Disclosure in Multistage Projects*

Omer Ali†

This version: November 2019

Abstract

I study a dynamic game between a sender (she) and a receiver (he). The receiver incurs a cost in every period in which he chooses to stay on the project. If he quits, the game ends and both players receive a continuation payoff of zero. The project is completed when two successes occur, at which point the game ends and both players receive positive continuation payoffs. Only good projects yield successes, and neither player knows the project's type at the beginning. Only the sender observes the first success, and she chooses when (and how) to share this information with the receiver. When she commits to revealing the first success as soon as it arrives, the receiver's beliefs about the quality of the project decline rapidly as he waits for news, leading him to quit prematurely. The main results show that the sender can incentivize the receiver to stay on the project longer by promising to reveal information with a delay. I also show that the mechanism that min-maxes the agent is not necessarily optimal.

*I would like to thank Maciej Kotowski, Ichiro Obara, Marek Pycia and Romain Wacziarg for their support and encouragement throughout this project. I would also like to thank Byeong-hyeon Jeong and Daehyun Kim for many fruitful discussions. The project benefitted from comments by Simon Board, Andrea Di Miceli, Andreas Gulyas, Jay Lu, Moritz Meyer-ter-Vehn, Nic Motz, Guillaume Sublet, and audiences at the UCLA Theory Proseminar. Remaining errors are the sole responsibility of the author.

†omer.ali@duke.edu; Samuel DuBois Cook Center on Social Equity, Duke University

1 Introduction

When and how individuals receive information about their progress on a task has significant effects on the effort they provide. While the incentive role of monetary payments is well understood, economists have only recently begun exploring the incentive role of information. In this paper, I study how information supplied by a sender affects a receiver's incentives to keep working on a project.

There are many applications of this framework: an entrepreneur who must convince her investors that the project is going well has an incentive to share good news as soon as it arrives. However, if she commits to this disclosure policy, the absence of good news makes investors more pessimistic about the prospects of the venture, and they may withdraw their funding. A physician who would like to keep her patient on an experimental treatment for as long as possible and promises to disclose positive results of the treatment risks making her patient more pessimistic when there is no good news to share. Finally, a firm that benefits from R&D may want researchers to keep working on a new technology even when there is uncertainty about the prospects of successfully marketing the innovation. When the firm receives information indicating that the prospects are good, it would like to make this known to its researchers. Committing to do so, however, means that as long as no good news is received, researchers conclude that things are not going well.

In all of the above examples, the sender must balance between the benefits of sharing good news and the discouragement effect of its absence. I study this tradeoff in a model in which the sender learns about the underlying quality of a project through the arrival of successes over time and must choose whether (and how) to share her information with the receiver. The project is completed when two successes arrive, which captures the idea that it has two distinct stages. While the first success is only observable to the sender, the second success is public and ends the game. Only good projects yield successes, which arrive stochastically in every period in which the receiver chooses to work. At the beginning of the relationship, neither player knows the type of the project and they share a common prior belief, p_0, that it is, in fact, a good project. The sender does not incur any cost herself, but also benefits from the completion of the project, an event I will refer to as a 'breakthrough'. When the sender and receiver choose their disclosure policy and quitting strategy (respectively) simultaneously, there is a class of equilibria characterized by the length of time the receiver is willing to work from the beginning of the game without receiving any information, and the threshold beliefs they use thereafter to decide whether or not to quit. We study two focal disclosure mechanisms: promise policies keep the receiver working until a pre-specified time at which point the sender releases some (or all) information; the indifference policy involves releasing just enough information in each period to make the receiver indifferent between working and quitting. While neither policy is the optimal mechanism in all cases, a number of partial results indicate that promise strategies are superior for the sender earlier in the game, while the indifference policy is better in its later stages.

One of the main tradeoffs faced by the sender in the model is that information can have both a positive as well as a negative effect on incentives. Good news (the arrival of the first success) motivates the receiver to keep working, but its absence is discouraging. This tradeoff is present in static models of persuasion such as Kamenica and Gentzkow (2011).[1] An additional tradeoff that the sender takes into account is introduced by the dynamics of the problem and relates to the timing of information revelation. While the sender always prefers to delay revealing information, the receiver prefers to learn about the outcome history as early as possible. Suppose that there is a certain amount of information that the sender can reveal to ensure that the receiver stays in the game whenever they are asked to do so. Can the sender decrease this amount in return for giving the receiver more information later in the game? The answer to this question determines whether promise policies improve upon the indifference policy described above. When it is possible to delay information revelation with a promise of a future reward, the sender considers whether the reward is affordable and, if so, changes their disclosure policy accordingly. A novel feature of the framework I study is that this tradeoff changes over time because of the non-stationarity of the problem. While in the early stages of the game, the receiver's beliefs are in a region of the belief space where the necessary rewards are affordable, over time, they move into a region where the sender prefers to stick to the indifference policy.

A growing literature studies how the provision of information can induce a receiver to take certain actions, even when the incentives of the receiver and the information provider are not aligned. Kamenica and Gentzkow (2011) study a static model in which a sender knows the state and commits to a disclosure policy that partially reveals it to the receiver. Whereas the sender's most preferred action is state-independent, the receiver conditions their choice on their beliefs about the state. The sender must calibrate their disclosure policy in a way that maintains the credibility of their messages. While the sender would always prefer to inform the receiver that the state is such that the sender's most preferred action is the appropriate one, the receiver would, in response, disregard this information and instead choose according to their prior. This principle is also at play in the setting I study: the sender would like the

[1]Similar frameworks are also studied in Calzolari and Pavan (2006) and Ivanov (2010).

receiver to continue working for as long as possible, but sending them good news regardless of the outcome history observed renders the messages unpersuasive.

A dynamic version of this problem is studied in Ely (2017): a sender observes the true state as it evolves according to a Markov process. The receiver knows the transition rule, but does not observe the state. Messages from the sender can be conditioned on their information in such a way that sometimes induces the receiver to take the sender's preferred action. In this setting, the myopic optimal disclosure policy is also optimal when the receiver is patient and strategic. This is in stark contrast with the results I find in this paper: the sender can do strictly better when they promise to reveal information in the future, thereby taking advantage of the receiver's patience.[2]

While the framework studied in this paper is novel, it combines features present in existing work. The multistage nature of the project is similar to the models in Bimpikis, Ehsani, and Mostagir (2016) and Green and Taylor (2016). Both of these papers, however, assume that the first success is the receiver's private information, not the sender's. Dynamic information disclosure is present in Ely (2017) and Orlov, Skrzypacz, and Zryumov (2016). Learning about the underlying productivity of the project and choosing how to share this information with the receiver is present in Orlov (2015) and Smolin (2015). In Pei (2016), an intermediary shares their information about the quality of the project/agent with the market. Che and Horner (2015) study the optimal recommendation strategy when a platform maximizes the welfare of consumers who sample a product over time.

Below, we describe some examples that motivate the structure of the game described above. The rest of the paper proceeds as follows: section 2 describes the framework and the evolution of beliefs; section 3 describes the mechanism design problem facing the sender; section 4 describes the simultaneous move game between sender and receiver when the latter also has commitment power, and section 5 concludes. Proofs are relegated to the appendix.

[2] Smolin (2015) studies the optimal feedback policy when the sender receives private information about the receiver's performance over time. This is the closest framework in spirit to the one studied in this paper. Payoffs accrue to the players in every period and the game is played for an infinite number of periods with no end date. These features ensure that, unlike the case here, values are stationary in beliefs and independent of calendar time. When the receiver prefers to quit earlier than the sender would like, an optimal policy involves a coarsening of the information received by the sender: they only communicate to the receiver whether or not their beliefs are above a certain threshold. In the model I study, however, the sender can sometimes increase their payoff by randomly disclosing bad news and keeping the receiver in the game with strictly positive probability in every period. In this way, the receiver continues working even when the sender's beliefs drift arbitrarily close to zero.

1.1 Examples

1.1.1 Venture capital

Consider a venture capitalist involved in funding a project run by an entrepreneur. Suppose the project requires building a prototype before the final product is produced. Funding occurs in discrete stages (periods) and allows the entrepreneur to continue working. If the project is good, a prototype is successfully built with some probability in each stage; a bad project never yields a functioning prototype. The entrepreneur observes whether or not the prototype built in each period is functioning and can choose whether (and how) to share this information with the investor. While the investor never observes the state of prototypes, the ultimate success of the project (say, the development of the marketable product) is public information. This occurs with some probability in every period but only after a prototype has been successfully developed. The prototype, therefore, represents an intermediate stage of the production process that reveals the quality of the project being undertaken by the entrepreneur, as well as the news that development of the product is halfway complete. The entrepreneur would like to work on the project for as long as possible, but the investor would only like to invest if the rewards are sufficiently high. How should the entrepreneur reveal information about the state of the prototype to the investor?

1.1.2 Experimental treatment

Consider a physician treating a patient with a condition that has no known cure. There is an experimental drug that can either be effective ($\omega = 1$) or not. Since the drug is unproven, neither party knows its quality for sure and they both begin with a shared common prior. The drug is taken once a week in pill form and is guaranteed to generate painful side effects that only the patient suffers ($c > 0$). During the week, the patient undergoes some tests that only the physician has access to. The weekly tests perfectly reveal whether the week's treatment was successful, but the patient doesn't feel any better. The condition is cured when two weeks of treatment succeed, only at which point does the patient know that the drug must have worked. The physician can choose whether and how to share information about the weekly tests with the patient. The patient would like to continue with the treatment but only if there is a reasonably high chance of success. The physician, on the other hand, would prefer to keep the patient on the drug for as long as possible. How should the physician communicate the results of the weekly tests to the patient?

1.1.3 Research and development

Consider a firm with an active research department. The firm ultimately monetizes the results of successful research projects and incentivizes its researchers by rewarding them with a financial stake in the project. They receive a payoff only if/when their research is successfully developed into a marketable product. Researchers can choose whether or not to contribute to a project, but they can only keep their financial stake alive if they continue working. Let's examine a researcher's decision whether or not to start and continue working on a particular project. If the project is good, it can be successfully sold by the marketing department, but in each time period this happens only with some probability. If the project is bad, it can never be marketed successfully and no matter how long the researcher remains involved, their stake will never bear fruit. The firm understands the researchers' incentives, but would like them to continue exploring an idea even when marketing efforts prove unsuccessful, because their planning horizon is longer than the researcher's, who may not necessarily spend their entire career at the firm. The marketing department receives information about the success of their efforts over time. Suppose there are two stages of a successful marketing campaign: (i) developing a strong core group of enthusiastic customers, and (ii) mass market success. The second stage can never be achieved before the first, and neither stage is possible with a bad project. When a project is successfully mass marketed, the researchers receive their payoffs. However, they do not learn about whether the marketing department has successfully mobilized a core group of enthusiasts. The firm's management can commit to sharing information about the first stage with researchers to keep them involved in the project for as long as possible. How should the firm share information about the outcome of its marketing efforts with its researchers?

2 Model

2.1 Environment

A sender and receiver collaborate on a joint project over discrete time. The project, which can either be good quality or bad, has the following characteristics: when the project is good, there is a positive probability of a success in every period in which the receiver chooses to work. When the project is bad, however, successes never arrive. The receiver must decide whether to work or quit in each period, and once they choose to quit, they can never work on the project again. Completion of the project - which we refer to as a "breakthrough" -

requires the arrival of two successes, only at which point do the players receive a payoff. It is helpful to view the first success as a project milestone, such as a functioning prototype, and the second success as the end of the project. Examples of such projects are common in venture financing, medicine, and R&D in firms (see subsection 1.1). While working makes the realization of a payoff more likely, it is costly for the receiver. As a result, they only choose to work when the likelihood of a breakthrough is sufficiently high.

Only the sender observes the arrival of the first success and decides whether (and how) to share this information with the receiver. We assume that the sender commits to a disclosure policy at the beginning of the game and cannot deviate from their chosen policy. After observing the outcome in each period, they can send a message to the receiver, which may depend on the history of outcomes realized and messages sent thus far. The receiver interprets this message in light of the sender's commitment to send certain messages only when certain histories arise. Messages inform the receiver about likely outcome histories and about the quality of the project. A message that the sender commits to sending more often when a success has arrived makes the receiver more optimistic about the quality of the project. The absence of such a message, however, has the opposite effect. The sender's decision problem reflects the tradeoff faced, for example, by an entrepreneur who commits to sharing information about her progress on a project of uncertain viability with her investors. While a breakthrough (in this case, the ultimate success of the venture) is observable to both the entrepreneur as well as her investors, progress is only known to the entrepreneur and she must consider the effect of regularly sharing news of this progress (or lack thereof) on the investors' enthusiasm for her project, and the likelihood of securing their continued support. The sequence of actions for the sender and receiver within a given period are described in figure 1.

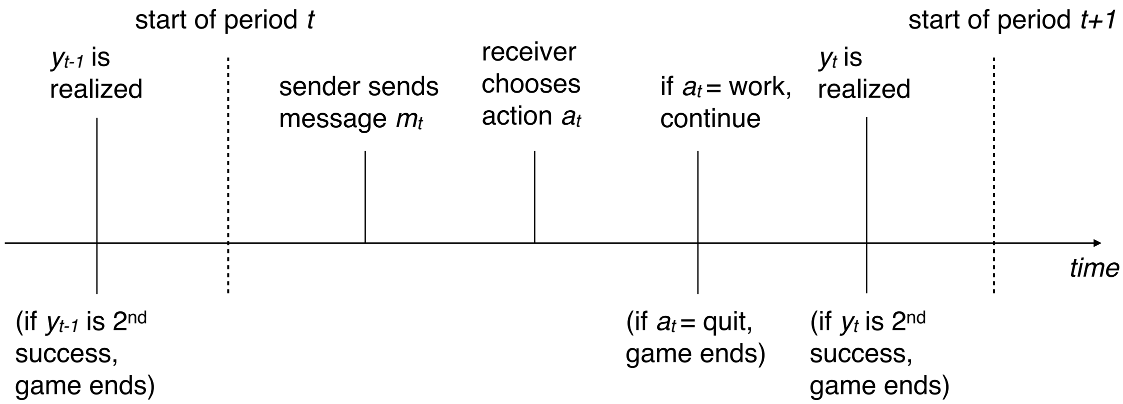
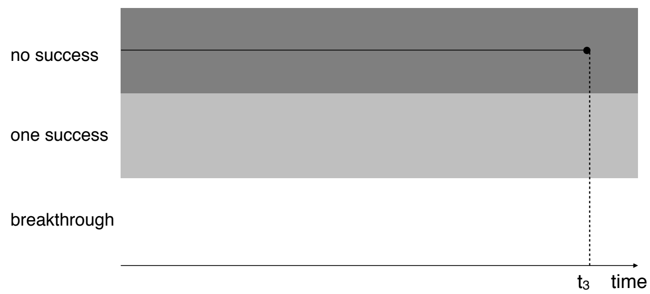


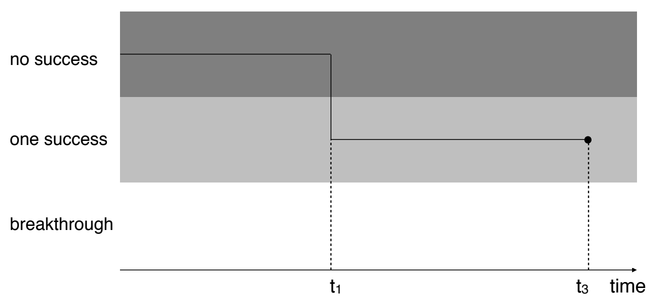
Figure 1: The sequence of actions within a period.

Let $\omega \in \Omega = \{0, 1\}$ denote the quality of the project. Neither player knows this parameter for sure at the beginning of the game.[3] The project is good ($\omega = 1$) with common prior probability p_0, in which case successes arrive with probability $\theta \in (0, 1)$ in every period in which the receiver chooses to work. The receiver incurs a per-period cost of $c > 0$ whenever they choose to work. If the receiver chooses to quit, the game ends and each player receives a payoff of zero. While the receiver never observes the arrival of an intermediate success, a breakthrough is observable to both players and ends the game. When a breakthrough occurs, each player receives a lump sum payoff of $B > 0$. The sender observes the entire history of outcomes (successes and failures in every period). They choose when and how to share this information with the receiver. As such, both players hold beliefs about the quality of the project that evolve over time. Notice that the benefit of a breakthrough is the same for both players, yet a misalignment of incentives exists because only the receiver incurs the cost of working.

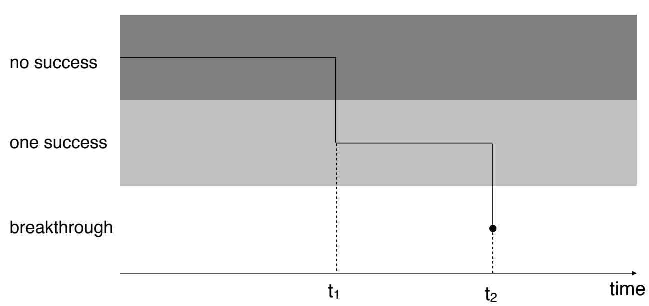
[3]This can alternatively be interpreted as the receiver's type, which is unknown to both players.



(a) An outcome history with no successes.



(b) An outcome history with one success, but no breakthrough.



(c) An outcome history with a breakthrough.

Figure 2: Possible outcome histories when the receiver quits at time t_3.

2.2 Outcomes, payoffs and preferences

Recall that the sender and receiver only receive a payoff of B when a breakthrough occurs. Along the way, the receiver incurs a cost of c in every period, but the sender does not. Let $y_t \in \{0, 1\}$ denote the *outcome* in period t when the receiver chooses to work, with 1 representing a success and 0 a failure. The *outcome history*, $y^t = (y_1, \ldots, y_t)$, is the sequence of outcomes up to time t. Let the set of all histories up to time t be Y^t. Define the function $S : Y^t \to \{0, 1, 2\}$ in the following way:

$$S\left(y^{t}\right)=\begin{cases}2 & if\ \sum_{s=1}^{t} y_{s}=2\ and\ y_{t}=1\\ 1 & if\ \sum_{s=1}^{t} y_{s}=1\\ 0 & if\ \sum_{s=1}^{t} y_{s}=0\end{cases}$$

The function S counts the number of successes in a history. Since the project ends when two successes arrive, histories have at most two successes, and the second must arrive in the final period. Y^t is the set of all t-period histories satisfying this restriction:

$$Y^{t}=\left\{y^{t}=\left(y_{1},\ldots,y_{t}\right)\in\{0,1\}^{t}: S\left(y^{t}\right)\leq 2,\ S\left(y^{t}\right)=2 \iff y_{t}=1\right\}$$

Notice that both players have preferences over outcome histories, not outcomes per se, since intermediate successes are not valued independently of the history in which they occur. A 10-period history with a success in the 5th period and no successes thereafter does not yield a payoff, whereas a 10-period history with successes in the 5th and 10th periods does. The 5th period success, therefore, is not independently valuable.

In addition to the players' utilities being defined over the space of outcome histories, they are also time-dependent. Consider the 10-period outcome history with a breakthrough in the final period. A receiver evaluating this history at the beginning of the game can look forward to 10 periods in which they incur the per period cost c and a reward at the end. Their utility from this outcome history at the beginning of the game is, therefore, $\delta^{9}B-\sum_{t=0}^{9}\delta^{t}c$, where the reward B is discounted, and there is a cost of c in every period leading up to the breakthrough. Now consider the same outcome history evaluated at the beginning of the 6th period. The costs incurred in the first five periods are now sunk and the reward is closer in time; the receiver's utility is now $\delta^{4}B-\sum_{t=0}^{4}\delta^{t}c$; the reward is closer in time, and there are only 5 more periods in which to incur the cost of working. In general, denote the receiver's payoff from outcome history y^t at the beginning of period $\tau\leq t$ by $u\left(y^{t}\middle|\tau\right)$:

$$u\left(y^{t}\middle|\tau\right)=\begin{cases}\delta^{t-\tau}B-\left(\sum_{s=0}^{t-\tau}\delta^{s}\right)c & if\ S\left(y^{t}\right)=2\\ -\left(\sum_{s=0}^{t-\tau}\delta^{s}\right)c & otherwise\end{cases}$$

When the receiver evaluates an outcome history, they take into account (i) how long they work, and (ii) whether or not a breakthrough occurs. On the other hand, when the sender evaluates an outcome history, they only care about whether or not a breakthrough occurs, since they do not incur the per-period cost of working, c. Denote the sender's payoff from outcome history y^t at time $\tau\leq t$ by $v\left(y^{t}\middle|\tau\right)$:

$$v\left(y^t \middle| \tau\right) = \begin{cases} \delta^{t-\tau}B & if\ S\left(y^t\right) = 2 \\ 0 & otherwise \end{cases}$$

This payoff structure is a departure from similar models in the literature, which generally feature flow payoffs that accrue to players over time. In Smolin (2015), the principal and agent receive flow payoffs generated by an underlying state and the agent's actions in each period. In Ely (2017), the players receive flow payoffs determined by an evolving state and the agent's actions in each period. When flow payoffs depend on the state, or a player's beliefs about the state, the problem can be expressed in a recursive manner with these beliefs as state variables. In this paper, both the receiver's as well as the sender's beliefs are relevant state variables. Moreover, the the non-linear evolution of beliefs precipitated by the information released at the start of each period implies that tools used in linear environments cannot be applied here.

2.3 Strategies

It is clear that there is a misalignment of incentives between the sender and receiver. The former would like the receiver to work on the project forever, since even arbitrarily long sequences of failures never convince the sender that the project is bad for sure (as long as $p_0 > 0$). The receiver, on the other hand, would only like to work when the probability that the project is good is sufficiently high, since working is costly. As such, the sender would like to maintain the receiver's beliefs about the quality of the project by sending good news. The receiver interprets the messages they receive with this in mind. In the model we study, the sender must commit to an information disclosure policy at the beginning of the game.[4] This policy specifies the distribution of messages sent to the receiver as a function of the outcome histories and message histories up to that point in the game. Recall that Y^t is the space of outcome histories up to time t. Let M_t denote the message space containing all messages sent by the sender in period t. $M^{t-1} = M_1 \times \ldots \times M_{t-1}$ denotes the space of message histories up to time $t-1$. Then $H^{t-1} = Y^{t-1} \times M^{t-1}$ is the space of histories containing past outcomes as well as past messages. Formally, the sender's strategy, $\sigma = \{\sigma_t\}_{t=1}^{\infty}$, is a sequence of functions mapping histories to probability distributions over messages:

[4] This is similar to the sender choosing a signal in Kamenica and Gentzkow (2011), the principal committing to a feedback policy in Smolin (2015), or the principal committing to an information policy in Ely (2017).

$$\sigma_t : H^{t-1} \to \Delta(M_t),$$

where $\Delta(M_t)$ is the space of probability distributions over an arbitrary message space, M_t. With a slight abuse of notation, we will denote the probability distribution over M_t induced by history h^{t-1} by $\sigma_t(.|\, h^{t-1})$. Knowing σ and observing $m_t \in M_t$ at time t informs the receiver about the likely outcome histories that arose. A perfectly informative disclosure policy, for example, would specify distributions with disjoint support whenever histories are distinct: $\sigma_t(m|\, h^{t-1}) > 0,\ \sigma_t\left(m|\, \hat{h}^{t-1}\right) > 0 \iff h^{t-1} = \hat{h}^{t-1}$. Receiving a message would perfectly reveal the outcome history to the receiver. Denote such a policy by σ^{full}. A disclosure policy that is completely uninformative would specify the same distribution over messages for every history. Denote a policy that provides no information to the receiver by σ^{null}, and let Σ be the space of all disclosure policies.

The receiver's strategy specifies whether or not they choose to work in each period after observing the sender's history of messages. Formally, their strategy, $\beta = \{\beta_t\}_{t=1}^{\infty}$, is a sequence of functions mapping each message history to a probability distribution over their action space:

$$\beta_t : M^t \to \Delta\{work, quit\},$$

Let $\mathcal{B}$ denote the space of all possible strategies, and a_t denote the chosen action in period t.

3 Analysis

3.1 Beliefs

Messages inform the receiver (who does not directly observe outcomes) about outcome histories that are likely to have arisen thus far in the game. Given a disclosure policy, σ, a message $m_t \in M_t$ induces a probability distribution, $d_t(.|\, m_t; \sigma) \in \Delta(Y^{t-1})$, over possible outcome histories via Bayes' rule. This distribution, which represents the receiver's beliefs, depends on σ, the sender's disclosure policy, and the message, m_t, received in period t. The former is the framework through which the receiver interprets messages, and the latter is the new information released as the game progresses. This, in turn, induces beliefs about the quality of the project, and whether an intermediate success has arrived - these derived beliefs turn out to be convenient parameters in the analysis of the game. We denote the receiver's beliefs at the beginning of period t about outcome histories of length $t-1$ *before* they receive a message by $d_t \in \Delta\left(Y^{t-1}\right)$. This is their belief about outcome histories given that the game has not ended in period t, and given all their prior information.

We begin by defining the receiver's beliefs over outcome histories, $d_t\left(.\mid m_t;\sigma\right)$:

$$d_t\left(y^{t-1}\middle| m_t;\sigma\right) = \frac{\mathbb{P}\left(y^{t-1}, m_t;\sigma\right)}{\mathbb{P}\left(m_t;\sigma\right)},$$

where $\mathbb{P}\left(m_t;\sigma\right)$ is the probability that message m_t is sent in period t given that the sender is using disclosure policy σ, and $\mathbb{P}\left(y^{t-1}, m_t;\sigma\right)$ is the probability that message m_t is sent and outcome history y^{t-1} is observed, given that the sender is using disclosure policy σ. The former is a function of the receiver's beliefs about outcome histories at the beginning of period t ($d_t \in \Delta\left(Y^{t-1}\right)$). The expression for $\mathbb{P}\left(m_t;\sigma\right)$ is given below.

$$\mathbb{P}\left(m_t;\sigma\right) = \sum_{y^{t-1}\in Y^{t-1}} \sigma_t\left(m_t\middle| y^{t-1}\right) d_t\left(y^{t-1}\right)$$

The numerator of the expression for $d_t\left(y^{t-1}\mid m_t;\sigma\right)$ can be further expanded into the following:

$$\mathbb{P}\left(y^{t-1}, m_t;\sigma\right) = \mathbb{P}\left(m_t\middle| y^{t-1};\sigma\right) d_t\left(y^{t-1}\right),$$

where $\mathbb{P}\left(m_t\mid y^{t-1};\sigma\right)$ is the conditional probability of observing message m_t following outcome history y^{t-1}. We can now give the expression for $d_t\left(y^{t-1}\mid m_t;\sigma\right)$ in terms of the receiver's beliefs.

$$d_t\left(y^{t-1}\middle| m_t;\sigma\right) = \frac{\sigma_t\left(m_t\mid y^{t-1}\right) d_t\left(y^{t-1}\right)}{\sum\limits_{y^{t-1}\in Y^{t-1}} \sigma_t\left(m_t\mid y^{t-1}\right) d_t\left(y^{t-1}\right)}$$

The receiver is interested in the project's quality, and whether a success has already occurred, conditional on the project being good. For any outcome history, y^{t-1}, there is a corresponding belief about the likelihood that this history was generated by a project of quality ω. It is clear to see that whenever $S\left(y^{t-1}\right) = 1$, it must be the case that $\omega = 1$.[5] However, when $S\left(y^{t-1}\right) = 0$, it may well be that $\omega = 1$, but a success has not yet arrived; it could also be the case that $\omega = 0$. Let the function $\phi : Y^t \to [0,1]$ be the mapping from

[5] We don't need to deal with the case in which $S\left(y^{t-1}\right) = 2$, since those histories are revealed to all players when they occur.

outcome histories (of any length) to beliefs about ω, so that $\phi(y^t)$ is the receiver's beliefs about ω when they know for sure that y^t occurred. For example, when $S(y^t) = 1$, $\phi(y^t) = 1$; on the other hand, when $S(y^t) = 0$, $\phi(y^t) = \frac{p_0(1-\theta)^t}{p_0(1-\theta)^t+1-p_0}$, which is the probability that the project is good and t failures have occurred. Similarly, let the function $\psi : Y^t \to [0,1]$ map outcome histories to beliefs about the arrival of the first success. $S(y^t) = 1$ means that $\psi(y^t) = 1$, and the receiver knows for sure that a success has arrived; $S(y^t) = 0$ implies that $\psi(y^t) = 0$, and the receiver knows for sure that no success has arrived. Interior values for this belief arise when the receiver holds non-degenerate beliefs about outcome histories. Suppose, for example, that the receiver believes that y_0^t arose with probability $1-\rho$ and y_1^t arose with probability ρ, where $S(y_0^t) = 0$, and $S(y_1^t) = 1$. Their beliefs about the number of successes conditional on $\omega = 1$ is then $\rho \times 1 + (1-\rho) \times 0 = \rho$.

Over the course of the game, the receiver's beliefs about outcome histories evolve in response to information about the continuation of the game as well as messages from the sender. In the first period, the fact that the game has not yet ended is uninformative because two successes are required to end the game and only one success can arrive per period. In the absence of any message from the sender, their beliefs about ω remain unchanged from p_0, and their belief that a success has occurred conditional on $\omega = 1$ is simply θ. Denote these beliefs by (p_0, θ) in the space $\Delta(\Omega) \times \Delta(X)$. When the sender plays the no-information strategy, σ^{null}, these beliefs persist until the beginning of the second period. Suppose instead that the sender plays the fully informative strategy, σ^{full}. When $y_1 = 1$, the receiver's beliefs before they observe the sender's message are (p_0, θ) and evolve to $(1, 1)$ after they observe it. When $y_1 = 0$, the receiver's beliefs evolve from (p_0, θ) to $\left(\frac{p_0(1-\theta)}{1-p_0\theta}, 0\right)$. Notice that the receiver's beliefs about ω obey the martingale property, since $p_0 = p_0\theta \times 1 + (1-p_0\theta) \times \frac{p_0(1-\theta)}{1-p_0\theta}$. Their beliefs about the arrival of the first success, conditional on $\omega = 1$ also obey the martingale property, so long as we condition all probabilities on the event $\omega = 1$: $\theta = \theta \times 1 + (1-\theta) \times 0$.

We can now define the mapping from beliefs about outcome histories, $d_t \in \Delta(Y^{t-1})$, to beliefs about the project's type, and whether a success has arrived, conditional on the project being good: $\Delta(\Omega) \times \Delta(X)$. We will use p_t and π_t to denote the receiver's belief that $\omega = 1$, and q_t and μ_t to denote their beliefs that a success has occurred, conditional on $\omega = 1$. The pair (p_t, q_t) will represent their beliefs after they learn that the game has not yet ended in period t, but before they observe the sender's message. Meanwhile, the pair (π_t, μ_t) will represent their beliefs after they observe the sender's period-t message. Given any distribution over outcome histories in Y^{t-1}, the pair of functions (ϕ, ψ) derive a pair of beliefs in $\Delta(\Omega) \times \Delta(X)$ induced by this distribution. With some abuse of notation, let $\phi(d_t)$

denote the beliefs about ω induced by the probability distribution over outcome histories, $d_t \in \Delta\left(Y^{t-1}\right)$:

$$\phi\left(d_t\right) = \sum_{y^{t-1} \in Y^{t-1}} \phi\left(y^{t-1}\right) d_t\left(y^{t-1}\right)$$

Similarly, let $\psi\left(d_t\right)$ denote the beliefs about the number of successes achieved, conditional on $\omega = 1$ when the receiver's beliefs about outcome histories is d_t:

$$\psi\left(d_t\right) = \sum_{y^{t-1} \in Y^{t-1}} \psi\left(y^{t-1}\right) d_t\left(y^{t-1}\right)$$

We summarize these observations below. A pair of disclosure policy and message, (σ, m_t), induce a belief distribution over outcome histories, which, in turn, induces beliefs about the quality of the project and whether a success has arrived.[6]

$$\begin{tikzcd}
(\sigma, m_t) \in \Sigma \times M_t \arrow[d, "d"] \\
d_t\left(.\mid m_t; \sigma\right) \in \Delta\left(Y^{t-1}\right) \arrow[d, "\phi"', "\psi"] \\
(\pi_t, \mu_t) \in \Delta\left(\Omega\right) \times \Delta\left(X\right)
\end{tikzcd}$$

The belief pairs (p_t, q_t) and (π_t, μ_t) are key parameters in the model and will feature prominently in the analysis that follows. We refer to (p_t, q_t) as the receiver's *period-t prior beliefs* and (π_t, μ_t) as their *period-t posterior beliefs*. Figure 3 depicts the evolution of the receiver's beliefs within a given period.

[6]The message need not be informative. If $m_t = \emptyset$, for example, then $d_t\left(.\mid m_t; \sigma\right) = d_t$, which are the receiver's beliefs after learning that the game has not yet ended, but before receiving a message from the sender.

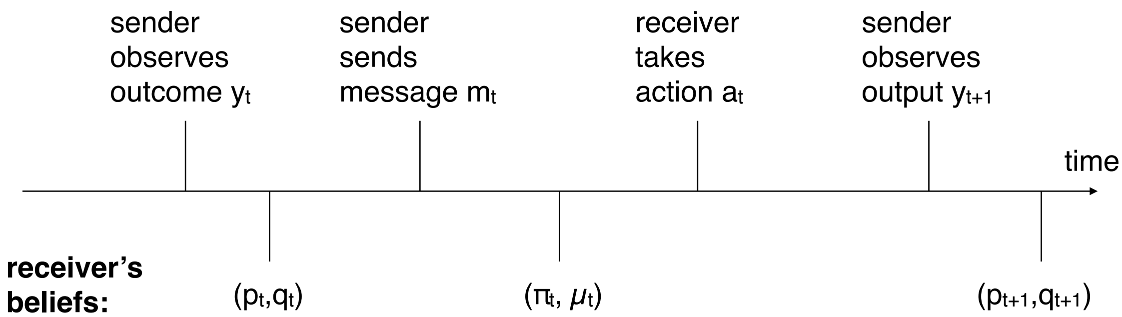


Figure 3: Within-period actions and beliefs

There are two benchmark disclosure policies briefly mentioned earlier: σ^{null}, the no-information disclosure policy, and σ^{full}, the full information disclosure policy. These two policies represent the worst and best payoffs, respectively, for the receiver. Whereas the full information policy yields the lowest expected payoff to the sender, the no-information policy is not their most preferred. In the following subsection, we describe how the receiver's beliefs evolve over time under the two policies.

3.2 Expected payoffs

A pair of strategies, $(\sigma, \beta) \in \Sigma \times \mathcal{B}$, along with a player's beliefs about the project induce a probability distribution over future outcome histories. The collection of possible outcome histories from the beginning of period τ onwards is $Y_\tau = \bigcup_{t=\tau}^{\infty} Y^t$. The receiver's expected payoff at time τ, given the strategy profile (σ, β), can be expressed as follows:

$$U\left(\sigma, \beta \middle| \tau\right) = \mathbb{E}_{\sigma,\beta}\left[u\left(\tilde{y} \middle| \tau\right) \middle| \left\{\sigma_t\right\}_{t=1}^{\tau-1}\right],$$

where $\tilde{y} \in Y_\tau$ is random, and the expectation is taken with respect to the distribution over Y_τ determined by the receiver's beliefs and the strategy pair (σ, β). The receiver's beliefs at time τ are determined by the sender's disclosure policy up to time τ. Similarly, the sender's expected payoff at time τ is $V\left(\sigma, \beta \middle| \tau\right)$:

$$V\left(\sigma, \beta \middle| \tau\right) = \mathbb{E}_{\sigma,\beta}\left[v\left(\tilde{y} \middle| \tau\right) \middle| \left\{y_t\right\}_{t=1}^{\tau-1}\right],$$

where the sender's beliefs are determined by the actual outcome realizations. Players' expected payoffs at the beginning of the game will be expressed by $U\left(\sigma, \beta \middle| 0\right)$ and $V\left(\sigma, \beta \middle| 0\right)$.

Although the sender can use an unrestricted message space, they ultimately care about whether or not the receiver chooses to work. Moreover, since the sender's strategy is known to the receiver, every message sent induces a unique choice of action by the receiver.[7]

As such, the message space may as well be the receiver's action space. This insight was leveraged by Kamenica and Gentzkow (2011) and Smolin (2015) to define and consider *straightforward signals* and *recommendation policies*, respectively. We define recommendation policies in our context as those policies with codomain equal to probability distributions over the receiver's action space.

Definition 1. A disclosure policy, $\sigma = \{\sigma_t\}_{t=1}^{\infty}$, is a **recommendation policy** if:

$$\sigma_t : H^{t-1} \to \Delta\left(\{work,\ quit\}\right)$$

Considering these policies is without loss of generality, as the next result will show. By considering only those policies that recommend an action for the receiver to take, we can reduce the space of disclosure policies over which the sender maximizes, simplifying their decision problem.

Proposition 1. *Consider the pair of disclosure policy and strategy* (σ, β)*. There exists a recommendation policy,* R^{σ}*, such that:*

$$V\left(\sigma, \beta|\, t\right) = V\left(R^{\sigma}, \beta|\, t\right)\ \ \forall t \in \mathbb{N}$$

$$U\left(\sigma, \beta|\, t\right) = U\left(R^{\sigma}, \beta|\, t\right)\ \ \forall t \in \mathbb{N}$$

Proof. See subsection A.1. □

In addition to the simplification afforded by proposition 1, we can also show that recommendation policies need only depend on the outcome history. As the receiver learns about outcome histories through the sender's messages over time, they update their beliefs about the state. Often, models in which a player learns about a parameter associated with the productivity, quality, type of a project or bandit arm reduce to optimal stopping problems. The optimal strategy involves tracking some parameter, usually the player's beliefs, which triggers the player to stop investing, experimenting, or working when it falls below some threshold.[8] Since our framework bears many similarities with these models, optimal strate-

[7]We assume that ties are broken in favor of working, a technical assumption that does not drive any of the results.

[8]See, for example, Smolin (2015) among many others.

gies in our setting also reduce to cut-off strategies, but only in the later stage of the game. We define a generalization of cut-off strategies, which we call *eventually cut-off strategies*. The relevant belief in our case is the receiver's belief that a breakthrough will occur in the next period. In the beginning of the game, the receiver knows that not enough time has elapsed for a breakthrough to occur, even if they are fairly confident that the project is good. In the extreme, they are sure that a breakthrough will never occur in the first period because two successes are required, and a maximum of one can arrive in one period. In the second period, their beliefs that a breakthrough will arrive can be positive but small. In the third period, beliefs will be positive and larger still etc. Once a sufficient amount of time has elapsed, their beliefs about the quality of the project begin to deteriorate, and their belief that a breakthrough will occur in the next period decreases. Under some disclosure policies, beliefs decline monotonically for the remainder of the game. In this phase of the game, they find it optimal to employ a cut-off strategy (see Proposition 2).

Definition 2. A strategy, $\beta \in \mathcal{B}$, is **eventually a cut-off strategy** if there exists some $T \in \mathbb{N}$ and $b \in [0,1]$ such that:

$$\beta_t(m_t) = \begin{cases} work & \forall t < T \\ work & if\ t \geq T,\ \pi_t \mu_t \theta \geq b \\ quit & if\ t \geq T,\ \pi_t \mu_t \theta < b \end{cases}$$

The receiver works in every period $t < T$, and for all $t \geq T$, they choose to work if and only if they believe that a breakthrough will occur in the next period with probability greater than b.

Recall from subsection 3.4 that the receiver's subjective belief that a breakthrough will occur in the current period is $\pi_t \mu_t \theta$. This belief decreases monotonically as long as $\pi_t \mu_t \geq \pi_{t+1} \mu_{t+1}$ for all remaining t. Notice that it is always the case that $\mu_{t+1} > \mu_t$, since the longer the receiver stays in the game, the more likely it is that a success has arrived, conditional on the project being good. However, μ_t is bounded above by 1, and over time, in the absence of good news from the sender, π_t tends towards 0. This moves the receiver's belief about the probability of a breakthrough downwards.

Proposition 2. *Suppose the sender chooses a recommendation policy R^σ. If $\{(\pi_t, \mu_t)\}_{t=1}^{\infty}$ is the sequence of the receiver's posterior beliefs after they receive the message "work" and $\pi_t \mu_t \geq \pi_{t+1} \mu_{t+1}$ $\forall t \geq T$ for some $T \in \mathbb{N}$, then a strategy that is eventually a cutoff strategy is optimal for the receiver.*

Proof. See subsection A.2. $\square$

Every disclosure policy chosen by the sender induces a sequence of posterior beliefs for the receiver. The receiver evaluates whether (and for how long) to work by using these beliefs to calculate their expected payoff. For a pair of posterior beliefs (π, μ) such that π is the receiver's belief about ω and μ is their belief about whether a success has already occurred, we refer to the quantity $\pi\mu\theta B - c$ as the *receiver's flow payoff*. This is the expected payoff the receiver anticipates less the cost of working for an additional period. The first term is the probability of a breakthrough arriving in the next period $(\pi\mu\theta)$ multiplied by the benefit (B). When the flow payoff is positive, it is myopically optimal for the receiver to work for an additional period. When the flow payoff is negative, a myopic receiver would quit, since working for one period costs strictly more than they expect to gain. However, a forward looking receiver may stay if the sum of future flow payoffs is positive. At the beginning of the game, this consideration determines whether the receiver begins working on the project or not. Let $\{(\pi_t, \mu_t)\}$ be the sequence of posterior beliefs induced by the sender's disclosure policy. If there is some T such that:

$$-c + (\pi_1\mu_1\theta B - c) + \sum_{t=2}^{T}\prod_{s=1}^{t-1}(1 - \pi_s\mu_s\theta)(\pi_t\mu_t\theta B - c) \geq 0,$$

the receiver chooses to work until period T.

Before we present the main result of this section, let $\bar{T}$ be the time at which the receiver chooses to quit when they receive no information from the sender. Recall that under $\bar{\sigma}$, the receiver's beliefs about the project deteriorate monotonically over time (see figure 4). Their beliefs about the arrival of a breakthrough in the next period $(\pi_t\mu_t\theta)$ also decrease monotonically beyond some point in time. In this phase of the game, they employ a cut-off strategy and quit whenever the benefit they can expect to gain in the next period $(\pi_t\mu_t\theta B)$ no longer exceeds the cost of working. Since the probability of a breakthrough is decreasing, if this inequality fails in any period, it also fails in all future periods. As a result, the receiver quits in the first period in which this inequality fails. We call this period, $\bar{T}$ and define it formally below. When the receiver does not learn any new information, their posterior beliefs are equivalent to their priors: $\pi_t = p_t$ and $\mu_t = q_t$.

$$\bar{T} = \min_t \left\{ t : p_s q_s < \frac{c}{\theta B}\ \forall s \geq t;\ \sigma = \sigma^{null} \right\}$$

To ensure that the receiver begins working on the project even under a no information policy, we make the following assumption:

Assumption 1. *The receiver always chooses to start working on the project:*

$$-c + p_1 q_1 \theta B - c + \sum_{t=2}^{\bar{T}-1} \prod_{s=1}^{t-1} \left(1 - p_s q_s \theta\right) \left(p_t q_t \theta B - c\right) \geq 0$$

3.3 Evolution of beliefs under σ^{null} and σ^{full}

Under σ^{null}, the only source of information about ω and the number of successes achieved comes from the fact that a breakthrough has not yet occurred. In the first period, there is no new information, since a breakthrough never occurs after only one success. The receiver's beliefs about ω remain at their initial level, p_0. However, in the second period, if the game has not yet ended, the receiver can begin to draw inferences about the project's quality. Let the number of successes achieved through time t be X_t. The probability that the project is good, conditional on having less than two successes by the second period is:

$$p_2 = \mathbb{P}\left(\omega = 1 \middle|\, X_t < 2\right) = \frac{\mathbb{P}\left(\omega = 1, X_t < 2\right)}{\mathbb{P}\left(X_t < 2\right)} = \frac{\left(1 - \theta^2\right) p_0}{\left(1 - \theta^2\right) p_0 + 1 - p_0},$$

where $(1 - \theta^2)$ is the probability that two successes have not yet occurred, given that the project is good. When the project is good, the number of successes up to time t follows a Binomial distribution with t draws and success probability θ: $X_t | \, \omega = 1 \sim B\left(t, \theta\right)$. Let p_t denote the receiver's beliefs that $\omega = 1$ at time t, when the game has not yet ended:

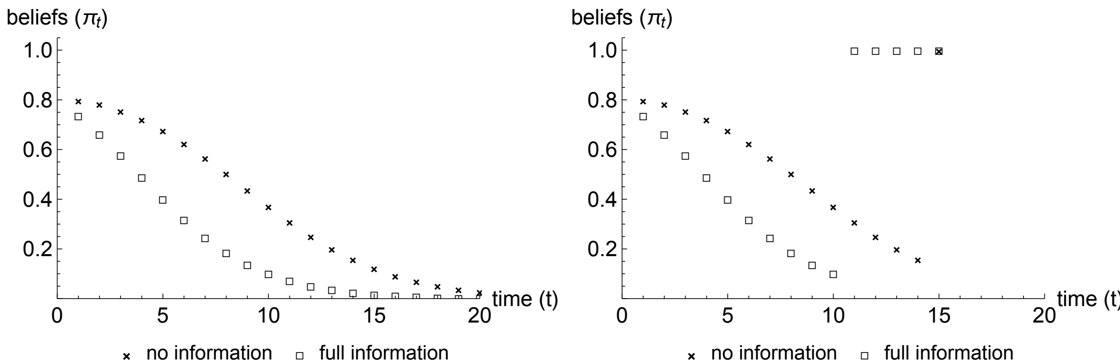
$$p_t = \mathbb{P}\left(\omega = 1 \middle|\, X_t < 2\right) = \frac{\mathbb{P}\left(\omega = 1, X_t < 2\right)}{\mathbb{P}\left(X_t < 2\right)}.$$

Since $\mathbb{P}\left(X_t < 2\right) \to 0$ as $t \to \infty$, $p_t \to 0$, and the receiver becomes increasingly pessimistic about the quality of the project the longer they stay in the game.

Now consider the case in which the sender reveals all information to the receiver (σ^{full}). This differs from the preceding case in that now the receiver knows when the first success arrives, not just when the second success ends the game. As such, beliefs evolve differently: in the absence of good news, the receiver becomes more pessimistic about the quality of the project. When they do see the first success, the receiver's beliefs jump to 1. Denote the receiver's period-t beliefs when they observe no successes by $\hat{p}_t$.

$$\hat{p}_t = \mathbb{P}\left(\omega = 1 \middle|\, X_t < 1\right) = \frac{\mathbb{P}\left(X_t = 0\right) p_0}{\mathbb{P}\left(X_t = 0\right) p_0 + 1 - p_0}$$

It can be shown that $\hat{p}_t \leq p_t$ for all t for histories in which no successes arrive. However, any history in which a success arrives induces beliefs $\hat{p}_t$ to jump to one, while beliefs p_t continue to deteriorate. Below, we depict the evolution of beliefs under no information and full information for two outcome histories to illustrate this point.



(a) No successes.

(b) Successes in the 11th and 15th periods.

Figure 4: The receiver's beliefs for two outcome histories under no information (σ^{null}) and full information (σ^{full}) when the prior belief that $\omega = 1$ is $p_0 = 0.8$, and the probability of success in each period when the receiver works on a good project is $\theta = 0.3$.

3.4 Probability of a breakthrough

The sender's period-t message induces period-t posterior beliefs (π_t, μ_t) about $\omega \in \Omega$ (the quality of the project) and $x \in X$ (the number of successes conditional on the project being good). Given these beliefs, the receiver can calculate the probability of achieving a breakthrough in the current period if they choose to work. The probability of a breakthrough in period t depends on their beliefs (π_t, μ_t), since a breakthrough only occurs when (i) the project is good, and (ii) one success has already occurred in the past. Indeed, it is simply the probability that the project is good, π_t, multiplied by the probability that a success has already occurred, conditional on the project being good, μ_t, multiplied by the probability that another success arrives this period, conditional on the project being good, θ: $\pi_t \times \mu_t \times \theta$. When the receiver does not observe an informative message in period t, their prior beliefs are unchanged, and hence their belief that a breakthrough will occur in the current period is $p_t q_t \theta$ instead.

Consider the mechanism design problem in which the sender commits to a disclosure policy and the receiver best responds. One class of disclosure policies, which we refer to as *promise policies*, consists of the following two stages: (i) a phase of no information disclosure, then (ii) a promise by the sender to reveal some information after a certain amount of time has elapsed. When the sender reveals all of their information at the promised time, we call the policy a *perfect promise policy*; otherwise, it is a *partial promise policy.*

Another focal policy we analyze is the *indifference policy*: the sender conceals all information until the receiver is at the cusp of quitting, at which point they begin releasing information randomly with just enough probability to keep the receiver indifferent between staying and quitting. These two policies are described below.

3.5 Indifference policy

Under the indifference policy, the sender conceals all information (by always recommending *work*) until time $\bar{T}$, after which point they (i) always recommend *work* after outcome histories with one success, and (ii) sometimes recommend *work* after outcome histories with no successes. Let $1 - r_t$ be the probability with which the sender recommends work following a $t-1$ period outcome history with no success. The probability r_t is chosen such that the receiver is just indifference between working and quitting when they receive the message *work*.

3.5.1 Payoffs from the indifference policy

Notice that since the sender always recommends to the receiver that they should work following histories with one success, their continuation value is B following any such history. An argument for this statement is provided below.

Proposition 3. *Under the indifference policy, the sender's value is B following any history with a success.*

Proof. See subsection A.3. □

Let R^* be the indifference policy and β^* the receiver's best response. There is always some positive probability that the receiver keeps working. This probability rapidly approaches 0 as time progresses. Nevertheless, this means that the sender's value is composed of an infinite sum. We describe it in full below, then provide upper and lower bounds that become arbitrarily tight the more terms we include. The sender's value at time $\bar{T}$ from the indifference policy is described below:

$$V\left(R^*, \beta^* | \bar{T}\right) = (1 - r_{\bar{T}})\left(B\theta b_{\bar{T}} + (1 - \theta b_{\bar{T}})\, V\left(R^*, \beta^* | \bar{T} + 1\right)\right)$$

The sender's value at the beginning of the game is below:

$$V\left(R^*,\beta^*|\,0\right)=p_0\mathbb{P}\left(X_{\bar{T}-1}>0\right)B+\left[p_0\mathbb{P}\left(X_{\bar{T}-1}=0\right)+1-p_0\right]V\left(R^*,\beta^*|\,\bar{T}\right),$$

where $X_{\bar{T}-1}$ follows a binomial distribution with $\bar{T}-1$ trials and success probability θ. The expression above captures the idea that, since the receiver stays for sure until $\bar{T}$, the sender can look forward to $\bar{T}-1$ trials that may yield a success. Beyond period $\bar{T}$, the sender's value is described above and depends on the sequence of revelation probabilities, $\{r_t\}_{t=\bar{T}}^{\infty}$. Since all policies considered by the sender involve an initial $\bar{T}$ period phase of no-information, this does not play a role in our comparison between indifference and promise policies below.

The receiver's value from the indifference policy depends on the revelation probabilities chosen by the sender, $\{r_t\}_{t=\bar{T}}^{\infty}$. Let $U\left(R^*,\beta^*|\,\bar{T}\right)$ be their continuation value at the beginning of period $\bar{T}$. Since the sender chooses these probabilities to ensure that the receiver is just indifferent between staying and quitting, his continuation value is equal to zero:

$$U\left(R^*,\beta^*|\,\bar{T}\right)=\left(1-r_{\bar{T}}\left(1-p_{\bar{T}}q_{\bar{T}}\right)\right)\left[\pi_{\bar{T}}\mu_{\bar{T}}\theta B-c+\left(1-\pi_{\bar{T}}\mu_{\bar{T}}\theta\right)U\left(R^*,\beta^*|\,\bar{T}+1\right)\right]=0$$

In fact, this is true in every period beyond $\bar{T}$. The next result shows that when the receiver plays an eventually cutoff strategy with $T=\bar{T}$, it is optimal for the sender to choose the indifference policy.

Theorem 1. *Suppose assumption 1 holds, and that the receiver chooses an eventually cutoff strategy with $T=\bar{T}$. The recommendation policy, R^* is optimal for the sender:*

$$R_t^*\left(y^{t-1}\right)=\begin{cases}work & if\ t<\bar{T}\\ work & if\ t\geq\bar{T}\ and\ S\left(y^{t-1}\right)=1\\ quit\ with\ probability\ r_t=\frac{c-p_tq_t\theta B}{c(1-p_tq_t)} & if\ t\geq\bar{T}\ and\ S\left(y^{t-1}\right)=0\\ work\ with\ probability\ \left(1-r_t\right) & if\ t\geq\bar{T}\ and\ S\left(y^{t-1}\right)=0\end{cases}$$

$$\beta_t^*\left(m_t\right)=\begin{cases}work & if\ m_t=work\\ quit & if\ m_t=quit\end{cases}$$

Proof. See subsection A.5. □

When this restriction on the receiver's strategies is relaxed, the indifference policy may no longer be optimal. For example, if the receiver can choose a voluntary initial period larger than $\bar{T}$, the sender can benefit by delaying information. To study this, we introduce promise policies in the next section.

3.6 Promise policies

3.6.1 Payoffs from a promise policy

Consider a promise policy that the sender proposes to the receiver at the beginning of period t, in which the sender promises to reveal the entire history of outcomes at time $T > t$. Denote the sender's beliefs about ω by b_t, and the receiver's period-t prior beliefs by (p_t, q_t). If the receiver chooses to stay until period T, the values of the players are described as follows, where P^T is the perfect promise strategy that reveals all information at time T, and β^T is the receiver's best response:

$$U\left(P^T, \beta^T \middle| t\right) = (p_t q_t \theta B - c) + (1 - p_t q_t \theta)\left(p_{t+1} q_{t+1} \theta B - c\right) + \ldots + \prod_{s=t}^{T-1} (1 - p_s q_s \theta)\, p_T q_T \left(B - \frac{c}{\theta}\right)$$

$$V\left(P^T, \beta^T \middle| t\right) = b_t \theta B + (1 - b_t \theta)\, b_{t+1} \theta B + \ldots + \prod_{s=t}^{T-2} (1 - b_s \theta)\, b_{T-1} \theta B$$

3.6.2 Incentive compatibility

At the time the sender offers a promise policy to the receiver, the receiver must decide whether or not to accept. This decision involves not only deciding whether or not to work in the period of the offer, it also involves deciding whether or not to work in every subsequent period before receiving the information from the sender. The sender's recommendation that the receiver should work in every period before the one in which they reveal the history of outcomes must be incentive compatible if the receiver is expected to follow through. The next result shows that if the receiver prefers to follow the recommendations of the promise strategy at time t, then they also prefer to do so at every subsequent time before the final period.

Proposition 4. *Let P^T be a promise policy proposed by the sender to the receiver at time $t > \bar{T}$. If $U_s\left(P^T\right) \geq 0$, then $U_{s+1}\left(P^T\right) \geq 0$ for all $t \leq s < T$.*

Proof. See subsection A.4. □

3.7 Conditions under which information delay is optimal

Sender

In general, the sender prefers to delay revealing information to the receiver, since a negative message delivered earlier in the game hastens the receiver's departure. We ask the following question: for an ϵ decrease in revelation probability today, what is the maximum δ increase in revelation probability tomorrow that the sender is willing to give, assuming the receiver continues obeying the recommendations? Let the sender's value at time t from a status-quo revelation policy (with revelation probabilities r_t and r_{t+1}) be V_t, and let their value from the new $\epsilon - \delta$ perturbed policy be $\hat{V}_t$.

$$
\begin{aligned}
V_t &= (1 - r_t)\,\theta b_t B + \ldots \\
&\quad (1 - r_t)\,(1 - r_{t+1})\,(1 - \theta b_t)\,\theta b_{t+1} B + \ldots \\
&\quad (1 - r_t)\,(1 - r_{t+1})\,(1 - \theta b_t)\,(1 - \theta b_{t+1})\,V_{t+2} \\
\hat{V}_t &= (1 - r_t + \epsilon)\,\theta b_t B + \ldots \\
&\quad (1 - r_t + \epsilon)\,(1 - r_{t+1} - \delta)\,(1 - \theta b_t)\,\theta b_{t+1} B + \ldots \\
&\quad (1 - r_t + \epsilon)\,(1 - r_{t+1} - \delta)\,(1 - \theta b_t)\,(1 - \theta b_{t+1})\,\hat{V}_{t+2}
\end{aligned}
$$

Since we assume that the only differences between the two disclosure policies occur in periods t and $t+1$, it follows that the continuation value at time $t+2$ is the same. A sufficient condition for the perturbed policy to be preferred by the sender is for the following condition to hold: $(1 - r_t + \epsilon)\,(1 - r_{t+1} - \delta) \geq (1 - r_t)\,(1 - r_{t+1})$. This inequality can be rearranged into the expression in condition 1 for δ in terms of ϵ, r_t and r_{t+1}.

Condition 1. Sufficient condition for the sender to prefer decreasing the revelation probability in period t by ϵ and increasing it by δ in period $t+1$:

$$
\delta \leq \hat{\hat{\delta}}_{max} = \frac{\epsilon\,(1 - r_{t+1})}{1 - r_t + \epsilon}
$$

In general, the sender's period-$t+2$ values may differ, but it will always be the case that $\hat{V}_{t+2} \geq V_{t+2}$. To see why, recall that the sender must compensate the receiver for lower flow payoffs in period t by increasing the revelation probability in period $t+1$ and, therefore, increasing the receiver's period-$t+1$ posterior beliefs beyond their value under the status quo.

While the condition on δ is sufficient, it is not necessary. Decreasing revelation probability at time t by ϵ strictly increases the sender's flow payoff in period t. This strict increase can allow for the condition we require to be violated and for the sender to still prefer the perturbed policy to the status quo.

Receiver

Now consider the receiver's willingness to accept a decrease in revelation probability by ϵ in period t in exchange for an increase in period $t+1$ revelation probability by δ. A decrease in r_t decreases the receiver's posterior beliefs about the prospect of a breakthrough in period t, while an increase in r_{t+1} increases these beliefs. As such, the receiver's flow payoffs are lower in period t and higher in period $t+1$. Let their period-t prior beliefs be (p, q) and their value at time t from the status quo policy and the perturbed policy be U_t and $\hat{U}_t$, respectively.

$$\begin{aligned} &U_t\left(work|\, m_t = work; R, \beta\right) = \left(\pi\mu\theta B - c\right) + \ldots \\ &\ldots + \left(1 - \pi\mu\theta\right)\left(1 - r_{t+1}\left(1 - p'q'\right)\right)\left(\pi'\mu'\theta B - c + \left(1 - \pi'\mu'\theta\right)U_{t+2}\right) \end{aligned}$$

where (π, μ) are the receiver's period-t posterior beliefs when their prior beliefs are (p, q) and the revelation probability is r_t; (p', q') are the receiver's period-$t+1$ prior beliefs, and (π', μ') are their period-$t+1$ posterior beliefs.

$$\begin{aligned} &\hat{U}_t\left(work|\, m_t = work; \hat{R}, \hat{\beta}\right) = \left(\hat{\pi}\hat{\mu}\theta B - c\right) + \ldots \\ &\ldots + \left(1 - \hat{\pi}\hat{\mu}\theta\right)\left(1 - \left(r_{t+1} + \delta\right)\left(1 - \hat{p}'\hat{q}'\right)\right)\left(\hat{\pi}'\hat{\mu}'\theta B - c + \left(1 - \hat{\pi}'\hat{\mu}'\theta\right)\hat{U}_{t+2}\right) \end{aligned}$$

where terms with hats reflect the fact that under the perturbed policy, the evolution of the receiver's beliefs differs from their evolution under the status quo. Since the receiver's value at time $t+2$ is determined by their period-$t+2$ prior beliefs, and these, in turn, are determined by the revelation probabilities in time t and $t+1$, $\hat{U}_{t+2}$ may not be equal to U_{t+2}. Unlike the sender's, the receiver's beliefs (and, hence, their values) are affected by the revelation probabilities. For any $\epsilon > 0$, the corresponding δ that keeps the receiver's period t value from working weakly greater than the status quo value must raise their period-$t+1$ posterior beliefs about the prospects of a breakthrough above the status quo level. This is necessary to compensate the receiver for the lower flow payoff in period t under the perturbed strategy. In addition to generating higher flow payoffs in period $t+1$, higher period-$t+1$

posterior beliefs also induce high period -$t+2$ prior beliefs, which increase the receiver's value. As such, $\hat{U}_{t+2} \geq U_{t+2}$. In general, when the value of δ is chosen such that $U_t = \hat{U}_t$, say, it is possible to choose a value that increases $\hat{U}_t$ through a mixture of (i) higher period-$t+1$ flow payoffs, as well as (ii) a higher value of $\hat{U}_{t+2}$. It is sufficient to require the increase in period-$t+1$ revelation probability, δ, to guarantee that $\hat{U}_t \geq U_t$ through the flow payoffs in periods t and $t+1$ alone, because the receiver's period-$t+2$ continuation value will, by the previous argument, automatically be higher. For any decrease in period-t revelation probability, an increase in period-$t+1$ revelation at least as large as δ guarantees that the receiver's value at time t is at least as high as their value under the status quo policy.

Condition 2. Sufficient condition for the receiver to accept ϵ lower revelation probability in period t and δ higher probability in period $t+1$:

$$\delta \geq \hat{\delta}_{min} = \frac{\epsilon pq}{1-r_t+\epsilon}\left[\frac{B\theta\left[2\left(1-pq\left(1-\theta\right)\right)-\theta\left(1+p\right)\right]+c\left[\left(1-pq\right)\theta-\left(1-p\left(q\left(1-\theta\right)+\theta\right)\right)r_{t+1}\right]}{c\left(1-\left(1-pq\right)r_t\right)\left(1-p\left(q\left(1-\theta\right)+\theta\right)\right)}\right]$$

When the δ required to keep the receiver's payoff sufficiently high is lower than the maximal δ the sender is willing to offer, the sender can improve their payoff by decreasing the revelation probability. This is guaranteed to hold whenever $\hat{\delta}_{min} \leq \hat{\hat{\delta}}_{max}$.

3.8 Examples

A priori, it is not clear whether the sender earns a higher payoff from the indifference policy or from the maximal promise policy - the one in which they keep the receiver in the game for the longest time possible. The examples below show that for some parameter values, the sender can do strictly better by adopting the indifference strategy, and for others, they can do better by adopting the maximal promise strategy. Recall that any policy used by the sender involves an initial $\bar{T}$ period with no information sharing. The indifference policy differs from promise policies beyond this initial phase. Therefore, we consider the sender's value beyond this initial point in the comparison below. The sender's payoff from the indifference policy involves an infinite number of terms, since the receiver remains in the game with positive probability in every period. Below, we derive bounds on their value beyond period $\bar{T}$:

$$\left(1-r_t\right)\left[\theta b_t B+\left(1-\theta b_t\right)\left(1-r_{t+1}\right)\theta b_{t+1}B\right] \leq V\left(R^*,\beta^*\middle|\, t\right) \leq \frac{\left(1-r_t\right)\theta b_t B}{1-\left(1-r_t\right)\left(1-\theta b_t\right)}$$

These inequalities follow from the fact that $b_t \geq b_{t+1}$, and $r_t \leq r_{t+1}$ for all t. The sender's value from promise policies involve multiple stages of computation: first, we find the maximal time the receiver is willing to stay before receiving information. Given this time, we calculate the sender's expected payoff from the receiver remaining in the game until that time. When $B = 4$, $\theta = 0.4$, $c = 0.4$ and $p_0 = 0.8$, the sender conceals all information for the first 8 periods of the game. After this point, they can use a maximal promise policy to keep the receiver in the game for an additional period after which they fully reveal the outcome history. The receiver is only willing to stay in the game for one more period beyond the voluntary 8 during which they stay without receiving any information. However, the sender can do strictly better by using the indifference policy. When, on the hand, $\theta = 0.35$, the receiver still stays in the game for 8 periods without receiving any information, but the sender can now use a maximal promise policy to keep them in the game for an additional 3 periods beyond the voluntary phase. They now prefer the maximal promise policy to the indifference policy. An increase in θ has two effects: the probability of a success in a given period is now higher, conditional on the project being good. However, in the absence of good news, beliefs deteriorate faster. Lower values of θ make players more lenient on outcome histories. It may well be the case that the project is good but a success is yet to arrive.

4 Commitment

In this section, we consider the game in which the receiver has commitment power, and both players choose their strategies at the beginning of play. The sender chooses a disclosure policy, and the receiver chooses a mapping from disclosure policies to strategies in $\mathcal{B}$, simultaneously. To distinguish between these strategies and those described above, we refer to them as *commitment strategies.* A commitment strategy, C, is a mapping $C : \Sigma \to \mathcal{B}$, and the space of all such strategies is $\mathcal{C}$.

The strategies defined in subsection 2.3 map the sender's messages to distributions over actions. When the players commit to their strategies at the beginning of the game, messages can have different meanings, depending on the disclosure policy in use by the sender (recall that beliefs are induced by the pair of disclosure policy and message). The receiver's decisions are determined by their beliefs. When the disclosure policy is known in advance (in the mechanism design problem analyzed earlier), messages alone determine the receiver's beliefs. When the disclosure policy is not known, however, the pair of disclosure policy and message together determine their beliefs. Both *strategies* as well as *commitment strategies* are equivalent to mappings from beliefs to distributions over actions.

A pair, (σ^*, C^*), is an *equilibrium* if $(\sigma^*, C(\sigma^*))$ is a Nash equilibrium of the simultaneous move game, and the receiver's strategy, $\beta^* = C^*(\sigma^*)$, is sequentially rational on the equilibrium path. Namely, it must be the case that, given that the sender is committed to σ^*, the receiver would not want to deviate from β^* at any point during the game. Conversely, given that the receiver is committed to β^*, the sender would not want to deviate from σ^* at the beginning of the game. Whereas the receiver makes decisions along the path of play, the sender makes all their choices at the beginning of the game, since their disclosure policy determines how messages are interpreted. Changes in messages over time would render this commitment meaningless, and the receiver would not be able to correctly interpret messages.

Definition 3. An **equilibrium** is a pair (σ^*, C^*) such that:

$$U(\sigma^*, C^*(\sigma^*)|\, 0) \geq U(\sigma^*, C(\sigma^*)|\, 0) \;\; \forall\, C \in \mathcal{C},$$

$$V(\sigma^*, C^*(\sigma^*)|\, 0) \geq V(\sigma, C(\sigma)|\, 0) \;\; \forall\, \sigma \in \Sigma,$$

We restrict attention to commitment strategies that induce the same mappings from beliefs to actions for any $\sigma \in \Sigma$. For example, suppose that for some $\sigma_1 \neq \sigma_2 \in \Sigma$, and $C \in \mathcal{C}$, $C(\sigma_1) = \beta_1$ and $C(\sigma_2) = \beta_2$. Let (π_t, μ_t) be the receiver's period-t posterior beliefs under (σ_1, β_1) as well as under (σ_2, β_2) induced by messages m_{1t} and m_{2t}, respectively. If $\beta_1(m_{1t}) \neq \beta_2(m_{2t})$, then C does not induce the same mapping from beliefs to actions for σ_1 and σ_2. Essentially, we restrict attention to commitment strategies that map a receiver's beliefs to action distributions. Call this restricted space $\tilde{\mathcal{C}} \subseteq \mathcal{C}$.

The main result of this section shows that a range of outcomes, including outcomes obtained through promise policies, can be sustained in equilibrium. In particular, the first best (for the receiver) full information outcome can be sustained. Before we state the result, let $\hat{T}$ denote the time at which information is released in the maximal promise policy.

Theorem 2. *Suppose assumptions* **??***, and 1 hold, and suppose that the receiver chooses commitment strategies from $\tilde{\mathcal{C}}$. For every $T \in \left[\bar{T}, \hat{T}\right]$, there exists an equilibrium pair $\left(\beta^T, C^T\right)$ such that the sender conceals all information until time T, and the receiver chooses to stay until time T, at which point they quit if a success has not arrived.*

5 Conclusion

In this paper, I study the role of disclosure policies on the incentives of a receiver who decides when to quit a project. Although the payoffs that accrue to the players from the success of the project are the same, the sender does not incur the cost of working. As a result, she would like the project to continue indefinitely. The receiver, on the other hand, does incur a cost, and only prefers to stay on the project when he is sufficiently optimistic about the project's chances of success. There are two key forces at play: (i) the balance between promising to give the receiver good news, and the resulting bad news generated when this good news does not arrive; and (ii) the tradeoff between releasing some information today, and delaying revelation at the same time as increasing the amount of information released to the agent in order to compensate them for waiting. The first tradeoff appears in many settings in which an agent commits to an information revelation policy used to persuade another agent to take some actions that may not be in line with their interests. The second tradeoff appears in dynamic settings, but often does not play a crucial role. In Ely (2017), for example, the myopic optimal disclosure policy is also an optimal policy when the receiver is patient and strategic. In the above setting, I provide conditions under which, and examples in which, this is not the case. The sender can take advantage of the receiver's patience by promising them more information if they choose to stay in the project longer. Why does the sender benefit from the receiver's patience in this framework? There are two important differences between the model I study here and the one studied in Ely (2017): (i) the sender does not receive flow payoffs in my setting, and (ii) she is uncertain about the quality of the project. These two differences imply that the value to the sender from the receiver staying for one more period is large: she may receive a perfect good news signal about the project in this additional period, which would guarantee her receipt of the lump sum payoff, B. Promising to reveal information in the future allows the sender the opportunity to learn that the project is good, and relay that information to the receiver. These considerations do not arise when the sender knows the state and can choose when to communicate it to the receiver. Other than the flow payoff that accrues, there is no informational value from waiting an additional period.

I study two natural disclosure policies: the indifference policy and the maximal promise policy. I derive conditions under which the sender would prefer to deviate from the indifference policy by delaying revelation and compensating the receiver with more information later on. Through examples, I demonstrate that neither the indifference policy nor the maximal promise policy is optimal for all parameter values. This is surprising because the indifference policy keeps the receiver's continuation value at 0, making his individual rationality constraint bind. On the other hand, the maximal promise policy generates strictly positive continuation payoffs for the receiver: when they choose to stay until some period, T, their continuation value grows as they approach T and they anticipate receiving the promised information.

Many open questions remain. Are there sufficient conditions for the optimality of the indifference policy? This questions is more difficult to answer than whether or not there are conditions for the optimality of the maximal promise policy. This is because continuation payoffs after a deviation in which the sender delays releasing information are strictly positive, hence higher than those under the indifference policy. This gives the sender more room to delay information further, giving the promise policy an advantage over the indifference policy. A sufficient condition must, therefore, take this possibility into account. In general, expressions for the receiver's continuation values are less tractable than those in existing models since their relevant beliefs are two dimensional with non-linear transition rules. This makes studying deviations from the indifference policy more cumbersome. Nevertheless, the findings above point towards an important role that promise policies can play in dynamic persuasion games.

References

Bimpikis, Kostas, Shayan Ehsani, and Mohamed Mostagir (2016). *Designing dynamic contests*. Working Paper (cit. on p. 4).

Calzolari, Giacomo and Alessandro Pavan (Sept. 2006). "On the optimality of privacy in sequential contracting". In: *Journal of Economic Theory* 130.1, pp. 168–204 (cit. on p. 3).

Che, Yeon-Koo and Johannes Horner (Apr. 2015). *Optimal Design for Social Learning*. Working Paper (cit. on p. 4).

Ely, Jeffrey C. (Jan. 2017). "Beeps". In: *American Economic Review* 107.1, pp. 31–53 (cit. on pp. 4, 11, 30).

Green, Brett and Curtis R. Taylor (Dec. 2016). "Breakthroughs, Deadlines, and Self-Reported Progress: Contracting for Multistage Projects". In: *American Economic Review* 106.12, pp. 3660–3699 (cit. on p. 4).

Ivanov, Maxim (Mar. 2010). "Informational control and organizational design". In: *Journal of Economic Theory*. Judgment Aggregation 145.2, pp. 721–751 (cit. on p. 3).

Kamenica, Emir and Matthew Gentzkow (Oct. 2011). "Bayesian Persuasion". In: *American Economic Review* 101.6, pp. 2590–2615 (cit. on pp. 3, 11, 17).

Orlov, Dmitry (Jan. 2015). *Optimal Design of Internal Disclosure*. Working Paper (cit. on p. 4).

Orlov, Dmitry, Andrzej Skrzypacz, and Pavel Zryumov (Feb. 2016). *Persuading the Regulator to Wait*. Working Paper (cit. on p. 4).

Pei, Harry Di (Aug. 2016). *Reputation with Strategic Information Disclosure*. Working Paper (cit. on p. 4).

Smolin, Alex (Nov. 2015). *Optimal Feedback Design*. Working Paper (cit. on pp. 4, 11, 17).

A Proofs

A.1 Proof of Proposition 1

Fix a disclosure policy, $\sigma = \{\sigma_t\}$, such that $\sigma_t\left(.|\, y^{t-1}\right) \in \Delta\left(M_t\right)$. We can partition M_t into Q_t and $M_t \backslash Q_t$ such that the receiver quits whenever they observe a message $m_t \in Q_t$ and chooses to work otherwise. Let $r_t = \int_{Q_t} \sigma_t\left(m_t|\, y^{t-1}\right) dm_t$, and define R_t^σ as follows:

$$R_t^\sigma\left(m|\, y^{t-1}\right) = \begin{cases} r_t & if\ m = quit \\ 1 - r_t & if\ m = work \end{cases}$$

Since the receiver chooses to quit whenever they receive a message $m_t \in Q_t$ under σ, they choose to quit whenever they receive the message *"quit"* under R^σ. Similarly, they choose to work whenever they receive the message *"work"* under R^σ, since working is preferred to quitting whenever they receive a message $m_t \in M_t \backslash Q_t$ under σ. To see why the players' values are the same under (σ, β) as they are under (R^σ, β), notice that the receiver stays in the game after each history with exactly the same probability under both pairs.

□

A.2 Proof of Proposition 2

Let T be the first period such that $\pi_t \mu_t \geq \pi_{t+1}\mu_{t+1}$ for all $t \geq T$. The receiver's expected payoff from T onwards when they choose to work whenever they receive the message "*work*" is the following:

$$\begin{aligned} U\left(R^\sigma, \beta|\, T-1\right) &= \pi_T\mu_T\theta B - c + \left(1 - \pi_T\mu_T\theta\right) U\left(R^\sigma, \beta|\, T\right) \\ &= \pi_T\mu_T\theta B - c + \left(1 - \pi_T\mu_T\theta\right)\left[\pi_{T+1}\mu_{T+1}\theta B - c + \left(1 - \pi_{T+1}\mu_{T+1}\theta\right) U\left(R^\sigma, \beta|\, T+1\right)\right] \\ &= \sum_{t=T}^{\infty} \left(\pi_t\mu_t\theta B - c\right) \prod_{s=1}^{t-T} \left(1 - \pi_{T+s-1}\mu_{T+s-1}\theta\right) \end{aligned}$$

Since $\pi_t\mu_t \geq \pi_{t+1}\mu_{t+1}$, it is clear that once $(\pi_t\mu_t\theta B - c) < 0$ (or $\pi_t\mu_t < \frac{c}{\theta B}$) for some t, the expression is negative for all subsequent periods, and quitting in period t maximizes the receiver's payoff conditional on reaching period T. For any choice rule prior to period T, the receiver maximizes their expected payoff from period T onwards with a cutoff strategy. Hence, their optimal strategy is eventually a cutoff strategy.

$\square$

A.3 Proof of proposition 3

Suppose that a success has occurred. The sender recommends that the receiver works in every period thereafter. Since the policy is incentive compatible, the receiver complies with the recommendation. Only good projects yield successes, which means that the project must be good. As such, in every period, there is a probability of θ that a breakthrough arrives, with a payoff of B. With probability $(1-\theta)$ there is a next period with the same prospects. The sender's value is below:

$$\begin{aligned} V &= \theta B + (1-\theta)\,\theta B + (1-\theta)^2\,\theta B + (1-\theta)^3\,\theta B + \ldots \\ &= \theta B\left(1 + (1-\theta) + (1-\theta)^2 + (1-\theta)^3 + \ldots\right) \\ &= \theta B\left(\frac{1}{1-(1-\theta)}\right) = B \end{aligned}$$

$\square$

A.4 Proof of proposition 4

Suppose that $U_s\left(P^T\right) \geq 0$ for some $t \leq s < T$. We can write this condition, while expanding $U_s\left(P^T\right)$, as follows:

$$U_s\left(P^T\right) = \left(p_s q_s \theta B - c\right) + \left(1 - p_s q_s \theta\right)\left(p_{s+1} q_{s+1} \theta B - c\right) + \ldots + \prod_{\tau=s}^{T-1}\left(1 - p_\tau q_\tau \theta\right) p_T q_T \left(B - \frac{c}{\theta}\right)$$

Since $t > \bar{T}$, every term of the form $\left(p_s q_s \theta B - c\right)$ in the above expression is negative. The expression for $U_{s+1}\left(P^T\right)$ can be written in terms of $U_s\left(P^T\right)$ in the following way:

$$U_{s+1}\left(P^T\right) = \frac{U_s\left(P^T\right) - \left(p_s q_s \theta B - c\right)}{\left(1 - p_s q_s \theta\right)} \geq U_s\left(p^T\right),$$

where the inequality follows from the fact that $\left(p_s q_s \theta B - c\right) \leq 0$ and $\left(1 - p_s q_s \theta\right) < 1$.

$\square$

A.5 Proof of Theorem 1

We will show that (1) given R^*, β^* is (i) sequentially rational on the equilibrium path, and (ii) optimal. (2) Given β^*, R^* is optimal for the sender.

Let b_t be the sender's period t beliefs that the project is good, where $t \geq \bar{T}$. Let p_t be the receiver's belief that $\omega = 1$, and q_t be their belief that there has been one success, conditional on $\omega = 1$, all at the beginning of period t (their period t prior beliefs - see subsection 3.1 for details). We consider the class of disclosure policies in which the sender always recommends to the receiver that they stay when $S\left(y^t\right) = 1$, and recommends that they quit with some probability r_t when $S\left(y^t\right) = 0$. The choice of r_t affects the receiver's period-t posterior beliefs, which are a function of their prior beliefs and the probability r_t in the following way (where we have suppressed time subscripts):

$$\pi\left(p, q, r\right) = \frac{p\left(1 - r\left(1 - q\right)\right)}{1 - r\left(1 - pq\right)}$$

$$\mu\left(q, r\right) = \frac{q}{1 - r\left(1 - q\right)}$$

It can be shown that both π and μ are increasing in r. The reason is that the probability r represents the credibility of the sender's messages, or the extent to which their disclosure policy is in line with the receiver's interests. For example, consider a disclosure policy with $r_t = 1$: this means that whenever the sender observes a history with no successes, they tell the receiver to quit, which is exactly what the receiver would like to do. When $r_t = 1$, and the sender recommends to the receiver that they should stay, the receiver knows for sure that the sender must have observed a success. The receiver's period-t posterior beliefs, therefore, jump to 1. In the other extreme, when $r_t = 0$, the sender always tells the receiver to stay. As a result, the message does not convey any credible information, and the receiver's period-t posterior beliefs are unchanged - they quit even when the sender sends the message *"stay"*. The value of r_t, therefore, simultaneously determines (i) how often the sender recommends that the receiver quit when they observe no successes, and (ii) how likely it is that a success has occurred given that the sender recommends *"stay"*. The sender would like to minimize how often they recommend that the receiver quit (by minimizing r_t), as well as ensuring that their "*stay*" message is credible. This trade-off is captured by the sender's value function from the optimal strategy after histories in which no success has been observed. The value function, $V\left(b, p, q\right)$, takes their beliefs, as well as the prior beliefs of the receiver into account, and can be described as follows:

$$
\begin{aligned}
V(b,p,q) &= \max_{r\in[0,1]}\left\{(1-r)\left[b\theta V(1,\pi(p,q,r),\mu(q,r))+(1-b\theta)V(b',p'(\pi,\mu),q'(\mu))\right]\right\} \\
\textit{subject to} \quad & \pi(p,q,r)\,\mu(q,r)\,\theta B \geq c
\end{aligned}
$$

The constraint ensures that the sender's *"stay"* message is credible - the receiver's period-t posterior beliefs must be such that when they do receive the stay message, they actually prefer to stay. $V(1,\pi,\mu)$ is the sender's value when a success arrives and their beliefs that $\omega=1$ jumps to 1. When a success does not arrive, their beliefs about the project deteriorate in the next period to b':

$$
b' = \frac{(1-\theta)\,b}{1-\theta b}
$$

Recall that the transition rule for receiver's beliefs is the following:

$$
p' = \frac{\pi(1-\mu)+\pi\mu(1-\theta)}{1-\pi\mu\theta}
$$

$$
q' = \frac{(1-\mu)\,\theta+\mu(1-\theta)}{1-\mu\theta}
$$

Namely, whenever the game does not end, the receiver's prior beliefs in the next period are given by (p',q') given that their previous period's posterior beliefs were (π,μ). Suppose the sender plays the strategy σ^* in which they reveal the state to the receiver with probability r_t in period t such that $\pi_t\mu_t\theta B=c$. Namely, the sender reveals just enough information to bring the receiver's beliefs up to the level at which they are indifferent between staying and quitting (we assumed that the receiver stays when they are indifferent). In period t, given a history with no success, the sender's value function is described below:

$$
V_t^* = (1-r_t)\left[\theta b_t V(1,\pi_t,\mu_t)+(1-\theta b_t)(1-r_{t+1})\left[\theta b_{t+1}V(1,\pi_{t+1},\mu_{t+1})+(1-\theta b_{t+1})V_{t+2}\right]\right]
$$

Consider an alternative strategy, $\hat{\sigma}$, in which, whenever a success has not yet occurred, the sender reveals the state with higher probability in period t ($\hat{r}_t > r_t$), pushing the receiver's posterior beliefs above the threshold level required for them to stay. In period $t+1$, the sender reveals the state with smaller probability ($\hat{r}_{t+1} < r_{t+1}$), since the receiver's period-$t+1$prior beliefs are now higher than they would have been had the sender sent the message *"quit"* with probability r_t instead of $\hat{r}_t$. This is beneficial for the sender, since they now send the message *"stay"* with higher probability. The trade-off is the following: the sender must send the message *"stay"* with lower probability in period t in order to send it with higher probability in period $t+1$. The sender's value from this strategy, $\hat{V}$, is described below.

$$\hat{V}_t = (1-\hat{r}_t)\left[\theta q_t V(1,\hat{\pi}_t,\hat{\mu}_t) + (1-\theta q_t)(1-\hat{r}_{t+1})\left[\theta q_{t+1} V(1,\hat{\pi}_{t+1},\hat{\mu}_{t+1}) + (1-\theta q_{t+1}) V_{t+2}\right]\right]$$

First, notice that $\pi_{t+1} = \hat{\pi}_{t+1}$ and $\mu_{t+1} = \hat{\mu}_{t+1}$, since the sender only needs to push the receiver's period-$t+1$ posterior beliefs to the threshold required for them to stay. To determine whether such a deviation is profitable, we calculate the difference $V_t^* - \hat{V}_t$:

$$\begin{aligned}
V_t^* - \hat{V}_t &= \theta q_t\left[(1-r_t) V(1,\pi_t,\mu_t) - (1-\hat{r}_t) V(1,\hat{\pi}_t,\hat{\mu}_t)\right] + \ldots \\
&\quad \ldots + (1-\theta q_t)\,\theta q_{t+1} V(1,\pi_{t+1},\mu_{t+1})\left[(1-r_t)(1-r_{t+1}) - (1-\hat{r}_t)(1-\hat{r}_{t+1})\right] + \ldots \\
&\quad \ldots + (1-\theta q_t)(1-\theta q_{t+1}) V_{t+2}\left[(1-r_t)(1-r_{t+1}) - (1-\hat{r}_t)(1-\hat{r}_{t+1})\right] \\
&= \theta q_t\left[(1-r_t) V(1,\pi_t,\mu_t) - (1-\hat{r}_t) V(1,\hat{\pi}_t,\hat{\mu}_t)\right] + \ldots \\
&\quad \ldots + (1-\theta q_t)\left[(1-r_t)(1-r_{t+1}) - (1-\hat{r}_t)(1-\hat{r}_{t+1})\right] \times \ldots \\
&\quad \ldots \times \left[\theta q_{t+1} V(1,\pi_{t+1},\mu_{t+1}) + (1-\theta q_{t+1}) V_{t+2}\right]
\end{aligned}$$

First, note that the first bracketed expression is positive since $(1-r_t) > (1-\hat{r}_t)$ and $V(1,\pi_t,\mu_t) = V(1,\hat{\pi}_t,\hat{\mu}_t)$. Indeed, it can be shown that $V(1,\pi_t,\mu_t) = V(1,\hat{\pi}_t,\hat{\mu}_t) = V(1,\pi_{t+1},\mu_{t+1}) = B$. To see this, notice that the sender always recommends to the receiver that they should stay after every period with one success. The receiver obeys this recommendation and stays forever, eventually achieving a breakthrough. Since there is no discounting, and a breakthrough arrives with probability 1 whenever the receiver stays forever, the sender's payoff is simply B (see proposition 3 for a proof).

The expression above can be simplified as follows:

$$V_t^* - \hat{V}_t = \theta q_t\left[\hat{r}_t - r_t\right] B + (1-\theta q_t)\left[(1-r_t)(1-r_{t+1}) - (1-\hat{r}_t)(1-\hat{r}_{t+1})\right]\left[\theta q_{t+1} B + (1-\theta q_{t+1}) V_{t+2}\right]$$

Note that r_t is determined by the condition $\pi_t \mu_t \theta B = c$, which is a function of p_t, q_t, θ, B and c. Solving this identity for r_t yields the following expression:

$$r_t = \frac{c - p_t q_t \theta B}{c(1 - p_t q_t)},$$

which is positive since $c > p_t q_t \theta B$ for all $t \geq \bar{T}$. Consider any $\hat{r}_t = r_t + \epsilon$ for some $\epsilon > 0$, and notice that the resulting period-t posterior beliefs are as follows:

$$\begin{aligned}
\hat{\pi}_t\left(p_t, q_t, \hat{r}_t\right) &= \frac{p_t\left(1 - \hat{r}_t\left(1 - q_t\right)\right)}{p_t\left(1 - \hat{r}_t\left(1 - q_t\right)\right) + \left(1 - \hat{r}_t\right)\left(1 - p_t\right)} \\
\hat{\mu}_t\left(q_t, \hat{r}_t\right) &= \frac{q_t}{1 - \hat{r}_t\left(1 - q_t\right)}
\end{aligned}$$

Applying the transition rules for b', p' and q' from above yields the following expressions, where hats denote quantities under the alternative strategy, $\hat{\sigma}$:

$$\begin{aligned}
b_{t+1} &= \frac{\left(1 - \theta\right) b_t}{\left(1 - \theta\right) b_t + 1 - b_t} \\
\hat{p}_{t+1} &= \frac{\hat{\pi}_t\left(1 - \hat{\mu}_t\right) + \hat{\pi}_t \hat{\mu}_t\left(1 - \theta\right)}{\hat{\pi}_t\left(1 - \hat{\mu}_t\right) + \hat{\pi}_t \hat{\mu}_t\left(1 - \theta\right) + 1 - \hat{\pi}_t} \\
\hat{q}_{t+1} &= \frac{\left(1 - \hat{\mu}_t\right)\theta + \hat{\mu}_t\left(1 - \theta\right)}{1 - \hat{\mu}_t \theta}
\end{aligned}$$

Next, $\hat{r}_{t+1}$ chosen such that $\hat{\pi}_{t+1}\hat{\mu}_{t+1}\theta B = c$. Recall that $\hat{\pi}_{t+1}$ is a function of period-$t+1$ prior beliefs, $(\hat{p}_{t+1}, \hat{q}_{t+1})$, as well as $\hat{r}_{t+1}$:

$$\begin{aligned}
\hat{\pi}_{t+1} &= \frac{\hat{p}_{t+1}\left(1 - \hat{r}_{t+1}\left(1 - \hat{q}_{t+1}\right)\right)}{\hat{p}_{t+1}\left(1 - \hat{r}_{t+1}\left(1 - \hat{q}_{t+1}\right)\right) + \left(1 - \hat{r}_{t+1}\right)\left(1 - \hat{p}_{t+1}\right)} \\
\hat{\mu}_{t+1} &= \frac{\hat{q}_{t+1}}{1 - \hat{r}_{t+1}\left(1 - \hat{q}_{t+1}\right)}
\end{aligned}$$

We now have formulae for all the terms in the expression for $V_t^* - \hat{V}_t$ expressed in terms of p_t, q_t, θ, B, c and ϵ. It can be shown that the expression in brackets, which is the only one that may turn out to be negative, simplifies to the following fraction, which is always positive:

$$\left(1 - r_t\right)\left(1 - r_{t+1}\right) - \left(1 - \hat{r}_t\right)\left(1 - \hat{r}_{t+1}\right) = \frac{\epsilon p_t \theta\left(B\theta - c\right)\left(1 - q_t\right)}{c\left(1 - p_t\left(q_t + \left(1 - q_t\right)\theta\right)\right)} > 0$$

This shows that no "one-shot" two period, deviation by the sender improves upon their payoff from policy σ^*. Now consider multiple period deviations in which the sender keeps the receiver's posterior beliefs after sending message *"stay"* above the threshold belief required to keep them in the game when they stick to strategy β^*. This is the only other type of admissible deviation, because the receiver quits whenever their beliefs fall below this threshold.

Consider an alternative strategy, $\tilde{\sigma}$, in which the sender deviates in periods t, $t+1$, and $t+2$, such that $\tilde{r}_t = \hat{r}_t > r_t$, $\tilde{r}_{t+1} > \hat{r}_{t+1}$ and $\tilde{r}_{t+2} < \hat{r}_{t+2} = r_{t+2}$. Let $\tilde{V}_t$ denote the sender's value in period t from the strategy $\tilde{\sigma}$, $\hat{V}_t$ their value from $\hat{\sigma}$, and V_t^* their value from σ^*. Since $\hat{r}_t = \tilde{r}_t$, it follows that the receiver's period-$t+1$ prior beliefs are the same under the two strategies: $(\hat{p}_{t+1}, \hat{q}_{t+1}) = (\tilde{p}_{t+1}, \tilde{q}_{t+1})$. Starting at period $t+1$, the strategies $\tilde{\sigma}$ and $\hat{\sigma}$ now correspond to a pair of strategies one of which restores the receiver's beliefs to the threshold level and the other inflates their beliefs for one period before restoring it in the next, respectively. But these are precisely what the strategies σ^* and $\hat{\sigma}$ do at time t. The result above shows that $V_t^* > \hat{V}_t$ for any $\hat{\sigma}$ with $\hat{r}_t > r_t$ and $\hat{r}_{t+1}$ such that $\hat{\pi}_{t+1} = \pi_{t+1}$. This implies that $\hat{V}_{t+1} > \tilde{V}_{t+1}$. Using this fact, and the expression for $\tilde{V}_t$, we show that $V_t^* > \tilde{V}_t$.

$$\tilde{V}_t = (1-\tilde{r}_t)\left[\theta b_t B + (1-\theta b_t)\,\tilde{V}_{t+1}\right] < (1-\tilde{r}_t)\left[\theta b_t B + (1-\theta b_t)\,\hat{V}_{t+1}\right] = \hat{V}_t < V_t^*$$

Now consider any strategy, σ^r, that deviates from σ^* for r periods - the sender changes the recommendation probabilities r_t for $r-1$ periods, potentially changing the receiver's posterior beliefs, then restores the receiver's posterior beliefs in the r^{th} period to the threshold level and returns to the strategy σ^* thereafter. We have shown that the sender does not benefit from such a deviation when $r = 2$ or 3.[9] We will prove that such a strategy is never beneficial for any $r \in \mathbb{N}$ by induction. We begin by assuming that it is not beneficial for $r = k$. Denote the sender's value at time s by V_s^k when the sender deviates deviates for k periods, and by V_s^{k+1} when they deviate for $k+1$ periods. Let the first period of deviation be t and assume that $V_t^* \geq V_t^k$; we will show that $V_t^* \geq V_t^{k+1}$. Notice that V_t^k can be expressed as follows:

$$V_t^k = \theta B \left[\sum_{s=0}^{k-2} q_{t+s} \prod_{i \leq s} (1-\hat{r}_{t+i}) \prod_{j \leq s-1} (1-\theta b_{t+j})\right] + \prod_{i \leq k-2} (1-\hat{r}_{t+i})\,(1-\theta b_{t+i})\, V_{t+k-1}^k,$$

[9] It is immediate that the sender does not benefit when $r = 1$, since inflating the receiver's beliefs for one period then returning to σ^* is costly for that one period (since $\hat{r}_t$ is higher than r_t), and there is no benefit in subsequent periods.

where $t+k-1$ is the last period in which the sender deviates from σ^* under σ^k - the period in which the receiver's posterior beliefs are restored to the threshold level. Let the receiver's prior beliefs at the start of this period be $(\hat{p}_{t+k-1}, \hat{q}_{t+k-1})$ and the sender's be $b_{t+k-1},$; the sender chooses $\hat{r}_{t+k-1}$ such that the receiver's posterior beliefs are at the threshold. Indeed, V^k_{t+k-1} can be expressed as follows:

$$V^k_{t+k-1} = (1 - \hat{r}_{t+k-1}) \left[B\theta b_{t+k-1} + (1 - \theta b_{t+k-1}) V^*_{t+k} \right] = V^*_{t+k-1} (b_{t+k-1}, \hat{p}_{t+k-1}, \hat{q}_{t+k-1})$$

Now consider V^{k+1}_t and suppose that the first $k-1$ deviations are exactly those employed in σ^k (such a strategy is always possible to construct). The sender's value in period t, V^{k+1}_t, can be expressed as follows:

$$V^{k+1}_t = \theta B \left[\sum_{s=0}^{k-2} b_{t+s} \prod_{i \le s} (1 - \hat{r}_{t+i}) \prod_{j \le s-1} (1 - \theta b_{t+j}) \right] + \prod_{i \le k-2} (1 - \hat{r}_{t+i}) (1 - \theta b_{t+i}) V^{k+1}_{t+k-1},$$

Notice that V^{k+1}_{t+k-1} can be further unpacked:

$$V^{k+1}_{t+k-1} = (1 - \tilde{r}_{t+k-1}) \left[B\theta q_{t+k-1} + (1 - \theta q_{t+k-1}) (1 - \tilde{r}_{t+k}) \left[B\theta q_{t+k} + (1 - \theta q_{t+k}) V^*_{t+k+1} \right] \right],$$

where $\tilde{r}_{t+k-1} \ge \hat{r}_{t+k-1}$ and $\tilde{r}_{t+k} \le \hat{r}_{t+k}$. But this is the same as the sender's value at time $t+k-1$ with priors $(b_{t+k-1}, \hat{p}_{t+k-1}, \hat{q}_{t+k-1})$ of a 2-period deviation from σ^*, which we have already shown is never beneficial. This implies the following:

$$V^*_{t+k-1} (b_{t+k-1}, \hat{p}_{t+k-1}, \hat{q}_{t+k-1}) \ge V^{k+1}_{t+k-1}$$

Combining the inequalities established above, we can conclude that σ^* is preferred by the sender to any $k+1$ period deviation:

$$V^*_t \ge V^k_t \ge V^{k+1}_t$$

This completes the proof.

$\square$