is no information distortion.

When $v_g\lambda_1^{**} \leq c < v_g\tilde{\lambda_1} \wedge \widehat{\tilde{c}}$, the constrained program of the social planner is:

$$\max -K(\lambda_0) + \lambda_0\bar{\mu_0} - c + (1-\lambda_0)\left[-K(\tilde{\lambda_1} \wedge \mu_1) + \tilde{\lambda_1} \wedge \mu_1 - c\right] \quad (E_{2I})$$
$$\text{s.t. } (IR_0), (F_0), \mu_1 \geq \frac{c}{v_g}$$

Using similar methods of finding the optimal sender strategy in the main text, one can show that the solution to (E_{2I}) is $(\lambda_{0,e}, \lambda_{1,e}) = (\frac{c}{v_g}, \frac{c}{v_g})$. Therefore, there is no information distortion.

When $c < v_g\lambda_1^{**}$, one can see that we can restrict μ_1 to be less than or equal to $\tilde{\lambda_1}$ without loss of generality. The constrained program of the social planner is:

$$\max -K(\lambda_0) + \lambda_0\bar{\mu_0} - c + (1-\lambda_0)\left[-K(\mu_1) + \mu_1 - c\right] \quad (E_{2S_0})$$
$$\text{s.t. } (IR_0), (F_0), \mu_1 \geq \frac{c}{v_g}$$

Using similar methods of finding the optimal sender strategy in the main text, one can show that the solution to (E_{2S_0}) is $(\lambda_{0,e}, \lambda_{1,e}) = (\frac{\mu_0 - \frac{c}{v_g} - c}{p - \frac{c}{v_g}}, \frac{c}{v_g})$. Therefore, the sender provides too much information (relative to the efficient solution) in the second period and too little information in the first period. □

5 Extensions

5.1 Discounting

In the main model, we assume that there is no discounting. In reality, information acquisition and provision usually happen in a short period, and that assumption is reasonable. However, some communications between the sender and the receiver can take longer. Here, we study the information provision strategy when the sender and the receiver have the same discount factor, $\delta \in (0,1)$. With discounting, the sender's problem becomes:

$$\max_{\lambda_0, \bar{\mu_0}, \mu_1, \lambda_1, \bar{\mu_1}} -K(\lambda_0) + p\lambda_0 + \delta(1-\lambda_0)\left[-K(\lambda_1) + p\lambda_1\right] \quad (P_{2,\delta})$$
$$\text{s.t. } \lambda_0(\bar{\mu_0} + v_b) + \delta(1-\lambda_0)[\lambda_1(\bar{\mu_1} + v_b) - c] \geq c \quad (IR_{0,\delta})$$
$$(F_0), (\lambda_1, \bar{\mu}_1) \text{ solves (P1)}$$

One can see that both the objective function and the first-period participation constraint change. When the players become less patient (the discount factor δ decreases), the present value of the second-period sender surplus decreases, and the first-period participation constraint becomes tighter. Thus, it is less attractive for the sender to sell the goods in the second period.

Proposition 4. *When the search cost is high, $c \geq v_g\lambda_1^{**}$, the optimal strategy does not depend on the discount factor, δ. When the search cost is low, $c \leq \tilde{c}$, the sender increases the amount of* *persuasion in the first period and reduces it in the second period, as players become less patient.*

Proof. When $c \geq \widehat{c}$, the sender does not provide information for any δ.

(1) $v_g \lambda_1^{**} \leq c < \widehat{c}$

By the same argument as the proof of Proposition 6 in the main text, one can see that the optimal strategy of the sender does not depend on δ. For low prior, $\mu_0 \in [\frac{c}{v_g}, \frac{2v_g - c}{(v_g)^2}c)$, the sender provides information in one period, $(\lambda^*, \bar{\mu}^*) = (\frac{c}{v_g}, 1)$. For high prior, $\mu_0 \geq \frac{2v_g - c}{(v_g)^2}c$, the sender provides information in both periods, $(\lambda_t^*, \bar{\mu_t}^*) = (\frac{c}{v_g}, 1), t = 0, 1$.

(2) $c \leq \tilde{c}$

The sender's problem can be divided into 2 cases:

$$\begin{aligned} &\max -K(\lambda_0) + p\lambda_0 + \delta(1-\lambda_0)\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] \\ &\text{s.t. } (IR_{0,\delta}), (F_0), \mu_1 \in [c - v_b\lambda_1^{**}, \lambda_1^{**}] \end{aligned} \tag{$P_{2S_+}^{\delta}$}$$

$$\begin{aligned} &\max -K(\lambda_0) + p\lambda_0 + \delta(1-\lambda_0)\left[-K(\frac{\mu_1 - c}{-v_b}) + \frac{(\mu_1 - c)p}{-v_b}\right] \\ &\text{s.t. } (IR_{0,\delta}), (F_0), \mu_1 \in [\frac{c}{v_g}, c - v_b\lambda_1^{**}] \end{aligned} \tag{$P_{2S_0}^{\delta}$}$$

Consider the solution to ($P_{2S_0}^{\delta}$). By the same argument as the proof of Proposition 5 in the main text, ($P_{2S_0}^{\delta}$) is equivalent to

$$\begin{aligned} &\max -K(\lambda_0) + p\lambda_0 + \delta(1-\lambda_0)\left[-K(\frac{\mu_1 - c}{-v_b}) + \frac{(\mu_1 - c)p}{-v_b}\right] \\ &\text{s.t. } \lambda_0 \in \left[\frac{\mu_0 - \mu_1}{1 - \mu_1}, \frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1}\right] \\ &\qquad \mu_1 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0 - c}{v_g - c}\right] \end{aligned} \tag{$P_{2S_0}^{\delta'}$}$$

By arguments similar to the proof of Proposition 5 in the main text, $\lambda_0 \leq \frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1}$ is binding for $\mu_0 < \widehat{\mu_0}$. Suppose that μ_1's constraints are not binding.

The Lagrangian is $\mathcal{L} = -K(\lambda_0) + p\lambda_0 + \delta(1-\lambda_0)\left[-K(\frac{\mu_1 - c}{-v_b}) + \frac{(\mu_1 - c)p}{-v_b}\right] + \eta\left(\frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1} - \lambda_0\right)$ *s.t.* $\eta \geq$

$0, \eta\left(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1} - \lambda_0\right) = 0$. F.O.C. $\Rightarrow$

$$\begin{aligned}
\eta &= \delta(p-\mu_1) - \delta\frac{p-\mu_1}{p}K'(\frac{\mu_1-c}{-v_b}) \\
&= -K'(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}) + p + \delta\left[K(\frac{\mu_1-c}{-v_b}) - \mu_1 + c\right] \\
\Rightarrow & \delta\left[K(\frac{\mu_1-c}{-v_b}) + \frac{p-\mu_1}{p}K'(\frac{\mu_1-c}{-v_b}) - p + c\right] - K'(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}) + p = 0 \qquad (*_\delta)
\end{aligned}$$

The sum of the first two terms of the LHS of ($*_\delta$) strictly increases in δ, and the LHS of ($*_\delta$) strictly increases in μ_1. So, the optimal λ_0 strictly decreases in δ; the optimal μ_1 and λ_1 strictly increase in δ. When one of μ_1's constraints is binding, λ_0 and λ_1 does not depend on δ.

We finish the proof of this case by showing that the S_0 strategy dominates the S_+ strategy. Denote the sender surplus of the optimal S_0 (S_+) strategy when the discount factor is $\delta \in (0,1)$ by $\Pi_{S_0}(\delta)$ ($\Pi_{S_+}(\delta)$) and denote the corresponding optimal S_0 (S_+) strategy at time t by $\lambda_{t,S_0}(\delta)$ ($\lambda_{t,S_+}(\delta)$). We have:

$$\begin{aligned}
\Pi_{S_0}(\delta) =& -K(\lambda_{0,S_0}(\delta)) + p\lambda_{0,S_0}(\delta) + \delta(1-\lambda_{0,S_0}(\delta))\left[-K(\lambda_{1,S_0}(\delta)) + p\lambda_{1,S_0}(\delta)\right] \\
\geq& -K(\lambda_{0,S_0}(1)) + p\lambda_{0,S_0}(1) + \delta(1-\lambda_{0,S_0}(1))\left[-K(\lambda_{1,S_0}(1)) + p\lambda_{1,S_0}(1)\right] \\
=& -K(\lambda_{0,S_0}(1)) + p\lambda_{0,S_0}(1) + (1-\lambda_{0,S_0}(1))\left[-K(\lambda_{1,S_0}(1)) + p\lambda_{1,S_0}(1)\right] \\
& -(1-\delta)(1-\lambda_{0,S_0}(1))\left[-K(\lambda_{1,S_0}(1)) + p\lambda_{1,S_0}(1)\right] \\
\overset{(\dagger)}{\geq}& -K(\lambda_{0,S_+}(1)) + p\lambda_{0,S_+}(1) + (1-\lambda_{0,S_+}(1))\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] \\
& -(1-\delta)(1-\lambda_{0,S_+}(1))\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] \\
=& -K(\lambda_{0,S_+}(1)) + p\lambda_{0,S_+}(1) + \delta(1-\lambda_{0,S_+}(1))\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] \\
\geq& -K(\lambda_{0,S_+}(\delta)) + p\lambda_{0,S_+}(\delta) + \delta(1-\lambda_{0,S_+}(\delta))\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] \\
=& \Pi_{S_+}(\delta)
\end{aligned}$$

, where the inequality ($\dagger$) holds because we have shown in the proof of Proposition 6 in the main text that in this case $-K(\lambda_{0,S_0}(1)) + p\lambda_{0,S_0}(1) + (1-\lambda_{0,S_0}(1))\left[-K(\lambda_{1,S_0}(1)) + p\lambda_{1,S_0}(1)\right] \geq -K(\lambda_{0,S_+}(1)) + p\lambda_{0,S_+}(1) + (1-\lambda_{0,S_+}(1))\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right]$ and we have $\lambda_{0,S_0}(1) \geq \lambda_{0,S_+}(1)$, $-K(\lambda_{1,S_0}(1)) + p\lambda_{1,S_0}(1) \leq -K(\lambda_1^{**}) + p\lambda_1^{**}$.

□

When the search cost is high, it is hard to satisfy the participation constraints of the receiver. Providing enough information to persuade the receiver to search dominates the force of discounting. Therefore, the sender's strategy remains the same as the no-discounting case, and the sender perfectly smooths information provision. When the search cost is low, the sender reduces the amount of persuasion in the second period when the players are less patient because of discounting.

The sender increases the amount of persuasion in the first period as she becomes more tempted to convert the receiver early.

We assume in the main model that the receiver is forward-looking and takes into account the potential payoff of the second period when he chooses his action in the first period. This assumption is reasonable if the information environment is transparent and the receiver knows that gradual learning is possible. We now consider the possibility of asymmetric discount factors. In particular, the sender is perfectly patient while the receiver is perfectly impatient (myopic). If the receiver is myopic, then he trades off only the current-period benefit and cost in deciding whether or not to search. The information in the second period cannot relax the first-period participation constraint. Therefore, when the receiver is myopic, the feasible information structure is a subset of when the receiver is forward-looking. If the receiver's surplus in the second period is strictly positive when the receiver is forward-looking, the optimal information structure may not be feasible when the receiver is myopic. Hence, the sender is (weakly) worse off if the receiver is myopic rather than forward-looking. This result has managerial implications, as it suggests that the sender should try to better inform the receiver of the possibility of gradual learning. Common knowledge of gradual information revelation improves the sender's surplus.

5.2 Infinite Number of Periods

The two-period model can capture information smoothing and gradual learning. We extend the main model to a model with an infinite number of periods and show that the main insights extend to this richer model. This also provides some additional insights.

Time is discrete, $t = 0, 1, 2, ...$ In each period, the sender determines and commits to the information structure of the current period but cannot commit to the information structure in the future. The receiver can search for information before deciding. Unlike the two-period model, there is no deadline. The receiver can search for as long as he wants. Since the sender's payoff is bounded by p, the payoff function is well-defined even without the discount factor. So, we do not consider discounting for consistency with the main model. We can analyze the problem similarly if we include discounting.

Proposition 5. *When the search cost is high, $v_g\lambda_1^{**} \leq c < \widehat{c}$, the sender provides perfectly smooth information for $k := \lfloor \frac{ln(1-\mu_0)}{ln(1-c/v_g)} \rfloor$ periods, and a positive signal fully reveals the state, $(\lambda_t^*, \bar{\mu_t}^*) = (c/v_g, 1)$, for $t = 0, 1, ..., k-1$. When the search cost is low, the sender adds noise to positive signals. As the search cost approaches zero, the sender could obtain the equilibrium payoff as if the persuasion cost were zero.*

Proof.

(1) High search cost $v_g\lambda_1^{**} \leq c < \widehat{c}$

Consider an arbitrary period t in which the receiver takes action G. Suppose the belief at the beginning of period t is μ_t.[20] One can see that the receiver must take action G after observing a positive signal and take action B or S after observing a negative signal in period t. Denote the sender's (receiver's) continuation value after the receiver observes a negative signal in period t by V_t (W_t). $V_t, W_t \geq 0$ and (weakly) increase in $\underline{\mu}_t$. The sender's problem in period t is:

$$\max_{\lambda_t, \bar{\mu}_t} -K(\lambda_t) + p\lambda_t + (1-\lambda_t)V_t$$

$$s.t.\ \lambda_t(\bar{\mu}_t + v_b) + (1-\lambda_t)W_t \geq c \tag{IR_t}$$

$$\lambda_t \bar{\mu}_t + (1-\lambda_t)\underline{\mu}_t = \mu_t \tag{F_t}$$

Denote the optimal λ_t without constraints by $\lambda^{**}_{t,H}$. $\lambda^{**}_{t,H} = \arg\max_{\lambda_t} -K(\lambda_t) + p\lambda_t + (1-\lambda_t)V_t$. The F.O.C. of $-K(\lambda_t) + p\lambda_t + (1-\lambda_t)V_t \Rightarrow K'(\lambda^{**}_{t,H}) = p - V_t < p = K'(\lambda^{**}_1) \Rightarrow \lambda^{**}_{t,H} < \lambda^{**}_1 \leq c/v_g$. For any information structure in which $\lambda_t > c/v_g$, by reducing λ_t (and potentially increasing $\bar{\mu}_t$ to satisfy the participation constraint), the sender can increase her payoff if V_t is fixed. One can see that $\underline{\mu}_t$ and W_t will (weakly) increase, as long as we keep (IR_t) binding. Hence, V_t will also (weakly) increase. So, the sender's payoff will be even higher. So, under the optimal information structure, λ^*_t must be no greater than c/v_g. Since it holds for any period in which the receiver takes action G, and the receiver surplus from that period is $\lambda^*_t(\bar{\mu}_t + v_b) - c \leq \frac{c}{v_g}(1+v_b) - c = 0$. The receiver gets zero surplus in each period, $W_t = 0$. Therefore, the receiver takes action G immediately after observing a positive signal in any period he searches for information (otherwise, the expected gain from search is strictly negative and he will not search). This implies that the optimal information structure is k periods of one-shot signals. To satisfy the receiver's particiaption constraints, $(\lambda^*_t, \bar{\mu}^*_t) = (c/v_g, 1)$, for $t = 0, 1, ..., k-1$.

The sender's expected payoff of providing k periods of such information is: $\sum_{i=0}^{k-1}(1-\frac{c}{v_g})^i\left[\frac{cp}{v_g} - K(\frac{c}{v_g})\right]$, which increases in k. Thus, the sender will provide as many periods of information as possible. Now we characterize the maximum number of periods.

Denote the initial belief at the beginning of period t by $\mu_t := \underline{\mu}_{t-1}$. The feasibility costraint (F_t) and $(\lambda^*_t, \bar{\mu}^*_t) = (c/v_g, 1)$ imply that $\mu_t = \frac{\mu_{t-1} - \frac{c}{v_g}}{1-\frac{c}{v_g}}$. By induction, one can show that $\mu_k = \frac{\mu_0 - 1 + (1-\frac{c}{v_g})^k}{(1-\frac{c}{v_g})^k}$. For it to be feasible to provide information in k periods, we need $\mu_{k-1} \geq c/v_g \Leftrightarrow k \leq \frac{ln(1-\mu_0)}{ln(1-c/v_g)}$. Hence, the maximum number of periods is $\lfloor \frac{ln(1-\mu_0)}{ln(1-c/v_g)} \rfloor$.

(2) Low search cost

The receiver needs to decide between G and B at the end of the game. We first derive an upper bound on the probability that the receiver decides on G under any feasible information

[20] It is possible that there are more than one initial beliefs at the beginning of period t if the receiver searches for information regardless of the signal realization in a previous period. In that case, μ_t is any one of them. We will show that the optimal information structure is one-shot. So, there is actually one initial belief in any period.

structure. Denote the probability that the receiver takes action G in period t by q_t. Because the belief must be greater than or equal to $-v_b$ if the receiver takes action G, the mean-preserving property of the beliefs implies that $\mu_0 \geq \sum_{t=0}^{+\infty} q_t(-v_b) \Rightarrow \sum_{t=0}^{+\infty} q_t \leq -\frac{\mu_0}{v_b}$. The probability that the receiver takes action G eventually is bounded from above by $-\frac{\mu_0}{v_b}$. Thus, the sender's payoff is bounded from above by $-\frac{\mu_0 p}{v_b}$, even if the persuasion cost is zero. Now we show that the sender can achieve that payoff as the search cost vanishes. This means that she can obtain the equilibrium payoff as if the persuasion cost were zero.

Consider the following strategy: Given a search cost c, the sender provides the same one-shot signals for T consecutive periods ($t = 0, 1, ..., T-1$), where $(\lambda_t, \bar{\mu}_t) = (\lambda, \bar{\mu}) = (\sqrt{c}, -v_b + \sqrt{c})$ and $T = \frac{ln\left(1-\frac{\mu_0}{-v_b+\sqrt{c}}\right)}{ln(1-\sqrt{c})}$. In each period, the receiver's expected payoff from searching is $\lambda_t(\bar{\mu}_t + v_b) - c = 0$. So, the receiver will keep searching if he observes a negative signal, except in the last period. By setting $\underline{\mu}_{T-1} = 0$, one can verify that the mean-preserving property of the belief is satisfied, and that the variables are well-defined for c small. The probability that the receiver takes action G eventually is $\sum_{t=0}^{T-1} \lambda(1-\lambda)^t = 1-(1-\lambda)^T = \frac{\mu_0}{-v_b+\sqrt{c}} \to -\frac{\mu_0}{v_b}$ as $c \to 0$.

The sender only incurs the persuasion cost if the receiver has not received a good signal. The expected total persuasion cost of the sender is bounded from above by the costs of always providing the information in T periods, which is $TK(\sqrt{c})$.

$$\lim_{c\to 0} TK(\sqrt{c}) = \lim_{c\to 0} \frac{ln\left(1-\frac{\mu_0}{-v_b+\sqrt{c}}\right)}{ln(1-\sqrt{c})} K(\sqrt{c}) = ln\left(1+\frac{\mu_0}{v_b}\right) \lim_{c\to 0} \frac{K(\sqrt{c})}{ln(1-\sqrt{c})} \overset{\text{L'Hospital's rule}}{=} 0$$

Hence, the sender's payoff approaches $-\frac{\mu_0 p}{v_b}$ as $c \to 0$. The receiver's expected payoff from searching (net of the search cost) is zero in every period given the above strategy. One can see that the sender's payoff under the optimal strategy is no lower than that payoff, and it is bounded from above by $-\frac{\mu_0 p}{v_b}$. So, it also approaches $-\frac{\mu_0 p}{v_b}$ as $c \to 0$.

Now we show that the sender adds noise to positive signals when the search cost is low. Suppose not. The mean-preserving property of the beliefs implies that $\mu_0 \geq \sum_{t=0}^{+\infty} q_t \cdot 1 \Rightarrow \sum_{t=0}^{+\infty} q_t \leq \mu_0 < -\frac{\mu_0}{v_b}$. Then, the payoff of the sender is bounded from above by $\mu_0 p < -\frac{\mu_0 p}{v_b}$. But, we have shown that the sender's payoff approaches $-\frac{\mu_0 p}{v_b}$ as $c \to 0$. So, it cannot be optimal for small c. Therefore, the sender adds noise to positive signals when the search cost is low.

□

This proposition shows that the main insights are robust to the specification of the length of time. The two-period model corresponds to the case when there is a deadline in the information acquisition. When time is infinite, there is no limit on how long the receiver can search. Under high search costs, the optimal information structure fully convinces the receiver that the state is g when a positive signal arrives. Under low search costs, the optimal information structure adds some noise to the positive signal. The intuition is the same as the two-period case. Because the receiver can keep searching for a longer period, the sender can better smooth the information. When the search cost is high, the sender may provide information for more than two periods if she believes that the state is likely to be g and the receiver is willing to spend more time searching. The proposition shows that the sender smooths information provision over more periods for a higher prior belief. The ability to smooth the information is valuable for the sender, especially when the search cost approaches zero. In that case, she could obtain the equilibrium payoff as if the persuasion cost were zero. Because of the low search cost, the sender can convince the receiver to search with very little information in each period. The sender's cost becomes very low by smoothing the information over many periods.