DYNAMIC PROGRESS REPORT

WEI ZHAO

ABSTRACT. This paper studies information design in a dynamic moral hazard environment. An agent and an expert face a common uncertainty regarding the effectiveness of a collective decision. The agent bears the cost of effort of information acquisition and makes the final decision. The expert is the only observer of research outcomes and provides information over time to the agent. Both parties are equally affected by the decision. I show that one optimal information policy consists in disclosing truthfully with delay. In the first periods of time, the delay is zero, then strictly increases and finally vanishes. By the time the delay decreases back to zero, the agent has taken the decision with probability one.

KEYWORDS: Dynamic Information Disclosure, Optimal hold-up, Learning, Information hoarding, Persuasion JEL CLASSIFICATION: C73; D82.

Date: October 21, 2021.

1. Introduction

When a decision has to be made under uncertainty, information acquisition has to be conducted in order to improve the effectiveness of choice. Sometimes, experts are privately informed of research outcomes. On one hand, experts are delegated to conduct information acquisition. On the other hand, only experts are able to fully process the research outcome. Therefore, experts are normally supposed to publish research progress report periodically. On one hand, such report ”puts the whole group on the same page”: it helps the rest to understand research results and make more accurate decisions. On the other hand, it also helps evaluate progress and decide whether or not to carry out further exploration. There are many scenarios where the experts have more bargaining power than the agent therefore have more freedom in writing the report. In the first category of scenarios, an expert raises funds from many investors to conduct information acquisition. For instance, during the COVID-19 pandemic, vaccine producers raise funds from multiple governments to conduct research and they are supposed to publish their experiment outcomes periodically. There are few vaccines which have entered into the final period of experiments, while all the governments are eager to get the first batch of vaccines to re-open the economy. The second category of scenarios considers feedback provision. Supervisors make comments on PhD students' projects, based on which students decide whether or not to keep exploring. Financial advisors acquire information of retail investors' risk averse degree through asking several related questions. Then they provide recommendations on financial products and portfolios. In my paper, the agent faces uncertainty in making a once-for-all choice between two options. The agent exerts effort to conduct dynamic information acquisition and the expert is privately informed of the research outcome. There are two states of the world and both players prefer to match the option with the state. The statistical experiment, which is fixed at each period, generates either a breakthrough signal for one state or its absence. The expert needs to report the progress at each instant. I characterize the expert's optimal dynamic research progress reporting policy, if she can commit to one at the beginning of game.

While both players share the same preferences in decision making, the expert does not bear the cost of information acquisition. Therefore, she prefers a more accurate decision than the agent does. This preference mis-alignment leads to departures from full information disclosure. Information acquisition creates information value (benefit) but delays decision making (cost). As information acquisition continues, both the agent becomes more pessimistic of information value created in the future and the agent's payoff of immediate decision making increases. These two components discourage the agent's further exploration. What information design can do? On the benefit side, the expert can hoard information value created in previous periods and then release it in later periods. On the cost side, the expert can persuade the agent to become more uncertain of the underlying state and therefore make the agent reluctant to make decision right away. I elaborate on these two channels in a three-period model.

In the continuous time model, I show that there is an optimal disclosure policy where the expert truthfully reports with delay. The whole research process is divided into three phases. In the first phase, the delay level is kept at zero. Then it strictly accumulates in the second phase. In the third phase, the delay level shrinks strictly. At the end of the third phase, the delay has completely vanished and the decision is taken with probability 1. Moreover the delay level in the second and third phase is designed such that at each instant, the agent is indifferent between acquiring information for another instant and making decision right away.

Under the optimal reporting policy, a knowledge gap arises since the second phase. I adopt the expected hoarded information value to measure the gap. In the first phase, the expert reveals all she has learnt immediately. In the second phase,[1] the expert generates just enough information value to balance the agent's opportunity cost for further exploration. Besides, she hoards the rest of the created information value. Therefore, the expected hoarded information value accumulates. In the third phase, the expert generates all the information value created through information acquisition at that instant. In addition, she releases the previously hoarded information value, to fill up the gap between the opportunity cost and the created information value. At the end of the third phase, the expected hoarded information value vanishes completely.

To motivate the agent to acquire information in the absence of breakthrough signal, the expert has to postpone disclosing breakthrough signal even if she prefers to reveal it immediately. That's where commitment power plays a critical role in shaping the optimal policy. However, when the decision-making payoff upon receiving the breakthrough signal is large enough, the expert is more reluctant to postpone its disclosure. In other words, the expert will put off the critical time where information value hoarding begins.

[1]The turning point here between the second and third phase is different from the one in describing the delay level in the last paragraph.

Facing limited amount of hoarded information value, the expert then turns to persuasion to improve the agent's uncertainty of the underlying state. Such persuasion will make the agent reluctant to make decision right away and therefore decreases the opportunity cost. We show that there exists at most one instant at which persuasion in conducted.

Finally, I extend the basic model to information structures with breakthrough signals in both states.

1.1. **Related literature.** This paper studies information disclosure on a learning process. If the belief space is taken as the state space, then my paper falls into the category of disclosure of an evolving state where the agent responds to inter-temporal incentive. Please check Table 1 for classifying literature on dynamic Bayesian persuasion.

	Evolving state	Persistent state
Responding to inter-temporal incentive	**My paper**, Ball [2019] Smolin [Forthcoming]	Zhao et al. [2020] Basu [2017] Bizzotto et al. [Forthcoming] Orlov et al. [2020] Au [2015] Henry and Ottaviani [2019] Escudé and Sinander [2020] Che et al. [2020]
belief-based and short-lived decision-maker	Che and Hörner [2017] Glazer et al. [Forthcoming]	None
Belief-based and long-lived decision-maker	Renault et al. [2017] Ely [2017]	None

TABLE 1. Literature category

The first closest paper is Ball [2019]. His paper studies information design of an evolving process in a repeated game. He shows that the optimal disclosure policy can be implemented by truthful report with delay. There are two critical differences between his paper and mine. First, in his paper, the underlying process is mean-reverting Ornstein-Uhlenbeck process and both the sender and the receiver share a quadratic utility functions. It is therefore without loss of **generality** to focus on the policy set of truthful report with delay. In the basic model, we show that it is rather without loss of **optimality** to focus on this policy set. However, when the payoff at breakthrough signal is relatively large enough, the expert has to turn to persuasion to decrease the opportunity cost, which can not be implemented by through truthful report with delay. Second, the delay level in his paper strictly decreases and then maintained at zero. In other words, the sender just exploits his informational advantage at the very beginning of the game and then release it gradually. Therefore, the information value hoarding phase is missing in his model.

The second closest paper is Smolin [Forthcoming]. He studies information design on a learning process in a stopping game. He assumes that the stage payoff is linear in posterior while the outside option is fixed. In my paper, however, I assume that the stage payoff is fixed while the outside option is convex in posteriors. Moreover, the sender, in his paper, prefers to persuade the agent to keep working while my model assumes that their preferences are mis-aligned only in a belief interval. Therefore the trade-off between timeliness and accurateness is novel in my paper. Moreover, the devices of both information value hoarding and persuasion, targeting at increasing benefit and lowering opportunity cost respectively, is new in my model.

In the category of literature in which information disclosure is on an evolving state and the agent is belief-based and short-lived, Che and Hörner [2017] and Glazer et al. [Forthcoming] studies dynamic information design on social learning. In their papers, the sender adopts information disclosure to internalize the externality of information acquisition by each individual agent. There are two differences between my paper and these papers. First, the long-run receiver's memory implies that the information value acts as consumables. Once some information value has been generated in previous periods through information disclosure, then this part of information value can not be released again. Second, the inter-temporal incentive is missing in their papers when it comes to the design of ”carrot” and therefore the component on the benefit side is absent. The second component is also absent in the subcategory of literature of dynamic information design on Makovian transition process with belief-based agent (c.f. Renault et al. [2017] and Ely [2017]). In these papers, a patient sender discloses a Markovian process to a long-lived agent at each period. The agent is supposed to take action at each period, which is purely based on his updated beliefs at that period.

My paper is related to the literature category where information disclosure is on a persistent state and the agent responds to inter-temporal incentives. Zhao et al. [2020] studies how a sender can make the most of a single piece of private information to motivate a long-run receiver to take a single action as frequently as possible. Orlov et al. [2020] and Bizzotto et al. [Forthcoming] study how a sender discloses information of a persistent state in responding to the exogenous flow of public information. Au [2015] and Basu [2017] consider the scenario where the receiver is privately informed. Finally, there is a category of literature which imposes both the cost and the restrictions on information structure of statistical experiment conducted by the sender. Henry and Ottaviani [2019] and Escudé and Sinander [2020] assign this cost to the sender and restrict information structures to the set of Gaussian processes. Che et al. [2020] analyze the scenario in which the cost is borne by the receiver and restrict the information structure to the set of Poisson processes.

More broadly, my paper adopts the belief-based approach which is first developed by Aumann et al. [1995] and then utilized by Kamenica and Gentzkow [2011]. Finally, my paper is related to the literature of delegated information acquisition (c.f. Angelucci [2017], Yang [2019], Clark [2016], Chade and Kovrijnykh [2016] and Zhao and Zhao [2019]). However, the key difference lies in the fact that my paper assumes that it is the expert rather than the agent who has the commitment power, while their papers assume the opposite.

2. Model

The agent faces uncertainty in making a once-for-all decision, which affects both the agent and expert equally. Denote the set of options $A := \{a_1, a_2\}$ and the set of states as $\Omega := \{\omega_1, \omega_2\}$. Initially, both players assigns probability $p_0 \in [0, 1]$. The common preference of both the agent and expert is given by Table 2.

TABLE 2. Payoff structure

Payoff \ Policy / State	a_1	a_2
ω_1	1	0
ω_2	0	1

If the choice a matches the state $\tilde{\omega}$ of the world, then their payoff is 1. Otherwise their payoff is 0. Denote by $m(p)$ the maximal utility under optimal choice given belief p. Specifically, $m(p)$ is

$$m(p) = \begin{cases} 1 - p & p \in \left[0, \frac{1}{2}\right) \\ p & p \in \left[\frac{1}{2}, 1\right] \end{cases}$$

Finally, denote the discount factor by ρ. Therefore, if the agent puts off decision-making till time t, then their payoffs are discounted by $e^{-\rho t}$.

TABLE 3. Information structure

Signal / State	s_a	s_b
ω_1	$1 - \lambda dt$	λdt
ω_2	1	0

2.1. **Information acquisition.** The agent can exert effort with cost c to acquire information before making the collective choice. If effort is exerted between t and $t + dt$, a signal is generated according to the following information structure
This signal is privately observed by the expert. If the expert forms some belief p_t at time t, then with probability $(1 - p_t)\lambda dt$ does the expert privately observe the breakthrough signal s_b and believe that the state is exactly ω_1. Otherwise, the expert receives its absence signal s_a and updates her belief p_{t+dt} with increment $dp_t = p_{t+dt} - p_t = \frac{\lambda p_t(1-p_t)dt}{1-\lambda(1-p_t)dt}$. This leads to the ordinary differential equation (ODE)

$$\dot{p} = \lambda p(1 - p) \tag{1}$$

I call *research process* $\mathcal{P}$, the process with increment specified in Equation (1) and origin p_0. At time t, the process $\mathcal{P}$ passes belief $p_t = \frac{p_0 e^{\lambda t}}{(1-p_0)+p_0 e^{\lambda t}}$.

My model also accommodates the case where $p_0 < \frac{1}{2}$. However, there exists a lower bound $\underline{p}_0$ below which the agent will make decision immediately. We offer a full characterization of $\underline{p}_0$ in Appendix. Throughout the paper, we only consider the case where $p_0 > \underline{p}_0$.

2.2. **Reporting policy.** The expert, being privately informed of the research progress, is supposed to issue a report at the end of each period. Given any measurable message space M, denote a history component at time t by $h_t = (s_t, m_t) \in S \times M$, where s_t is the generated signal at time t and m_t is the sent message at time t. $h^t = (h_{t'})_{t'<t}$ summarizes all previous histories by time t, provided that the agent has acquired information up to time t. Besides, denote by H^t the set of all previous histories h^t by time t. A dynamic disclosure policy (π, M) then specifies

$$\pi_t : H^t \times S \to \Delta(M),\ \forall t \geq 0,$$

which maps the set of previous histories by time t and the signal generated at time t to probability distribution space on message space M. Throughout the paper, I assume that the expert has all the bargaining power and can commit to any dynamic reporting policies at the beginning of the game. Denote C as the the action of pending decision-making. I will show it later that it is without loss of generality to restrict M to $A \cup \{C\}$.

Given a dynamic reporting policy (π, M), denote by $h_m^t = (m_{t'})_{t'<t}$ previous history of messages by time t and H_m^t the set of all previous histories of message by time t. A decision rule α then specifies

$$\begin{cases} \alpha_0 : \emptyset \to \Delta\left(A \cup \{C\}\right) \\ \alpha_t : H_m^t \times M \to \Delta(A \cup \{C\}) \quad \forall t > 0 \end{cases}.$$

At time t, the agent, upon receiving message history h_m^t by time t and the message m_t at time t, decides whether or not keep acquiring information. If not, he needs to choose between option a_1 and a_2.

Given policy (π, M) and decision rule α, denote by $\pi_t^\alpha(h^t, s_t, m_t|\omega)$ the joint probability that the previous history by time t is h^t, the signal geneated at time t is s_t, the sent message at time t is m_t and the agent keeps acquiring information by time t, conditional on state ω. Finally, denote the aggregate discounted frequency of option $a \in A$ being selected conditional on state ω by $Y^\alpha(a|\omega)$, which is given as

$$Y^\alpha(a|\omega) = \int_{t=0}^{\infty} \int_{h^t \in H^t} \int_{m_t \in M} e^{-\rho t} \left[\pi_t^\alpha(h^t, s_b, m_t|\omega) + \pi_t^\alpha(h^t, s_a, m_t|\omega)\right] \alpha_t(a|h_m^t, m_t) dm_t dh^t dt.$$

The discounted information acquisition cost is given by

$$\begin{aligned} &c \sum_{\omega} p_0(\omega) \int_{t=0}^{\infty} \int_{h^t \in H^t} \int_{m^t \in M^t} e^{-\rho t} \left[\pi_t^\alpha(h^t, s_b, m_t|\omega) + \pi_t^\alpha(h^t, s_a, m_t|\omega)\right] [1 - \sum_{a \in A} \alpha_t(a|h_m^t, m_t)] dm^t dh^t dt \\ =&c \left[\frac{1}{\rho} - \sum_{\omega \in \Omega} \sum_{a \in A} p_0(\omega) Y^\alpha(a|\omega)\right] \end{aligned}$$

The expert then solves the following optimization problem

$$\max_{(\pi, M), \alpha} p_0 Y^\alpha(a_2|\omega_2) + (1 - p_0) Y^\alpha(a_1|\omega_1)$$

$$s.t.\ \alpha \in \textbf{argmax}_{\alpha'} p_0 Y^{\alpha'}(a_2|\omega_2) + (1 - p_0) Y^{\alpha'}(a_1|\omega_1) - c \left[\frac{1}{\rho} - \sum_{\omega \in \Omega} \sum_{a \in A} p_0(\omega) Y^{\alpha'}(a|\omega)\right]$$

Please note that whenever there are multiple optimal decision rule α given reporting policy (π, M), the agent is assumed to pick expert's most preferred ones.

2.3. **Application.** My model imposes a strong assumption on the expert's bargaining and committing power in designing dynamic reporting policies. It can be applied to two scenarios. In the first scenario, a single expert designs her dynamic research progress reporting policy when she raises funds from multiple investors to conduct information acquisition. The second scenario studies dynamic feedback design.

2.3.1. *Public fundraising on research.* During the COVID-19 pandemic, vaccine producers raise funds from multiple governments to conduct research on valid vaccines. However, few vaccine producers have survived and entered into the third phase of experiment. Many promising producers decided to stop further experimentation due to serious side-effects found in clinical trial. However, almost all governments want to get the first batch of vaccines to re-open the economy and ensure the safety of citizens. At each period, the vaccine producers have to issue a periodical progress report on antibody production effectiveness, side effects, clinical experiment results, etc. Based on the report, these governments then decide whether or not fund vaccine research for another period. If not, then these governments have to decide whether or not purchase the vaccine. The most advanced vaccine producer can have large bargaining power in designing the reporting policy.

Many startup firms raise funds in primary market to incubate and develop their business before IPO (initial pricing offering) issuance and entering into the secondary market. Fundraising in primary market always involves multiple rounds of financing. It starts with the angel round and continues with round A, B, etc. At the beginning of each round, the entrepreneurs have to conduct roadshow to report their projects' progress to potential investors. These investors then decide whether to fund the project for another round. If not, the investors decide whether to sell their possessed shares or not after the entrepreneurs enter into the secondary market. Many unicorns, who have dominated the market shares and/or proved itself of profit making, are popular in primary market. They have more bargaining power and flexibility in designing their roadshow.

Similarly, listed and high-technology firms in NASDAQ raises funds from both retail and institutional investors to conduct research. They are required to issue periodically mandatory report or they are also expected to issue voluntary report, which discloses their research progress. Such report affects either their share prices or the applicability of new share issuance, which then determines whether they have collected enough funds to conduct further exploration. These firms are always assumed in the literature to have more bargaining and commitment power in designing both the contract and disclosure policy.

Finally, crowdfunding is the practice of funding a project or venture by raising small amounts of money from a large number of people, typically via the Internet. Some projects may involve multiple rounds of financing. For instance, Linux is a computer system which promotes open source, security and free, in contrast to Mac OS and Windows. There are multiple ways to finance the development of their systems. Elementary OS, one of the most popular Linux distribution, offers free developed system but charges fees if customers wants to experience the beta system in advance. The developers charge first batch of 25 USD and then 10 USD each month for subscription. In reciprocal, the developers have to issue updates and share with all the backers.

Throughout the paper, I assume that the expert can commit to dynamic reporting policy in the beginning of the game. This assumption is justified by either law enforcement and monitoring or by reputation motivation. Reputation matters for experts and any inconsistency or lie may damage their career. Entrepreneurs also care about reputation and credit, otherwise they may find it harder or even impossible to get finance or sell their products in the future.

2.3.2. *Dynamic feedback design.* Retail investors may come to the bank to purchase some financial products. They have to turn to financial advisors for recommendations of financial products. However, both investors and financial advisors have no idea of investors' risk averse degree. The financial advisors then ask the investors several questions and the investors have to consume both time and effort to answer these questions. Investors' answers can reveal information about their risk averse degree but can only be evaluated by the financial advisors. Finally, the financial advisors then makes recommendations on financial products and portfolios to the retail investors. Currently, more and more retail investors conduct such purchase online and AI has already replaced these financial advisors.

First-year PhD students have to choose a field to conduct research, between micro-economics and macro-economics or between theory or empiric, etc. However, they may not know at which field they are more talented. Therefore, they exerts effort and time to attend courses in both fields, and course directors offers exercises and exams for feedbacks. Course directors can control informativeness of these feedbacks through two ways. First, they can design the difficulty and content coverage of these tests. On the other hand, they can also choose how to disclose students' performances in exams. The informativeness is restricted by the depth and width of students' knowledge in a given field, which is positively related to the amount of time and effort students have exerted in studying the subject.

Role-playing games (RPG) are also designed to accommodate different types of players. Therefore, game developers introduce several professions and positions. Players are supposed to choose the most appropriate one to improve their gaming experience. However, players may not know at which position they are more talented. In order to make a better choice, players have to exert effort to learn features of different professions and positions. Game developers then design feedback mechanism to help players learn more about themselves and therefore make more accurate choices. For instance, game developers control level of difficulty of quests and challenges during the experience phase. Many online MOBA gaming (for instance, DOTA and LOL, which offer a platform where a group of players combat with the other.) can design the matching mechanism through controlling the level of both teammates and opponents. Based on the performance, the players acquire more information of their own types, which then enables them to make more accurate decisions.

2.4. **Preference mis-alignment.** Even if the collective choice affects the expert and agent equally, the expert prefers a more accurate choice making since she does not bear the cost of information acquisition. Given belief p_t at time t and under perfect information disclosure, compare the payoff between acquiring information for dt period and making decision immediately. The payoff difference is

$$\begin{aligned} & -cdt + e^{-\rho dt}\left[\frac{p_t}{p_{t+dt}}m(p_{t+dt}) + \left(1-\frac{p_t}{p_{t+dt}}\right)m(0)\right] - m(p_t) \\ =& \underbrace{e^{-\rho dt}\left[\frac{p_t}{p_{t+dt}}m(p_{t+dt}) + \left(1-\frac{p_t}{p_{t+dt}}\right)m(0) - m(p_t)\right]}_{\text{Benefit}} - \underbrace{\left[cdt + \underbrace{\rho m(p_t)dt}_{\text{time cost}}\right]}_{\text{Opportunity cost}} \end{aligned}$$

On the benefit side, information acquisition enables more accurate choice making, therefore creating information value. On the cost side, it also puts off decision making, making it absent the flow payoff $\rho m(p_t)dt$ of decision making at time t. I denote this part of opportunity cost as time cost, which is strictly increasing in payoff of immediate decision-making. In addition, the opportunity cost for the agent also includes the information acquisition cost. Denote by $q^R = \frac{\lambda}{\lambda+\rho}$ the belief where the created information value (benefit) balances the time cost (opportunity cost for the expert). Denote by $q^P = \frac{\lambda-c}{\lambda+\rho}$ the belief where the created information value (benefit) balances the total of time cost and information acquisition cost (opportunity cost for the agent). Our next lemma solves the optimal stopping problem for both the expert and the agent respectively.

Lemma 1. *Under perfect revelation of signals, the agent stops information acquisition at q^P, upon not receiving breakthrough signal s_b. The expert prefers the agent to stop information acquisition at q^R upon not receiving breakthrough signal s_b.*

The fact that $q^P < q^R$ implies that the expert prefers a more accurate choice making than the agent. As the agent keeps acquiring information, it becomes harder to make research progress and the created information value therefore decreases. At the same time, the uncertainty about the state keeps resolving and the time cost therefore increases. Such trend will remain if the agent keeps acquiring information.

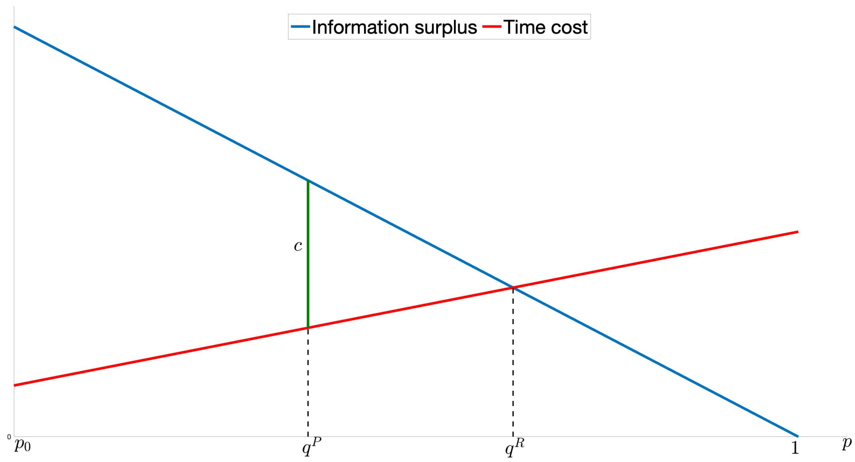

FIGURE 1. Information value and time cost

Their preferences are mis-aligned in interval (q^P, q^R), where the agent prefers to make decision right away while the expert prefers to acquire more information. The expert prefers to prolong information acquisition at that interval. What information design can do? On the benefit side, in addition to the information value necessary to balance the opportunity cost for information acquisition, there is remaining part in early periods. The expert can hoard this part of information value and release it in later periods. On the cost side, the expert can persuade the agent that his payoff of immediate decision-making is low enough. Such information disclosure can therefore decreases his time cost and therefore the opportunity cost for information acquisition. I will elaborate on these two components in a three-period toy model in the next section.

3. EXAMPLE

In this section, I consider a three period, discrete version of the model. We keep the payoff structure and signal space $\mathcal{S}$ the same. However, information structure of statistical experiment is revised as follows,

TABLE 4. Discrete version information structure

State \ Signal	s_a	s_b
ω_1	$1-e^{-\lambda}$	$e^{-\lambda}$
ω_2	1	0

Under the information structure, the prior belief p_t at time t jumps either to belief 0 or to belief p_{t+1} given as

$$p_{t+1} = f(p_t) := \frac{p_t}{(1-p_t)(1-e^{-\lambda})+p_t}.$$

To make things interesting, let me introduce the following assumptions on p_0.

Assumption 1. *The agent, being perfectly informed of research progress, prefers to stop exerting effort and make choice at the end of the first period. However, the expert prefers the agent to put off collective choice making till the end of the second period.*[2]

The information value *created* through information acquisition in the second period is not enough to cover the opportunity cost. To motivate the agent to acquire information in the second period, a component on the benefit side is proposed, i.e. information hoarding. Intuitively, the expert can *hoard* additional information value created in the first period and then *release* it in the second. Specifically, let me define q^* as follows,

$$\begin{aligned}
\underbrace{\frac{1}{1+\delta}m(q^*)}_{\text{time cost}} + \underbrace{\frac{1}{1+\delta}c}_{\text{information acquisition cost}} &= \underbrace{\frac{\delta}{1+\delta}\left[\frac{q^*}{f^{(2)}(p_0)}m(f^{(2)}(p_0)) + \left(1-\frac{q^*}{f^{(2)}(p_0)}\right)m(0) - m(q^*)\right]}_{\text{discounted aggregate information value}} \\
&= \underbrace{\frac{\delta}{1+\delta}\left[\frac{q^*}{f(p_0)}m(f(p_0)) + \left(1-\frac{q^*}{f(p_0)}\right)m(0) - m(q^*)\right]}_{\text{discounted hoarded information value}} \\
&\quad + \underbrace{\frac{\delta}{1+\delta}\frac{q^*}{f(p_0)}\left[\frac{f(p_0)}{f^{(2)}(p_0)}m(f^{(2)}(p_0)) + \left(1-\frac{f(p_0)}{f^{(2)}(p_0)}\right)m(0) - m(f(p_0))\right]}_{\text{discounted newly generated information value in third period}} \qquad (2)
\end{aligned}$$

Consider a belief martingale which starts from belief q^* and jumps to either belief 0 or $f^{(2)}(p_0)$. The first equality in Equation (2) means that the *generated* information value is just enough to balance the opportunity cost. The second equality in Equation (2) means

[2]Mathematically, the following inequalities hold

$$\begin{cases}
m(p_0) < -\frac{1}{1+\delta+\delta^2}c + \frac{\delta+\delta^2}{1+\delta+\delta^2}\left[\frac{p_0}{f(p_0)}m(f(p_0)) + \left(1-\frac{p_0}{f(p_0)}\right)m(0)\right] \\
m(f(p_0)) > -\frac{1}{1+\delta}c + \frac{\delta}{1+\delta}\left[\frac{f(p_0)}{f^{(2)}(p_0)}m(f^{(2)}(p_0)) + \left(1-\frac{f(p_0)}{f^{(2)}(p_0)}\right)m(0)\right] \\
m(f(p_0)) < \frac{\delta}{1+\delta}\left[\frac{f(p_0)}{f^{(2)}(p_0)}m(f^{(2)}(p_0)) + \left(1-\frac{f(p_0)}{f^{(2)}(p_0)}\right)m(0)\right]
\end{cases}$$

that the generated information value can be divided into two parts. The first part is the information value *hoarded* in the last period while the second part is the information value *created* through information acquisition in the second period (which is realized in the third period). Then I define the reporting policy τ^I through specifying its belief martingale. The induced belief martingale jumps to either belief 0 or q^* at the end of the first period.[3] If the belief q^* is reached, then the belief martingale jumps to belief 0 or $f^{(2)}(p_0)$ at the end of the second period. I illustrate the information value management for policy τ^I through Figure 2.

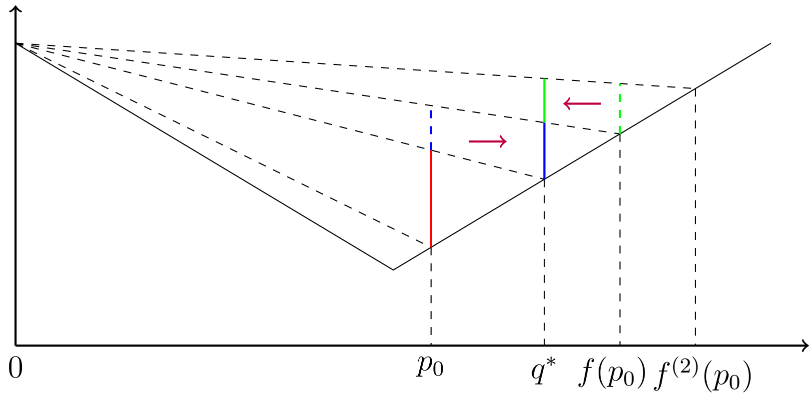

FIGURE 2. **Component on the benefit side, information hoarding**. The combination of red line and blue dashed line denotes the information value *created* in the first period while the green dashed line is the one *created* in the second period. At the end of the first period, the expert hoards the blue dashed part. Then she releases it at the end of the second period, generating the blue line. τ^I *generates* the red line at the end of the first period and the combination of blue and green line at the end of the second period.

However, the additional information value in the first period may not be enough to fill up the gap between opportunity cost and the information value *created* in the second period. To determine the maximal hoardable information value *created* in the first period, let me define the belief $\underline{p}$ such that

$$m(p_0) = -\frac{c}{1+\delta+\delta^2} + \frac{\delta+\delta^2}{1+\delta+\delta^2}\left[\frac{p_0}{\underline{p}}m(\underline{p}) + \left(1-\frac{p_0}{\underline{p}}\right)m(0)\right].$$

If $q^* < \underline{p}$, then the agent will not acquire information in the first period, under policy τ^I. To fix it, a component on the cost side is proposed, i.e. persuading the agent that the opportunity cost of information acquisition is low. Let me revise the belief martingale at the end of the first period induced by reporting policy τ^I. It is replaced by a compound belief martingale, which jumps to either belief 0 or $\underline{p}$. If $\underline{p}$ is reached, then it further

[3]If belief 0 is reached, then the belief martingale stays there.

jumps to either $f(p_0)$ or q^*. The splitting between belief $f(p_0)$ and q^* does not generate any information value. However, this splitting decreases the payoff of immediate choice making from $m(\underline{p})$ to $m(q^*)$ with probability $\frac{f(p_0)-\underline{p}}{f(p_0)-q^*}$, which therefore decreases the time cost and therefore the opportunity cost for information acquisition. Let me denote the corresponding reporting policy by τ^U, which is illustrated in the Figure 3.

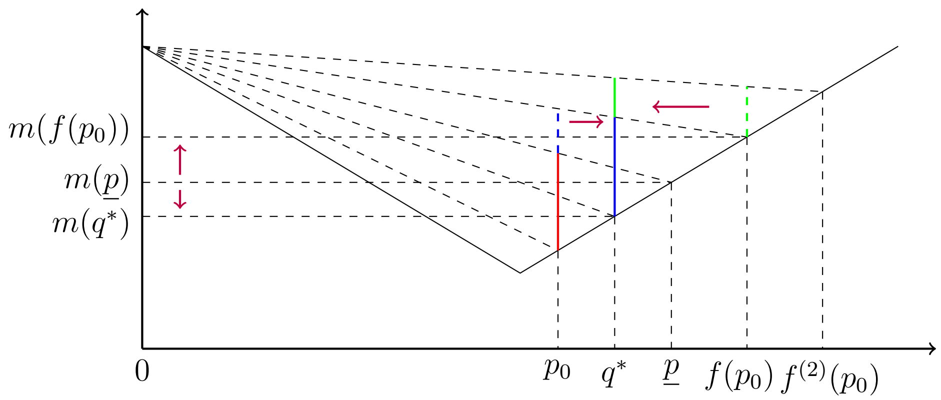

FIGURE 3. **Component on the cost side, persuasion**. The expert has to generate the information value of the red line to motivate the agent to acquire information in the first period. She hoards the information value of blue dashed line. Next, she persuades the agent to lower his payoff of immediate choice-making from $m(\underline{p})$ to $m(q^*)$ with probability $\frac{f(p_0)-\underline{p}}{f(p_0)-q^*}$. At the end of the second period, the expert generates information value in combination of blue and green line.

Proposition 1. *Whenever full information disclosure policy is not optimal,*[4] *if $q^* \geq \underline{p}$, then τ^I is optimal, if $q^* < \underline{p}$, then τ^U is optimal.*

In order to bridge the gap between created information value in the second period and opportunity cost, the expert has to hoard information value created in the first period. Only when this amount of hoarded information value is not enough does the expert persuade the agent that his payoff of immediate decision-making is low.[5]

There are two elements present in infinite period but missing in three-period model. First, to induce the agent to acquire more information, the expert can start to hoard

[4]This is true when $q^* < \underline{p}'$, where $\underline{p}'$ is such that,

$$\frac{\delta^2}{1+\delta+\delta^2}\frac{p_0}{f(p_0)}\left[\frac{f(p_0)}{f^{(2)}(p_0)}m(f^{(2)}(p_0))+\left(1-\frac{f(p_0)}{f^{(2)}(p_0)}\right)m(0)-m(f(p_0))\right]$$
$$-\frac{\delta}{1+\delta+\delta^2}\frac{p_0}{f(p_0)}m(f(p_0))=\frac{\delta}{1+\delta+\delta^2}\frac{p_0}{\underline{p}'}(1-\frac{\underline{p}'}{f(p_0)})m(0)$$

[5]The combination of information value as a carrot and persuasion in dynamic information design has been found in Ely and Szydlowski [2020] and Zhao et al. [2020]. In the former paper, the initial disclosure persuades the agent to be sufficiently optimistic of task easiness. In the later paper, the initial disclosure persuades the agent that the opportunity cost of exerting effort is low enough.

information in earlier periods, which therefore increases the pool of hoarded information value. Or, the expert, when facing limited amount of hoarded information value, can persuade the agent to lower his payoff of immediate decision-making and therefore the opportunity cost of information acquisition. Which tool should the expert turn to? Second, how does the expert hoard and release information value? Should the expert hoard information consecutively or intermittently? How much information value to be hoarded or released at each period? In the next section, we fully characterize the optimal disclosure policy in continuous time model.

4. Main result

This section is divided into four subsections. In the first subsection, I completely pin down the belief martingale induced by an optimal reporting policy. The second subsection sheds light on the information value management implied by the optimal reporting policy. The third subsection shows that one can implement the optimal reporting policy through truthful report with delay. In the forth subsection, I give a sketch of the proof.

4.1. **Construction of optimal reporting policy.** In this subsection, I construct one optimal reporting policy through the induced belief martingale. As first step, a belief process named "one-step-ahead" process, similar to the definition of q^*, will be defined. As second step, a reporting process is defined as the lower envelope of research process and one-step-ahead process. Finally, it is shown that one optimal reporting policy induces a belief martingale which either jumps to 0 or follows one reporting process.

4.1.1. *One-step-ahead process.* Recall that q^* is defined by the following property. Consider a belief martingale which starts from q^* and jumps to either 0 or $f^{(2)}(p_0)$, the generated information value is just enough to balance the opportunity cost for the agent. Similarly, let me define *one-step-ahead (belief) process* Q. Consider a belief martingale which starts from $q_t \geq \frac{1}{2}$ and jumps to either 0 or moves up by some increment dq_t, the *one-step-ahead process* then specifies the increment dq_twhich makes the agent indifferent between making decision immediately and acquiring information for another dt period. Mathematically, dq_t solves the following equation,

$$m(q_t) = -cdt + e^{-\rho dt}\left[\frac{q_t}{q_t + dq_t} m(q_t + dq_t) + \frac{dq_t}{q_t + dq_t} m(0)\right].$$

This leads to the ODE

$$\dot{q} = q(\rho q + c) \tag{3}$$

The increment is strictly increasing in both discount factor ρ and c. This is due to the fact that the increment is set such that the generated information value balances the total of time cost, which is increasing in ρ, and research cost c.

I call *one-step-ahead process* $\mathcal{Q}$, the process whose increment is specified in Equation (3) if it is below 1 and is maintained as 0 once it passes 1. More specifically, one-step-ahead process with origin $q_0 \in (0,1)$ is

$$q_t = \begin{cases} \frac{cq_0 e^{ct}}{\rho q_0(1-e^{ct})+c} & t \in \left[0, \frac{1}{c}\ln\left(\frac{\rho q_0 + c}{(\rho+c)q_0}\right)\right) \\ 1 & t \in \left[\frac{1}{c}\ln\left(\frac{\rho q_0+c}{(\rho+c)q_0}\right), \infty\right) \end{cases}.^{6} \tag{4}$$

Moreover, denote by $\mathcal{Q}^{\hat{t}}$ the one-step-ahead process which passes through belief $p_{\hat{t}}$ at time $\hat{t}$, or with origin $\frac{cp_{\hat{t}}e^{-c\hat{t}}}{\rho p_{\hat{t}}\left(1-e^{-c\hat{t}}\right)+c}$.[7]

4.1.2. *Hitting time.* For any process $\mathcal{Y}$ and any real number y, define *passing time* $T(y, \mathcal{Y})$ as the first time at which $\mathcal{Y}$ passes through y. Mathematically, it is defined as

$$T(y, \mathcal{Y}) := \inf\left\{t : Y_t = y\right\}.$$

For any $\hat{t} \leq T(q^P, \mathcal{P})$, define *hitting time* $\theta(\hat{t})$ as the last time at which $\mathcal{Q}^{\hat{t}}$ hits $\mathcal{P}$. Mathematically, $\theta(\hat{t})$ is defined as

$$\theta(\hat{t}) := \sup\left\{ t \geq \hat{t} : \underbrace{\frac{p_{\hat{t}}e^{\lambda(t-\hat{t})}}{(1-p_{\hat{t}}) + p_{\hat{t}}e^{\lambda(t-\hat{t})}}}_{\mathcal{P}\text{ at time } t} = \underbrace{\frac{cp_{\hat{t}}e^{c(t-\hat{t})}}{\rho p_{\hat{t}}(1-e^{c(t-\hat{t})}) + c}}_{\mathcal{Q}^{\hat{t}}\text{ at time } t} \right\}.$$

The definition of $\theta(\hat{t})$ is interpreted in Figure 4.

In Figure 4, we see that $\theta(\hat{t})$ decreases as $\hat{t}$ increases. I summarize this finding in the following Lemma.

Lemma 2. *For any $\hat{t} \leq T(q^P; \mathcal{P})$, the following holds*

(i). For any $t \in (\hat{t}, \theta(\hat{t}))$, $\mathcal{Q}^{\hat{t}}$ is strictly below $\mathcal{P}$, or $q_t < p_t,\ \forall t \in (\hat{t}, \theta(\hat{t}))$.

(ii). $\theta(\hat{t})$ is strictly decreasing in $\hat{t}$ and $\theta(T(q^P, \mathcal{P})) = T(q^P, \mathcal{P})$.

[6]The derivation of q_t can be found in Appendix C.4.

[7]One can easily verify that

$$q_{\hat{t}}^{\hat{t}} = \frac{c\frac{cp_{\hat{t}}e^{-c\hat{t}}}{\rho p_{\hat{t}}\left(1-e^{-c\hat{t}}\right)+c}e^{ct}}{\rho\frac{cp_{\hat{t}}e^{-c\hat{t}}}{\rho p_{\hat{t}}\left(1-e^{-c\hat{t}}\right)+c}(1-e^{ct})+c} = p_{\hat{t}}$$

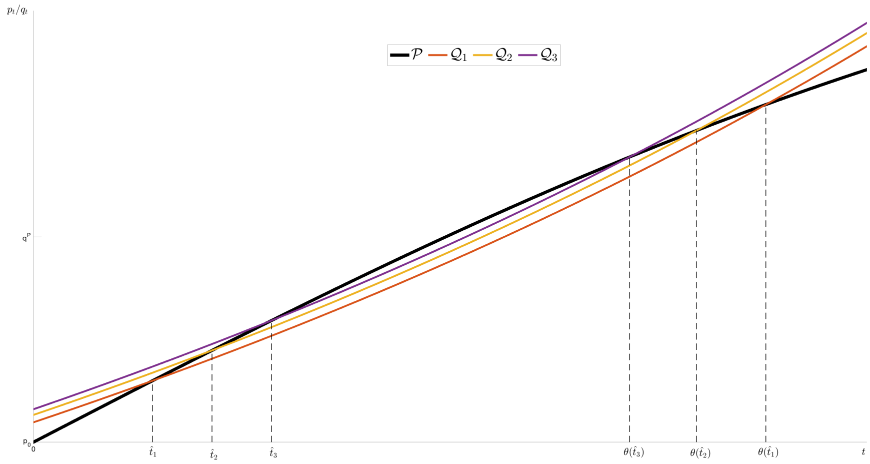

FIGURE 4. Hitting time $\theta(\hat{t})$. For any $\hat{t}_1 < \hat{t}_2 < \hat{t}_3$, we have that $\theta(\hat{t}_3) < \theta(\hat{t}_2) < \theta(\hat{t}_1)$. Moreover, $\mathcal{Q}^{\hat{t}}$ is strictly above than $\mathcal{P}$ below $\hat{t}$ or above $\theta(\hat{t})$. Otherwise, $\mathcal{Q}^{\hat{t}}$ is strictly below $\mathcal{P}$.

4.1.3. *Reporting process.* For any $\hat{t} \leq T(q^P, \mathcal{P})$, define a belief process $\mathcal{Q}(\hat{t})$ as the lower envelope of $\mathcal{P}$ and $\mathcal{Q}^{\hat{t}}$, or mathematically,

$$q_t(\hat{t}) = \min\left\{q_t^{\hat{t}}, p_t\right\}, \ \forall t \geq 0.$$

By Lemma 2, $\mathcal{Q}(\hat{t})$ is equal to $\mathcal{P}$ below $\hat{t}$ and above $\theta(\hat{t})$. Besides, $\mathcal{Q}(\hat{t})$ is equal to $\mathcal{Q}^{\hat{t}}$ in interval $[\hat{t}, \theta(\hat{t})]$. We call $\hat{t}$ as *transition time* and $\mathcal{Q}(\hat{t})$ as *reporting process with transition time* $\hat{t}$. $\mathcal{Q}(\hat{t})$ is shown in Figure 5.

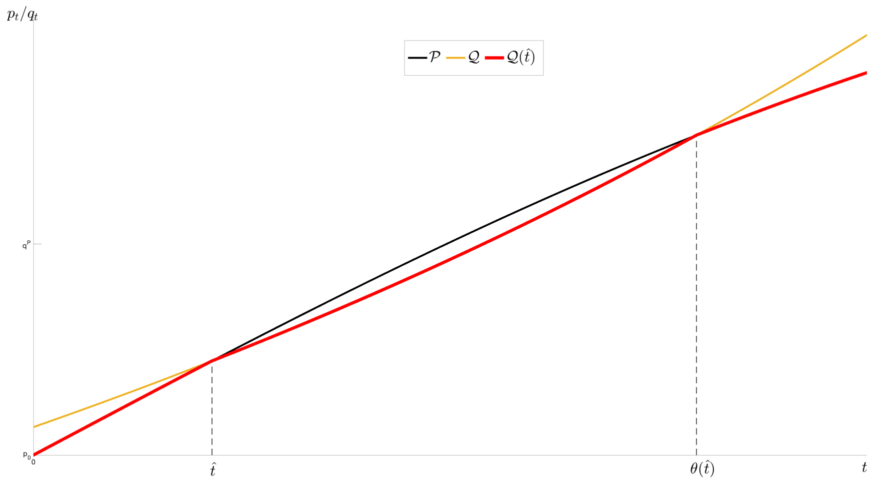

FIGURE 5. Reporting process $\mathcal{Q}(\hat{t})$ with transition time $\hat{t}$. $Q(\hat{t})$ (red curve), is a lower envelope of $\mathcal{P}$ (black curve) and $\mathcal{Q}^{\hat{t}}$ (orange curve).

Define a belief martingale $\gamma^{\hat{t}}$ which either jumps to 0 or follows $\mathcal{Q}(\hat{t})$. The following theorem shows that such a martingale is induced by some optimal reporting policy.

Theorem 1. *There exists a unique transition time t^* such that one optimal reporting policy induces a belief martingale γ^{t^*}. Moreover the agent will stop information acquisition and choose a_2 at $\theta(t^*)$, provided that his belief has not jumped to 0.*

4.1.4. *Determination of t^*.* For any prior belief p_0 and any time $s \geq 0$, define function $g(s;p_0)$ as the difference in payoff between acquiring information (and perfectly observing the learning outcome) till time s and making decision right away (and therefore never acquiring information). Mathematically, $g(s;p_0)$ is defined as

$$g(s;p_0) = \underbrace{\rho\int_{t=0}^{\infty} e^{-\rho t}\min\left\{\frac{p_0}{p_t},\frac{p_0}{p_s}\right\}m(0)dt + e^{-\rho s}m(p_s) - c\int_{t=0}^{s} e^{-\rho t}\frac{p_0}{p_t}dt}_{\text{Payoff of funding research till time } s} - m(p_0) \tag{5}$$

Then, define $\underline{t}(p_0)$ as the supremum of time s such that the payoff of acquiring information till time s is strictly lower than that of making decision immediately. Mathematically, $\underline{t}(p_0)$ is defined as

$$\underline{t}(p_0) := \sup\left\{s : \forall s' \in (0,s), g(s';p_0) < 0\right\}.$$

In the following corollary, we completely characterize t^* and show its uniqueness.

Corollary 1. *t^* is uniquely determined and can be characterized as*

$$t^* = \inf\left\{\hat{t} \geq \underline{t}(p_0) : \frac{\rho}{\rho+c} - ce^{-(\rho+c)(\theta(\hat{t})-\hat{t})}\left[\left(\frac{p_{\hat{t}}}{p_{\theta(\hat{t})}}\right)^2\frac{e^{c(\theta(\hat{t})-\hat{t})}}{(\rho+\lambda)(p_{\theta(\hat{t})}-q^P)} - \frac{1}{\rho+c}\right] \leq 0\right\}.$$

To motivate the agent to acquiring more information and make more accurate decision, the expert has to enlarge the pool of hoarded information and therefore bring forward the time at which information value hoarding begins. However, this act puts off decision making and therefore discounts the realization of information value. Determining the optimal transition time t^* and therefore the optimal policy balances the trade-off between timeliness and accurateness.

4.2. **Information value management.** Under the optimal reporting policy, the expert partially discloses what she has privately observed. Therefore a gap between the expert's and the agent's knowledge emerges.

Definition 1. *A knowledge gap is (p,q) with $q \leq p$ such that the agent's belief is q while the expert's belief is p with probability q/p and 0 with probability $1-q/p$.*

There is no knowledge gap before transition time t^*. Then the knowledge gap $(p_t, q_t(t^*))$ emerges with probability $\frac{p_0}{q_t(t^*)}$ for any $t \in (t^*, \theta(t^*))$. At the hitting time $\theta(t^*)$, the knowledge gap vanishes and the agent completely stops acquiring information. To measure the knowledge gap (p, q), define the *hoarded information value* $S(p, q)$ as

$$S(p,q) = \frac{q}{p} m(p) + \left(1 - \frac{q}{p}\right) m(0) - m(q) \tag{6}$$

At time t, the expected hoarded information value $s(t; t^*)$ is given

$$s(t;t^*) = \left(1 - \frac{p_0}{q_t(t^*)}\right) S(0,0) + \frac{p_0}{q_t(t^*)} S(p_t, q_t(t^*)) = \frac{p_0}{q_t(t^*)} S(p_t, q_t(t^*))$$

The expected hoarded information value $s(t; t^*)$ is plotted in Figure 6.

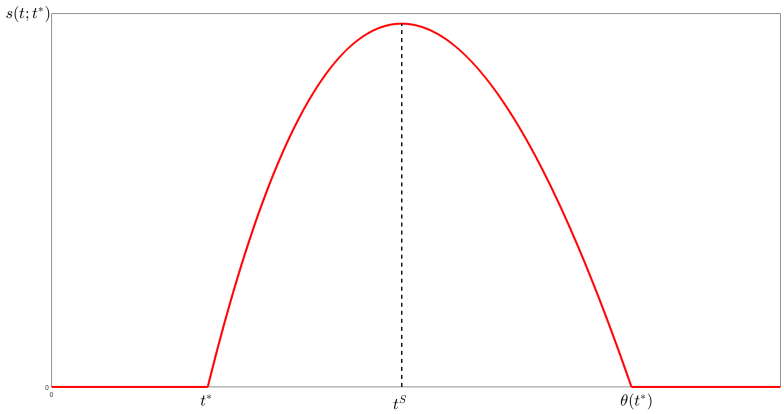

FIGURE 6. Aggregate hoarded information value

As shown on Figure 6, there exists a turning point t^S in the time interval $(t^*, \theta(t^*))$, below which the expected hoarded information value $s(t;t^*)$ strictly increases, and above which $s(t;t^*)$ strictly shrinks.

Corollary 2. *The expected hoarded information value $s(t;t^*)$ satisfies the following*

(i). $s(t;t^*) \equiv 0$ *for* $t \in [0, t^*] \cup [\theta(t^*), \infty)$*;*

(ii). Let $t^S = t^* + \frac{1}{\lambda - c} \ln \frac{\lambda(1-p_{t^*})}{\rho p_{t^*} + c}$ *be the turning point, we have that* $\frac{\partial s(t;t^*)}{\partial t} > 0$ *for all* $t \in (t^*, t^S)$ *and* $\frac{\partial s(t;t^*)}{\partial t} < 0$ *for all* $t \in (t^S, \theta(t^*))$.

From Corollary 2, the optimal reporting policy divides the whole time into three phases based on the management of information value. The first phase occurs before transition time t^*, the knowledge gap does not emerge and the expert truthfully reveal all what she has privately observed in time. The second phase occurs in time interval $(t^*, t^S]$. In this phase, the expert releases information value just enough to balance the opportunity cost, hoarding the additional information value. Therefore, the knowledge gap widens and the expected hoarded information value strictly accumulates. Finally, in the third phase, the information value generated at each instant is comprised of two parts. The first part is equivalent to the amount of information value created through information acquisition at that instant. The other part is the previously hoarded information value, which is released to bridge the gap between the first part and the opportunity cost. During the third phase, the knowledge gap narrows and the expected hoarded information value also shrinks. The third phase ends at the hitting time $\theta(t^*)$, when both the knowledge gap and the expected hoarded information value vanish completely.

4.2.1. *Comments on information value management.* First, information hoarding is concentrated into a time interval rather than scattered over the whole process. Besides, such time interval occurs just before the third phase, where the hoarded information value is gradually released. This property is mainly due to discounting. Intuitively, given the same amount of (undiscounted) information value being hoarded and the same releasing time, the later it is hoarded, the less it will be discounted away. In order to preserve the value of hoarded information value, the expert should put off hoarding information value as much as possible and the best way is to put it together in a continuous period just before releasing the information value.

Second, at each instant in both the second and third phase, the expert motivates the agent to conduct costly information acquisition by promising to generate information value in the future just enough to cover the opportunity cost for the agent. Then, at the next instant, the expert will at first release information value immediately to fulfill her promise at the last instant. Such procedure is iterated until the pool of hoarded information value is drained. To hoard information value, the experts has to garble between belief 0 and p_t and recommends for acquiring information. This implies that unnecessary information acquisition has to be conducted and decision-making has to be put off even if break-through signals have been received. These two distortions have to be imposed to motivate the agent to further explore upon receiving s_a.[8] Fix the time at which the agent stops information acquisition provided that his belief has not jumped to 0. Subject to incentive compatibility constraint, such information value arrangement leads to the fastest revelation of break-through signal s_b and therefore alleviates these two distortions to the most extent (which is shown in Lemma 4).

[8]Please note that the first distortion is not imposed on the expert directly since she does not bear the research cost. However, to compensate the agent for unnecessary information acquisition, the expert has to hoard more information value in the second phase. This therefore aggravates the second distortion.

Finally, device on the cost side, i.e. persuasion, is absent. Such device is like a double-edge sword. Given fixed amount of hoarded information value, persuasion, which splits beliefs, leads to earlier and therefore less accurate decision with some probability, even if it prolongs the length of acquiring information with the rest probability. In addition, to motivate the agent's further exploration, the expert can also bring forward the information hoarding phase and therefore hoard more information value. Information value hoarding is less costly than persuasion, which is shown in Lemma 5.

4.3. **Implementation of the optimal policy.** In this subsection, I show that the optimal reporting policy proposed by Theorem 1 can be implemented through truthfully reporting with strategic delay. Specifically, for any $t \in [0, \infty)$, define a delay function $D(t;t^*)$ such that $\mathcal{Q}(t^*)$ at time t is consistent with research progress process $\mathcal{P}$ at time $t - D(t;t^*)$. Mathematically, $D(t;t^*)$ is defined as the solution to the following equation,

$$p_{t-D(t;t^*)} = q_t(t^*) \tag{7}$$

First, since both $\mathcal{Q}(t^*)$ and $\mathcal{P}$ are strictly increasing in t, the delay function $D(t;t^*)$ is well defined. Figure 7 illustrates the definition of the delay function

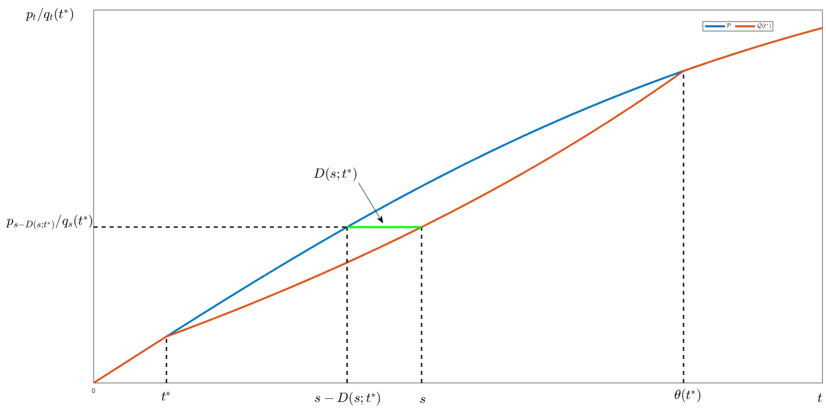

FIGURE 7. Definition of delay function

In the following corollary, I show that, in the open interval $(t^*, \theta(t^*))$, there exists a turning point t^D, below which the delay strictly increases and above which the delay strictly decreases. It also shows that the optimal reporting policy can be implemented by truthfully reporting the progress with delay $D(t;t^*)$.

Corollary 3. *Define the turning point* $t^D = T(q^P, \mathcal{Q}(t^*))$, *the optimal reporting policy satisfies the following,* *(i). The expert commits to truthfully reporting the signals she observes* $D(t;t^*)$ *periods before;*

(ii). The delay function $D(t;t^*)$ *is strictly increasing in* t *for any* $t \in (t^*, t^D)$ *and strictly decreasing for any* $t \in (t^D, \theta(t^*))$.

Corollary 3 implies that the implementation divides the whole process into three phases. Suppose that the agent has not received a message confirming state ω_1. In the first phase, over the time interval $[0, t^*)$, the expert commits to timely and truthfully report. In the second phase, the delay strictly increases and reaches its maximum at the turning point t^D, where the agent forms belief q^P. In the third phase, the delay strictly shrinks until the hitting time $\theta(t^*)$, where the delay level completely vanishes. By the end of the third phase, the agent stops information acquisition and make decision immediately. Figure 8 shows how the optimal disclosure policy adjusts the delay of reporting with time.

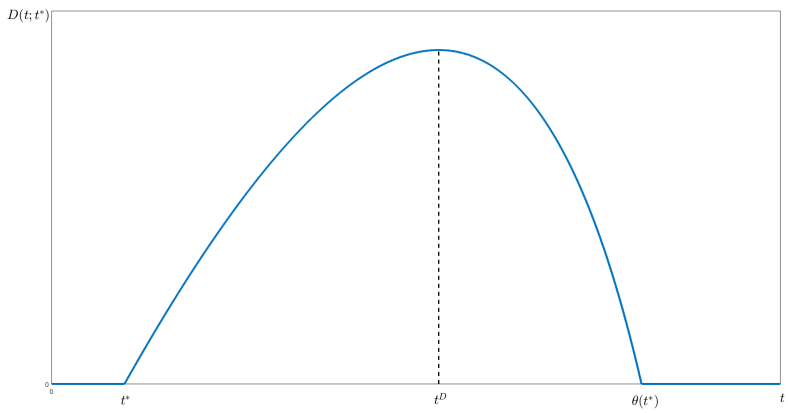

FIGURE 8. Delay function

To motivate the agent to explore further upon receiving s_a, the expert has to postpone the disclosure of breakthrough signal. Commitment power therefore plays a crucial role in determining our optimal policy. Furthermore, the expert balances the trade-off between timeliness of making decision at breakthrough signal and accurateness of decision-making upon not receiving breakthrough signal.

4.3.1. *Comparison to Ball [2019].* Even if the underlying state is persistent, the disclosure is on the generated signals. My paper can therefore be regarded as the disclosure of a Poisson process in a stopping game. Now I compare my paper to Ball [2019], who studies the disclosure of a Brownian motion in a repeated game. Even if he also shows that the optimal disclosure policy is truthful but delayed report, there are two critical differences, leaving behind the model formulation.

The first difference lies in dynamics of delay. In Ball [2019], a strictly positive delay emerges abruptly at the beginning of the game. Then, it gradually shrinks down to zero. After that, the delay is kept at zero. In my paper,the delay level is fixed at zero before the transition time t^* (in the first phase). Then, it strictly increases until the turning point t^D (in the second phase). Finally, it strictly shrinks down to zero, where the agent makes decision and the game stops.

Second, the restriction of disclosure policy to the set of truthful and delayed report is without loss of generality in Ball [2019] while it is, as I have shown, without loss of optimality in my paper. Specifically, the underlying stochastic process in Ball [2019] is mean-reverting Ornstein-Uhlenbeck process and utility functions are quadratic, therefore information disclosure boils down to precision control. Furthermore, the set of disclosure policy set of truthful and delayed report can recover any Bayesian plausible disclosure policy in terms of precision control (Please check Theorem 1 in Ball [2019]). However, since my model also involves designing what to be disclosed in addition to precision control, there exist many disclosure policies which can not be implemented through truthfully reporting with delay, like the policy τ^U in section 3 with the component on the cost side, i.e. persuasion. However, Corollary 3 implies that truthfully reporting with delay level $D(t;t^*)$ weakly dominates other disclosure policies.

4.4. **Sketch of the proof.** In the first step, I focus on a subset of disclosure policies.

Lemma 3. *Let Γ_1 be the set of reporting policies such that the agent is induced to stop acquiring information and makes a decision with probability 1. Furthermore, when he does this on the equilibrium path, then*

(1) The expert has already revealed all what she has observed.
(2) The agent's belief is either 0 or in the interval $[q^P, q^R]$.

There exists an optimal reporting policy in Γ_1.

First, the expert does not prefer the agent to delay decision forever since she is strictly impatient and gains utility only when decision is made.[9] Second, the expert prefers the decision to be made as accurately as possible. Therefore, on the equilibrium path,

[9]The assumption on the expert's preference is different from that in Ely and Szydlowski [2020], where the sender gains utility only when the receiver exerts effort.

she will not hold up any information when the agent makes decision.[10] Finally, their preferences are perfectly aligned at belief $p < q^P$ and at belief $p > q^R$. Both prefer to conduct further information acquisition at belief $p < q^P$ and prefer to stop further exploration at belief $p > q^R$.

Given any policy (π, M), denote $\boldsymbol{\gamma}^{(\pi,M)}$ the induced belief martingale. For any reporting policy in Γ_1, the agent has already stopped information acquisition and made decision with probability 1 at time $T(q^R;\mathcal{P})$. Denote $\tau^{(\pi,M)} = \gamma^{(\pi,M)}_{T(q^R;\mathcal{P})}$ the induced belief distribution at time $T(q^R;\mathcal{P})$. I call τ the *stopping belief distribution* since it also gives the distribution of beliefs at which the agent chooses to stop information acquisition. Further, let me denote $\lambda^{(\pi,M)}(t) = \gamma^{(\pi,M)}_t(0)$, the probability that the agent believes the state is exactly ω_1 at time t. I call λ *weight accumulation speed*. For any reporting policy in Γ_1, both the agent's and the expert's payoffs are fully determined by the pair (λ, τ).[11]

Definition 2. *A pair (λ, τ) is implementable if there exists a reporting policy in Γ_1 with weight accumulation speed λ and stopping belief distribution τ. Distribution τ is achievable if there exists an implementable pair (λ, τ).*

The second step of the proof derives the upper bound of λ for any achievable τ. I offer a sketch of the construction of the policy in Γ_1 and rigorous definition is in the appendix. The policy induces a belief martingale which either jumps to 0 or transits based on the following reporting process. It starts from the research process and then transits to the one one-step-ahead process at some transition time $\hat{t}(\tau)$. At time $T(q^P;\mathcal{P})$, it splits to generate beliefs with the corresponding weights, which recovers τ through the one-step-ahead belief process thereafter. Figure 9 illustrates the policy.

Let me denote $\boldsymbol{\gamma}^*$ as the induced belief belief distribution and $(\lambda^*(\cdot,\tau),\tau)$ the pair implemented by the constructed policy.

Lemma 4. *For any achievable τ, $\lambda^*(t;\tau) \geq \lambda(t)$, $\forall t \geq 0$, for any implementable (λ, τ).*

Let me offer an intuition of the proof in discrete time. Suppose there exists some disclosure policy $(\pi, M) \in \Gamma_1$ which induces a belief martingale $\boldsymbol{\gamma}$ and implements (λ, τ), such

[10]Please note that some information may be held up at off-equilibrium path to serve as a threat when the agent deviates.

[11]Specifically, the agent's payoff $U(\lambda,\tau)$ and the expert's payoff $V(\lambda,\tau)$ are given by

$$\begin{cases} U(\lambda,\tau) = m(0)\rho\int_{s=0}^{\infty} e^{-\rho s}\lambda(s)ds + \int_{q^P}^{q^R} e^{-\rho T(p;\mathcal{P})}m(p)d\tau - c\int_0^{T(q^R;\mathcal{P})} e^{-\rho t}\left[1-\lambda(t)-\mathbf{1}_{(t\geq T(q^P;\mathcal{P}))}\tau\left([q^P,p_t]\right)\right]dt \\ V(\lambda,\tau) = m(0)\rho\int_{s=0}^{\infty} e^{-\rho s}\lambda(s)ds + \int_{q^P}^{q^R} e^{-\rho T(p;\mathcal{P})}m(p)d\tau \end{cases}$$

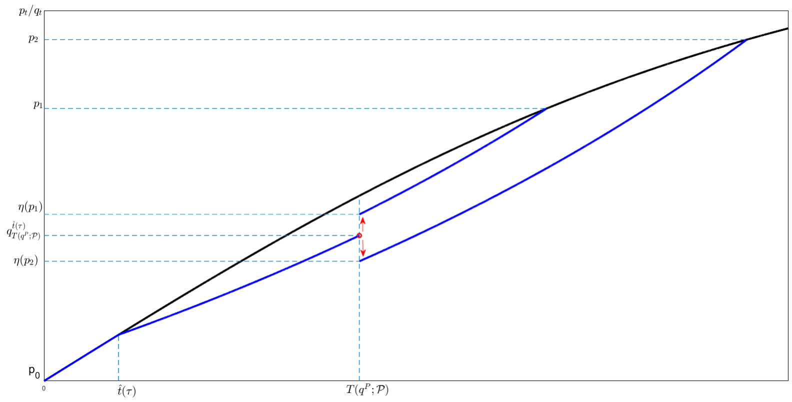

FIGURE 9. Reporting process achieving τ. The black curve is the research process $\mathcal{P}$ and the blue curve is the reporting process. The splitting at $T(q^P;\mathcal{P})$ is designed such that the belief martingale which either jumps to 0 or follows the reporting process can achieve τ.

that $\lambda(t) > \lambda^*(t;\tau)$. First, the weight accumulating speed under full information disclosure forms the upper bound among all plausible disclosure policies. Besides, $\lambda^*(\cdot;\tau)$ is the same as the one under full information disclosure before transition time $\hat{t}(\tau)$, therefore $t > \hat{t}(\tau)$.

Under disclosure policy (π, M), I call p active at time t if the agent still acquires information when he forms belief p at time t. Denote $P_{A,t}$ the set of active beliefs at time t. Moreover, define $p_{A,t} = \int_{P_{A,t}} \frac{p}{\gamma_t(P_{A,t})} d\gamma_t$ the expected active belief at time t. Similarly, one can also define $P^*_{A,t}$ the set of active beliefs and $p^*_{A,t}$ the expected active belief at time t under our prescribed disclosure policy. Bayesian plausibility then implies that

$$\begin{cases} \lambda(t)0 + \mathbf{1}_{(t\geq T(q^P;\mathcal{P}))} \int_{q^P}^{p_t} p d\tau + \gamma_t(P_{A,t})p_{A,t} = p_0 = \lambda^*(t;\tau)0 + \mathbf{1}_{(t\geq T(q^P;\mathcal{P}))} \int_{q^P}^{p_t} p d\tau + \gamma^*_t(P^*_{A,t})p^*_{A,t} \\ \lambda(t) + \mathbf{1}_{(t\geq T(q^P;\mathcal{P}))}\tau\left([q^P, p_t]\right) + \gamma_t(P_{A,t}) = 1 = \lambda^*(t;\tau) + \mathbf{1}_{(t\geq T(q^P;\mathcal{P}))}\tau\left([q^P, p_t]\right) + \gamma^*_t(P^*_{A,t}) \end{cases}$$

Therefore, we have that $\gamma_t(P_{A,t}) < \gamma^*_t(P^*_{A,t})$ and $p_{A,t} > p^*_{A,t}$. To motivate the agent to conduct further exploration at time t, the expert has to promise an expected continuation payoff strictly higher than $m(p^*_{A,t})$ under (π, M). However, $m(p^*_{A,t})$ is the future payoff generated under my prescribed disclosure policy. This is possible only if there exists another $t' > t$ such that $\lambda(t') > \lambda^*(t';\tau)$. If we iterate this procedure, then contradiction arises since $\lambda(t)$ and $\lambda^*(t;\tau)$ achieve the same upper bound $\tau(0)$ at the same time.

Let me denote Γ_2 the set of all constructed policies (obtained by varying all achievable τs). I then characterize the optimal achievable stopping belief distribution τ. I divide the characterization into two steps. In the first step, I derive the optimal stopping belief distribution τ given any transition time $\hat{t}$. In the second step, I derive the optimal transition time t^*.

Lemma 5. *For any transition time $\hat{t}$, it is optimal to choose supp*$(\tau) = \{0, p_{\theta(\hat{t})}\}$ *among all reporting policies in* Γ_2.

Lemma 5 implies that no further splitting of beliefs should be conducted at time $T(q^P; \mathcal{P})$. Kamenica and Gentzkow [2011] implies that the optimal τ is characterized through concavifying the expert's continuation payoff at time $T(q^P; \mathcal{P})$. I show that such payoff function is strictly concave.

Let me denote Γ_3 the subset of all policies in Γ_2 without belief splitting at $T(q^P; \mathcal{P})$. The payoffs under these reporting policies are completely determined by the transition time $\hat{t}$.

Lemma 6. *The optimal transition time is unique and is fully determined by the first order condition.*

Lemma 6 follows since the expert's payoff is strictly concave with respect to the transition time $\hat{t}$. Combining Lemma 3, 4, 5 and Corollary 1, Theorem 1 then follows.

4.5. **General payoff structure.** In this subsection, we consider a general payoff structure shown in Table 5.

TABLE 5. General payoff structure

Payoff \ Policy / State	a_1	a_2
ω_1	$u(a_1; \omega_1)$	$u(a_2; \omega_1)$
ω_2	$u(a_2; \omega_1)$	$u(a_2; \omega_2)$

Moreover, we assume a_1 (a_2) is optimal at state ω_1 (ω_2), i.e.

$$\begin{cases} u(a_1; \omega_1) > u(a_2; \omega_1) \\ u(a_1; \omega_2) < u(a_2; \omega_1) \end{cases}.$$

Similarly, the one-step-ahead process $\mathcal{Q}$ is defined as

$$-cdt + e^{-\rho dt}\left[\frac{q_t}{q_{t+dt}} m(q_{t+dt}) + \left(1 - \frac{q_t}{q_{t+dt}}\right) m(0)\right] = m(q_t).$$

We then derive the one-step-ahead process as

$$\dot{q}_t = \frac{dq_t}{dt} = \frac{q_t[c + \rho m(q_t)]}{u(a_1; \omega_1) - u(a_2; \omega_1)}. \tag{8}$$

We can also derive the closed-form of the process $\mathcal{Q}$ as

$$q_t = \frac{[\rho u(a_2;\omega_1)+c]\exp\left(\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}t\right)}{\rho[u(a_2;\omega_2)-u(a_2;\omega_1)]\left[1-\exp\left(\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}t\right)\right]+\frac{\rho u(a_2;\omega_1)+c}{q_0}}.$$

Similarly, we define $\mathcal{Q}^{\hat{t}}$ the one-step-ahead process which passes $p_{\hat{t}}$ at time $\hat{t}$. We also define $\hat{\mathcal{Q}}(q)$ the one-step-ahead process with starting point q. The next result shows that the main results in Theorem 1 still hold if $u(a_2;\omega_2)\geq u(a_1;\omega_1)$.

Proposition 2. *If $u(a_2;\omega_2)\geq u(a_1;\omega_1)$, then there exists some t^* such that one optimal policy induces a belief martingale which either jumps to belief 0 or follows the process $\min\{\mathcal{P},\mathcal{Q}^{t^*}\}$.*

If $u(a_2;\omega_2)<u(a_1;\omega_1)$, then the continuation payoff at time $T(q^P;\mathcal{P})$ may not be concave. We therefore need to adopt the persuasion channel to concavify this continuation payoff. The following proposition describes the optimal policy when $u(a_2;\omega_2)<u(a_1;\omega_1)$.

Proposition 3. *If $u(a_2;\omega_2)<u(a_1;\omega_1)$, then there exists some (t_1^*,t_2^*,q^*) such that one optimal policy induces a belief martingale which either jumps to 0 or*

(1) At time interval $[0,t_2^*)$, it follows $\min\{\mathcal{P},\mathcal{Q}^{t_1^*}\}$;
(2) At time t_2^*, it is split between $p_{t_2^*}$ and q^*;
(3) After time t_2^*, it then follows $\min\left\{\mathcal{P},\hat{\mathcal{Q}}(q^*)\right\}$.

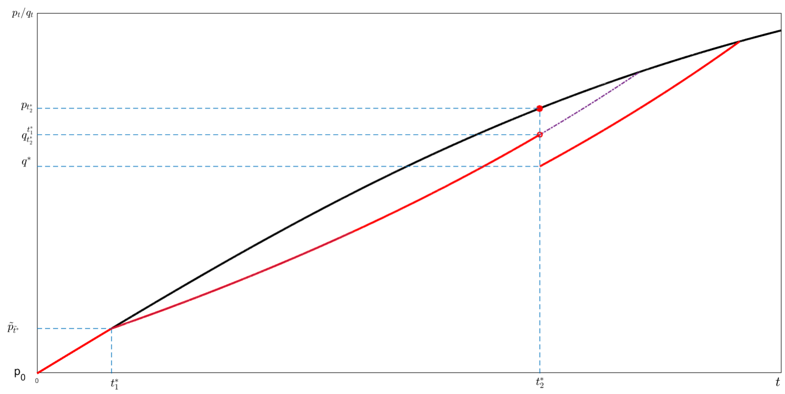

FIGURE 10. Optimal Reporting Process

To motivate the agent to acquire information upon receiving s_a, the expert has to postpone the disclosure of s_b. However, when $u(a_1;\omega_1)>u(a_2;\omega_2)$, the expert sacrifices more when she puts off disclosing s_b. Therefore, she prefers to put off the time at which information value hoarding starts, which then decreases the pool of hoarded information value. Facing limited amount of hoarded information value, the expert turns to persuasion to lower agent's belief and therefore the opportunity cost. We show that there exists at most one instant at which persuasion is conducted. Figure 10 illustrates the reporting process.

5. Extension

5.1. A more general signal structure. In this section, we allow for more general information structure with breakthrough signals for both states. The signal space is $S = \{b_1, b_2, a\}$ and the information structure is given by Table 6.

Information structure: State \ Signal	b_1	b_2	a
ω_1	$\lambda(1-\psi)dt$	0	$1-\lambda(1-\psi)dt$
ω_2	0	$\lambda\psi dt$	$1-\lambda\psi dt$

Table 6. Generalized information structure. The signal b_1 (resp. b_2) fully reveals the state ω_1(resp. ω_2)

Similarly, let me define *generalized research process* $\tilde{\mathcal{P}}$ as the one which tracks the evolution of beliefs, given that breakthrough signals have not been received. The corresponding ODE is

$$d\tilde{p}_t = \lambda\tilde{p}_t(1-\tilde{p}_t)(1-2\psi)dt \tag{9}$$

It is without loss of generality to assume $\psi < \frac{1}{2}$. Further, one can derive the threshold $\tilde{q}^P = \frac{\lambda(1-\psi)-c}{\rho+(1-\psi)}$ (resp. $\tilde{q}^R = \frac{\lambda(1-\psi)}{\rho+(1-\psi)}$), beyond which the agent (resp. the expert) prefers to stop information acquisition and make choice right away.

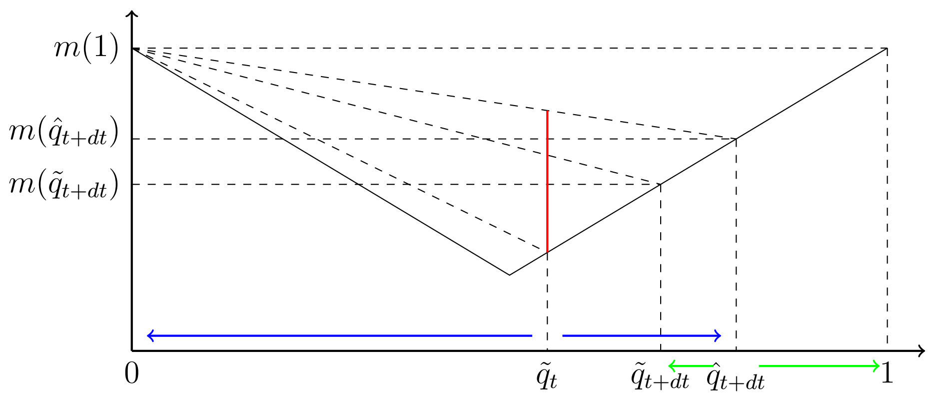

Figure 11. Generalized one-step-ahead process. In the first step, $\tilde{q}_t$ is split between 0 and $\hat{q}_{t+dt}$, depicted by the blue arrow. Such splitting generates information value of red line, which just covers the opportunity cost. In the second step, $\hat{q}_{t+dt}$ is split between 1 and $\tilde{q}_{t+dt}$, depicted by the green arrow.

Given some knowledge gap $(\tilde{p}_t, \tilde{q}_t)$ at time t, the expert will privately observe the signal b_2 with probability $\frac{\tilde{q}_t}{\tilde{p}_t}\tilde{p}_t\lambda\psi dt$ if the agent acquires information at t. Let me define a belief martingale by the following two-step procedure. The starting belief is $\tilde{q}_t \geq \max\left\{p_0, \frac{1}{2}\right\}$. In the first step, it either jumps to 0 or moves up to $\hat{q}_{t+dt}$, generating information value which just covers the opportunity cost. The increment $d\hat{q}_t = \hat{q}_{t+dt} - \hat{q}_t$ is given by by Equation (3). In the second step, $\hat{q}_t$ either jumps to 1 with maximal probability, i.e. $\lambda\psi\tilde{q}_t dt$, or moves down to $\tilde{q}_{t+dt}$. This *generalized one-step-ahead process* has an increment $d\tilde{q}_t = \tilde{q}_{t+dt} - \tilde{q}_t$, given by

$$d\tilde{q}_t = (\rho + \lambda\psi)\tilde{q}_t(\tilde{q}_t - \underline{q})dt, \tag{10}$$

where $\underline{q} = \frac{\lambda\psi - c}{\lambda\psi + \rho}$. The illustration of this belief martingale is depicted in Figure 11. For any valid (t, q), let me denote $\tilde{Q}(t, q)$ the generalized one-step-ahead process which passes q at t.

Proposition 4. *One optimal policy $(\tilde{\pi}^*, \tilde{M}^*)$ is specified as follows. There exists $(\tilde{t}^*, \theta^*, \tilde{q}^*)$ such that the splitting $\left(\tilde{q}_{T(\tilde{q}^P;\tilde{\mathcal{P}})}(\theta^*, \tilde{p}_{\theta^*}), \tilde{q}_{T(\tilde{q}^P;\tilde{\mathcal{P}})}(\theta^*, \tilde{q}^*)\right)$ of the belief $\tilde{q}_{T(\tilde{q}^P;\tilde{\mathcal{P}})}(\tilde{t}^*, \tilde{p}_{\tilde{t}^*})$ concavifies the expert's continuation payoff at time $T(\tilde{q}^P; \tilde{\mathcal{P}})$. The expert reveals b_2 perfectly whenever she has privately observed it. Otherwise, the induced belief martingale either jumps to 0 or transits according to the following reporting process,*

- For any $t \in [0, \tilde{t}^*)$, follow $\tilde{\mathcal{P}}$.
- For any $t \in [\tilde{t}^*, +\infty)$, follow the generalized one-step-ahead process until it hits $\tilde{\mathcal{P}}$, except that a splitting is conducted between $\tilde{p}_{\theta^*}$ and $\tilde{q}^*$ at time θ^*.

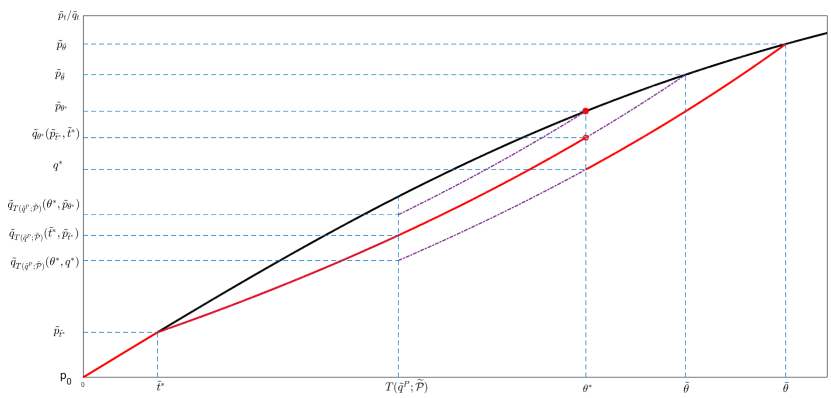

FIGURE 12. Optimal reporting process under generalized information structure. The black curve is the research process. The segment of red curve after the transition time $\tilde{t}^*$ is the generalized one-step-ahead process. The red curve is the optimal reporting process.

There are two key instants of the optimal policy in Proposition 4. The first key instant is $\tilde{t}^*$, when the expert begins to hoard information value. The second key instant is θ^*, when the expert persuades the agent to lower his payoff of immediate decision-making. Besides, such splitting is related to concavification of expert's utility at time $T(\tilde{q}^P;\tilde{\mathcal{P}})$. Figure 12 illustrates the optimal policy in Proposition 4.

References

C. Angelucci. Motivating agents to acquire information. *Available at SSRN 2905367*, 2017.

P. H. Au. Dynamic information disclosure. *The RAND Journal of Economics*, 46(4):791–823, 2015. ISSN 07416261. URL `http://www.jstor.org/stable/43895617`.

R. J. Aumann, M. Maschler, and R. E. Stearns. *Repeated games with incomplete information*. MIT press, 1995.

I. Ball. Dynamic information provision: Rewarding the past and guiding the future. *Available at SSRN 3103127*, 2019.

P. Basu. Dynamic bayesian persuasion with a privately informed receiver. 2017.

J. Bizzotto, J. Rüdiger, and A. Vigier. Dynamic persuasion with outside information. *American Economic Journal: Microeconomics*, Forthcoming.

H. Chade and N. Kovrijnykh. Delegated information acquisition with moral hazard. *Journal of Economic Theory*, 162:55 – 92, 2016. ISSN 0022-0531. doi: https://doi.org/10.1016/j.jet.2015.12.003. URL `http://www.sciencedirect.com/science/article/pii/S0022053115002057`.

Y.-K. Che and J. Hörner. Recommender Systems as Mechanisms for Social Learning*. *The Quarterly Journal of Economics*, 133(2):871–925, 12 2017. ISSN 0033-5533. doi: 10.1093/qje/qjx044. URL `https://doi.org/10.1093/qje/qjx044`.

Y.-K. Che, K. Kim, and K. Mierendorff. Keeping the listener engaged: a dynamic model of bayesian persuasion. *arXiv preprint arXiv:2003.07338*, 2020.

A. A. Clark. Contracts for information acquisition. 2016.

L. Doval and J. Ely. Sequential information design. 2018.

J. C. Ely. Beeps. *American Economic Review*, 107(1): 31–53, January 2017. doi: 10.1257/aer.20150218. URL `http://www.aeaweb.org/articles?id=10.1257/aer.20150218`.

J. C. Ely and M. Szydlowski. Moving the goalposts. *Journal of Political Economy*, 128 (2):468–506, 2020. doi: 10.1086/704387. URL `https://doi.org/10.1086/704387`.

M. Escudé and L. Sinander. Strategic research funding. *Working Paper*, 2020.

J. Glazer, I. Kremer, and M. Perry. The wisdom of the crowd when acquiring information is costly. *Management Science*, Forthcoming.

E. Henry and M. Ottaviani. Research and the approval process: The organization of persuasion. *American Economic Review*, 109(3):911–55, March 2019. doi: 10.1257/aer.20171919. URL https://www.aeaweb.org/articles?id=10.1257/aer.20171919.

E. Kamenica and M. Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, October 2011. doi: 10.1257/aer.101.6.2590. URL http://www.aeaweb.org/articles?id=10.1257/aer.101.6.2590.

M. Makris and L. Renou. Information design in multi-stage games. Working Papers 861, Queen Mary University of London, School of Economics and Finance, June 2018. URL https://ideas.repec.org/p/qmw/qmwecw/861.html.

R. B. Myerson. Multistage games with communication. *Econometrica: Journal of the Econometric Society*, pages 323–358, 1986.

D. Orlov, A. Skrzypacz, and P. Zryumov. Persuading the principal to wait. *Journal of Political Economy*, 128(7):2542–2578, 2020. doi: 10.1086/706687. URL https://doi.org/10.1086/706687.

J. Renault, E. Solan, and N. Vieille. Optimal dynamic information provision. *Games and Economic Behavior*, 104:329 – 349, 2017. ISSN 0899-8256. doi: https://doi.org/10.1016/j.geb.2017.04.010. URL http://www.sciencedirect.com/science/article/pii/S089982561730074X.

A. Smolin. Dynamic evaluation design. *AMERICAN ECONOMIC JOURNAL: MICROECONOMICS*, Forthcoming.

T. Sugaya and A. Wolitzky. The revelation principle in multistage games. 2019.

M. Yang. Optimality of Debt under Flexible Information Acquisition. *The Review of Economic Studies*, 87(1):487–536, 07 2019. ISSN 0034-6527. doi: 10.1093/restud/rdz035. URL https://doi.org/10.1093/restud/rdz035.

J. Zhao and W. Zhao. *External and Internal Information Acquisition*. HEC Paris, Apr. 2019.

W. Zhao, C. Mezzetti, L. Renou, and T. Tomala. Dynamic contracting over information. *Working Paper*, 2020.

Appendix A. Lower bound of prior $\underline{p}_0$

Consider the case where the prior $p_0 < \frac{1}{2}$. Note that if the prior p_0 is too small, then the agent prefers to make decision immediately instead of acquiring information. There are two reasons underlying the argument. For one thing, little uncertainty is involved, so the payoff of immediate decision making is large, leading to large time cost; For another, it takes too long for the learning process to create strictly positive information value (or value), which highly discounts the created information value. Let me define $\mathcal{P}(p_0)$ as the research progress process with the starting belief p_0. Furthermore, $\underline{p}_0 \in (0, \frac{1}{2})$ is defined as the solution to the following equation

$$m(\underline{p}_0) = \rho \int_{t=0}^{\infty} e^{-\rho t} \left(1 - \max\left\{\frac{\underline{p}_0}{p_t(\underline{p}_0)}, \frac{\underline{p}_0}{q^P}\right\}\right) m(0)dt + e^{-\rho T(q^P;\mathcal{P}(\underline{p}_0))} \frac{\underline{p}_0}{q^P} m(q^P) - c \int_0^{T(q^P;\mathcal{P}(\underline{p}_0))} e^{-\rho t} \frac{\underline{p}_0}{p_t(\underline{p}_0)} dt \quad (11)$$

If $p_0 < \underline{p}_0$, then the policymaker will never fund research and therefore always make decision immediately.

Our next Lemma shows that $\underline{p}_0$ is both well and uniquely defined.

Lemma 7. *There always exists a unique solution to the Equation* (11).

Proof. For any p_0, define function $M(p_0)$ as

$$M(p_0) := \rho \int_{t=0}^{\infty} e^{-\rho t} \left(1 - \max\left\{\frac{p_0}{p_t(p_0)}, \frac{p_0}{q^P}\right\}\right) m(0)dt + e^{-\rho T(q^P;\mathcal{P})} \frac{p_0}{q^P} m(q^P) - c \int_{t=0}^{T(q^P;\mathcal{P})} e^{-\rho t} \frac{p_0}{p_t(p_0)} dt$$

Since $M(1/2) > m(1/2)$ and $M(0) < m(0)$, there always exists some p_0 such that $M(p_0) = m(p_0)$ due to the continuity $M(p_0)$.

Furthermore, for any pair of (p_0, p_0') such that $0 < p_0 < p_0' < \frac{1}{2}$, one have the following

$$\begin{aligned} M(p_0) &= \rho \int_{t=0}^{\infty} e^{-\rho t} \left(1 - \max\left\{\frac{p_0}{p_t(p_0)}, \frac{p_0}{p_0'}\right\}\right) m(0)dt + e^{-\rho T(p_0';\mathcal{P})} \frac{p_0}{p_0'} M(p_0') - c \int_{t=0}^{T(p_0';\mathcal{P})} e^{-\rho t} \frac{p_0}{p_t(p_0)} dt \\ &= \rho \int_{t=0}^{\infty} e^{-\rho t} \left(1 - \max\left\{\frac{p_0}{p_t(p_0)}, \frac{p_0}{p_0'}\right\}\right) m(0)dt + e^{-\rho T(p_0';\mathcal{P})} \frac{p_0}{p_0'} m(p_0') - c \int_{t=0}^{T(p_0';\mathcal{P})} e^{-\rho t} \frac{p_0}{p_t(p_0)} dt \\ &+ e^{-\rho T(p_0';\mathcal{P})} \frac{p_0}{p_0'} \left(M(p_0') - m(p_0')\right) < m(p_0) + e^{-\rho T(p_0';\mathcal{P})} \frac{p_0}{p_0'} \left(M(p_0') - m(p_0')\right) \end{aligned}$$

The inequality is due to the fact that

$$\begin{aligned} &\rho \int_{t=0}^{\infty} e^{-\rho t} \left(1 - \max\left\{\frac{p_0}{p_t(p_0)}, \frac{p_0}{p_0'}\right\}\right) m(0)dt + e^{-\rho T(p_0';\mathcal{P})} \frac{p_0}{p_0'} m(p_0') \\ &< \rho \int_{t=0}^{\infty} e^{-\rho t} \left(1 - \frac{p_0}{p_0'}\right) m(0)dt + \frac{p_0}{p_0'} m(p_0') = m(p_0) \end{aligned}$$

Therefore, if $M(p_0') \leq m(p_0')$, then $M(p_0) < m(p_0)$. $\square$

APPENDIX B. TRADE-OFFS BALANCING

As is shown in Lemma 2, the hitting time $\theta(\hat{t})$ is decreasing in the transition time $\hat{t}$. Put it in other words, to motivate the agent to make more accurate decision, the expert has to begin hoarding information value earlier. This may lead to unnecessary delay in decision making since the expert has to delay disclosing private signal of breakthrough. Therefore, the optimal transition time t^* is set so as to balance the trade-off between earlier information hoarding and more accurate decision making. Specifically, one can represent the expert's payoff upon the reporting policy, specified in Theorem 1 but with transition time $\hat{t}$, as

$$\begin{aligned} W(\hat{t}) &= m(0)\int_0^{\theta(\hat{t})} e^{-\rho t}\frac{p_0}{q_t(\hat{t})}\frac{dq_t(\hat{t})}{q_t(\hat{t})dt}dt + e^{-\rho\theta(\hat{t})}\frac{p_0}{p_{\theta(\hat{t})}}m(p_{\theta(\hat{t})}) &(12)\\ &= \rho m(0)\int_0^{\infty} e^{-\rho t}\left(1-\frac{p_0}{\min\left\{q_t(\hat{t}),p_{\theta(\hat{t})}\right\}}\right)dt + e^{-\rho\theta(\hat{t})}\frac{p_0}{p_{\theta(\hat{t})}}m(p_{\theta(\hat{t})}) &(13)\end{aligned}$$

In the first line of the equation above, the term $\frac{p_0}{q_t(\hat{t})}\frac{dq_t(\hat{t})}{q_t(\hat{t})dt}$ gives the probability that the belief martingale jumps to belief zero exactly at time $(t, t+dt]$; In the second line, the term $1-\frac{p_0}{\min\left\{q_t(\hat{t}),p_{\theta(\hat{t})}\right\}}$ is the accumulating[12] probability that the belief martingale has already jumped to belief zero at time t. If the belief martingale is given by what is specified in Theorem 1 with transition time $\hat{t}$, then such probability is fixed at $1-\frac{p_0}{p_{\theta(\hat{t})}}$ since hitting time $\theta(\hat{t})$. By the definition of reporting process $\mathcal{Q}(\hat{t})$, we have that $1-\frac{p_0}{\min\left\{q_t(\hat{t}),p_{\theta(\hat{t})}\right\}} \leq 1-\frac{p_0}{\min\left\{p_t,p_{\theta(\hat{t})}\right\}}$, with strict inequality during time interval $(\hat{t},\theta(\hat{t}))$. Such gap in accumulating probability is exactly brought up by unnecessary delayed decision making.

By the definition of reporting process $\mathcal{Q}(\hat{t})$, the policymaker is indifferent between making decision at any moment during in time interval $[\hat{t},\theta(\hat{t})]$, which therefore implies that

$$\underbrace{\int_{\hat{t}}^{\theta(\hat{t})} e^{-\rho t}\frac{p_0}{q_t(\hat{t})}\frac{dq_t(\hat{t})}{q_t(\hat{t})dt}dt + e^{-\rho\theta(\hat{t})}\frac{p_0}{p_{\theta(\hat{t})}}m(p_{\theta(\hat{t})})}_{\text{Discounted payoff if making decision at } \theta(\hat{t})} - \underbrace{c\int_{\hat{t}}^{\theta(\hat{t})} e^{-\rho t}\frac{p_0}{q_t(\hat{t})}dt}_{\text{Discounted research cost if funding until } \theta\hat{t})} = \underbrace{e^{-\rho\hat{t}}\frac{p_0}{p_{\hat{t}}}m(p_{\hat{t}})}_{\text{Discounted payoff if making decision at } \hat{t}}.$$

Please note that the policymaker only funds the expert if he has not received a message confirming the state being ω_1 and the probability of the event at time t is given by $\frac{p_0}{q_t(\hat{t})}$. Introduce the equation above into (12), one can derive another form of the expert's payoff as

$$W(\hat{t}) = \underbrace{m(0)\int_0^{\hat{t}} e^{-\rho t}\frac{p_0}{p_t}\frac{dp_t}{p_t dt}dt + e^{-\rho\hat{t}}\frac{p_0}{p_{\hat{t}}}m(p_{\hat{t}})}_{\text{Discounted payoff of early making decision at } \hat{t}} + \underbrace{c\int_{\hat{t}}^{\theta(\hat{t})} e^{-\rho t}\frac{p_0}{q_t(\hat{t})}dt}_{\text{Research grant occupation}} \tag{14}$$

One can consider an artificial scenario in which the expert is able to occupy the research grant in addition to policy impact on social welfare. The expert chooses when to stop information acquisition and begin to embezzle research funds. The optimal transition

[12]By "accumulating", I mean that this probability is always increasing in time t;

time t^* is determined through balancing the payoff of early and therefore less accurate decision making, and longer periods of research grants embezzlement.

APPENDIX C. MISSING PROOFS

C.1. **Derivation of research progress process.** If $dp_t = \lambda p_t(1-p_t)dt$, then we have that

$$\begin{aligned}
&\frac{dp_t}{p_t} + \frac{dp_t}{1-p_t} = \lambda dt \Rightarrow d\ln p_t - d\ln(1-p_t) = d\lambda t \\
&\Rightarrow d\ln\frac{p_t}{1-p_t} = d\lambda t \Rightarrow \frac{p_t}{1-p_t} = \frac{p_0}{1-p_0}e^{\lambda t} \\
&\Rightarrow p_t = (1-p_t)\frac{p_0}{1-p_0}e^{\lambda t} \Rightarrow (1+\frac{p_0}{1-p_0}e^{\lambda t})p_t = \frac{p_0}{1-p_0}e^{\lambda t} \\
&\Rightarrow p_t = \frac{\frac{p_0}{1-p_0}e^{\lambda t}}{1+\frac{p_0}{1-p_0}e^{\lambda t}} = \frac{p_0 e^{\lambda t}}{p_0 e^{\lambda t} + (1-p_0)}
\end{aligned}$$

C.2. **Proof of Lemma 1.** Denote W_t as payoff to the agent, who perfectly observes research progress, if he chooses to stop research funding and make decision at time t, even if he has not observed break-through signals. One then have that

$$\begin{aligned}
&dW_t = W_{t+dt} - W_t \\
&= e^{-\rho t}\frac{p_0}{p_t}\left\{e^{-\rho dt}\left[\lambda(1-p)dt\ m(0) + [1-\lambda(1-p)dt]m(p+\lambda p(1-p)dt) - m(p)\right] - cdt - \rho m(p)dt\right\}
\end{aligned}$$

By the definition of q^P, one then have that $dW_t > 0$ when $p_t < q^P$ and $dW_t < 0$ when $p_t > q^P$. Therefore W_t reaches its maximal value when $p_t = q^P$.

C.3. **Proof of Proposition 1.** In this subsection, we at first use the belief-based approach developed by Kamenica and Gentzkow [2011] to characterize the set of Bayesian plausible information structure under imperfect information. Secondly, we then offer a standard proof of Proposition 1.

C.3.1. *Bayesian plausible information structure.* Denote V_t as payoff to the expert, who is delegated of decision making in research funding, if she chooses to stop research funding and make decision at time t. Similarly, one can show that V_t reaches its maximal value when $p_t = q^R$.

Before proving proposition 1, we at first extend Bayesian plausibility condition proposed by Kamenica and Gentzkow [2011] to scenario where sender is imperfectly informed.

Lemma 8. *Suppose both sender (She) and receiver (He) share a common prior p_0. Besides, the sender is further privately informed with information structure $(\pi, \mathcal{S} = \{s_1, s_2\})$. If receiving signal s_1, she updates her belief to p_1. Otherwise, she updates belief $p_2 \geq p_1$. Finally, she commits to some disclosure policy $(\tilde{\pi}, M)$ before observing private signals. Belief martingale γ can be induced by some disclosure policy if and only if supp$(\gamma) \in [p_1, p_2]$ and $\int_{p_1}^{p_2} \gamma(p)pdp = p_0$.*

Proof. Denote belief martingale $\alpha = (\alpha(p_1), \alpha(p_2))$ as the one induced by information structure $(\pi, \mathcal{S})$. Bayesian plausibility $\alpha(p_1)p_1 + \alpha(p_2)p_2 = p_0$ implies that

$$\begin{cases} \alpha(p_1) = \frac{p_2 - p_0}{p_2 - p_1} \\ \alpha(p_2) = \frac{p_0 - p_1}{p_2 - p_1} \end{cases}$$

In the first step, we prove "if" direction. For one thing, the receiver receives message $m \in M$ with probability $\gamma(m) = \alpha(p_1)\tilde{\pi}(m|s_1) + \alpha(p_2)\tilde{\pi}(m|s_2)$; For another, based on Bayesian updating principle, the receiver updates his belief to

$$p_m = \frac{\alpha(p_2)\tilde{\pi}(m|s_2)}{\alpha(p_1)\tilde{\pi}(m|s_1) + \alpha(p_2)\tilde{\pi}(m|s_2)} p_2 + \frac{\alpha(p_1)\tilde{\pi}(m|s_1)}{\alpha(p_1)\tilde{\pi}(m|s_1) + \alpha(p_2)\tilde{\pi}(m|s_2)} p_1 \in [p_1, p_2].$$

Moreover, we also have the following equation

$$\int_{m \in M} \gamma(m) p_m dm = \int_{m \in M} \gamma(m) \left[\frac{\alpha(p_2)\tilde{\pi}(m|s_2)}{\alpha(p_1)\tilde{\pi}(m|s_1) + \alpha(p_2)\tilde{\pi}(m|s_2)} p_2 + \frac{\alpha(p_1)\tilde{\pi}(m|s_1)}{\alpha(p_1)\tilde{\pi}(m|s_1) + \alpha(p_2)\tilde{\pi}(m|s_2)} p_1 \right] dm$$
$$= \alpha(p_2)p_2 \int_{m \in M} \tilde{\pi}(m|s_2)dm + \alpha(p_1)p_1 \int_{m \in M} \tilde{\pi}(m|s_1)dm = \alpha(p_2)p_2 + \alpha(p_1)p_1 = p_0$$

In the second step, we prove "only if" direction. For any belief martingale γ, let me define an information structure $(\tilde{\pi}, [p_1, p_2])$ as follows

$$\begin{cases} \tilde{\pi}(p|s_1) = \dfrac{\gamma(p)}{\alpha(p_1)} \dfrac{p_2 - p}{p_2 - p_1} = \dfrac{\gamma(p)}{\frac{p_2 - p_0}{p_2 - p_1}} \dfrac{p_2 - p}{p_2 - p_1} = \gamma(p) \dfrac{p_2 - p}{p_2 - p_0} \\ \tilde{\pi}(p|s_2) = \dfrac{\gamma(p)}{\alpha(p_2)} \dfrac{p - p_1}{p_2 - p_1} = \dfrac{\gamma(p)}{\frac{p_0 - p_1}{p_2 - p_1}} \dfrac{p - p_1}{p_2 - p_1} = \gamma(p) \dfrac{p - p_1}{p_0 - p_1} \end{cases}, \forall p \in [p_1, p_2] \tag{15}$$

At first, let me verify validity of the information structure $(\tilde{\pi}, [p_1, p_2])$. To show it, we have that

$$\int_{p \in [p_1, p_2]} \tilde{\pi}(p|s_1)dp = \int_{p \in [p_1, p_2]} \frac{\gamma(p)}{\alpha(p_1)} \frac{p_2 - p}{p_2 - p_1} dp = \frac{p_2 \int_{p \in [p_1, p_2]} \gamma(p)dp - \int_{p \in [p_1, p_2]} \gamma(p)pdp}{\alpha(p_1)(p_2 - p_1)}.$$
$$= \frac{p_2 - p_0}{\alpha(p_1)(p_2 - p_1)} = \frac{p_2 - p_0}{p_2 - p_0} = 1$$

Similarly, one can also show that $\int_{p \in [p_1, p_2]} \tilde{\pi}(p|s_2)dp = 1$. Secondly, we show that such information structure $(\tilde{\pi}, [p_1, p_2])$ induces the belief martingale γ. This is true since if the receiver receives message $p \in [p_1, p_2]$, then he updates his belief to

$$\frac{\alpha(p_1)\tilde{\pi}(p|s_1)}{\alpha(p_1)\tilde{\pi}(p|s_1)+\alpha(p_2)\tilde{\pi}(p|s_2)}p_1+\frac{\alpha(p_2)\tilde{\pi}(p|s_2)}{\alpha(p_1)\tilde{\pi}(p|s_1)+\alpha(p_2)\tilde{\pi}(p|s_2)}p_2$$
$$=\frac{\alpha(p_1)\frac{\gamma(p)}{\alpha(p_1)}\frac{p_2-p}{p_2-p_1}}{\alpha(p_1)\frac{\gamma(p)}{\alpha(p_1)}\frac{p_2-p}{p_2-p_1}+\alpha(p_2)\frac{\gamma(p)}{\alpha(p_2)}\frac{p-p_1}{p_2-p_1}}p_1+\frac{\alpha(p_2)\frac{\gamma(p)}{\alpha(p_2)}\frac{p-p_1}{p_2-p_1}}{\alpha(p_1)\frac{\gamma(p)}{\alpha(p_1)}\frac{p_2-p}{p_2-p_1}+\alpha(p_2)\frac{\gamma(p)}{\alpha(p_2)}\frac{p-p_1}{p_2-p_1}}p_2$$
$$=\frac{\frac{p_2-p}{p_2-p_1}p_1+\frac{p-p_1}{p_2-p_1}p_2}{\frac{p_2-p}{p_2-p_1}+\frac{p-p_1}{p_2-p_1}}=p$$

□

C.3.2. *Standard proof of Proposition 1.* The responsiveness and directness, as mentioned in texts and which will be further elaborated in later texts, implies that second round information disclosure will always be full information disclosure while the first round information disclosure involves at most three messages, one message for each action (including action a_1, a_2 and acquiring information). Furthermore, since both the expert and the agent share the same preference in choice-making, the agent should form a belief 0 upon receiving message at which he chooses action a_1. Similarly, he forms a posterior $f(p_0)$ upon receiving message at which he chooses action a_2. Let me denote the belief formed by the agent who receives the message and decides to acquire information for another period as p^F. Besides, let me denote the belief martingale induced by the first round information disclosure as $\boldsymbol{\gamma} = (\gamma(0), \gamma(p^F), \gamma(f(p_0)))$.

In the first step, we are going to show that it is without loss of **optimality** to restrict $p^F = q^*$. To prove this argument, for any belief martingale γ which can be induced by some disclosure policy, we construct another belief martingale γ' such that $\text{supp}(\gamma') \subseteq \{0, q^*, f(p_0)\}$. We then divide our proofs into three steps. First, we will show that the belief martingale γ' can be induced by some perturbed disclosure policy. Second, we show that the disclosure policy can motivate the agent to acquire information in the first period of the game. Finally, the expert weakly prefers belief martingale γ' to γ. Incentive compatibility implies that $p^F \leq q^*$. Let me consider a perturbed first round information disclosure policy with induced belief martingale γ' such that $\text{supp}(\gamma') \subseteq \{0, q^*, f(p_0)\}$ and

$$\begin{cases} \gamma'(0) = \gamma(0) + \gamma(p^F)\left(1 - \frac{p^F}{q^*}\right) \\ \gamma'(q^*) = \gamma(p^F)\frac{p^F}{q^*} \\ \gamma'(f(p_0)) = \gamma(f(p_0)) \end{cases}.$$

We argue that the belief martingale γ' can be induced by some perturbed disclosure policy. First, we have the following equation

$$\begin{aligned}\gamma'(0)+\gamma'(q^*)+\gamma'(f(p_0)) &= \gamma(0)+\gamma(p^F)\left(1-\frac{p^F}{q^*}\right)+\gamma(p^F)\frac{p^F}{q^*}+\gamma(f(p_0))\\ &= \gamma(0)+\gamma(p^F)+\gamma(f(p_0))=1\end{aligned}$$

Second, we also have Bayesian plausibility constraint as follows

$$\begin{aligned}\gamma'(0)\cdot 0+\gamma'(q^*)\cdot q^*+\gamma'(f(p_0))\cdot f(p_0) &= \left[\gamma(0)+\gamma(p^F)\left(1-\frac{p^F}{q^*}\right)\right]\cdot 0+\gamma(p^F)\frac{p^F}{q^*}\cdot q^*+\gamma(f(p_0))\cdot f(p_0)\\ &= \gamma(0)\cdot 0+\gamma(p^F)\cdot p^F+\gamma(f(p_0))\cdot f(p_0)=p_0\end{aligned}$$

Therefore, Lemma 8 implies that the belief martingale γ' can be induced by some disclosure policy. For any belief martingale γ, let me denote the expert's and agent's payoff as $V(\gamma)$ and $W(\gamma)$ respectively. Let me denote V^F (W^F) and V^N (W^N) the expert's (agent's) payoff under full and no information disclosure, respectively. Next, we argue that the belief martingale γ' is plausible in the sense that $W(\gamma') > W^N$. This is due to the following,

$$\begin{aligned}
W(\gamma') &= \left(1-\frac{\gamma'(q^*)q^*}{p_0}\right)W^F+\frac{\gamma'(q^*)q^*}{p_0}\left\{\begin{array}{l}-\frac{c}{1+\delta+\delta^2}+\frac{\delta+\delta^2}{1+\delta+\delta^2}\left(1-\frac{p_0}{q^*}\right)m(0)\\ +\frac{\delta+\delta^2}{1+\delta+\delta^2}\frac{p_0}{q^*}\left[-\frac{1}{1+\delta}c+\frac{\delta}{1+\delta}\left(\frac{q^*}{f^{(2)}(p_0)}m(f^{(2)}(p_0))+\left(1-\frac{q^*}{f^{(2)}(p_0)}\right)m(0)\right)\right]\end{array}\right\}\\
&= W^F-\frac{\gamma'(q^*)q^*}{p_0}\left\{\begin{array}{l}\frac{\delta+\delta^2}{1+\delta+\delta^2}\frac{p_0}{f(p_0)}\left[m(f(p_0))-\left(-\frac{c}{1+\delta}+\frac{\delta}{1+\delta}\left[\frac{f(p_0)}{f^{(2)}(p_0)}m(f^{(2)}(p_0))+\left(1-\frac{f(p_0)}{f^{(2)}(p_0)}\right)m(0)\right]\right)\right]\\ +\frac{\delta}{1+\delta+\delta^2}\left(\frac{p_0}{q^*}-\frac{p_0}{f(p_0)}\right)(c+m(0))\end{array}\right\}\\
&= W^F-\frac{\gamma(p^F)p^F}{p_0}\left\{\begin{array}{l}\frac{\delta+\delta^2}{1+\delta+\delta^2}\frac{p_0}{f(p_0)}\left[m(f(p_0))-\left(-\frac{c}{1+\delta}+\frac{\delta}{1+\delta}\left[\frac{f(p_0)}{f^{(2)}(p_0)}m(f^{(2)}(p_0))+\left(1-\frac{f(p_0)}{f^{(2)}(p_0)}\right)m(0)\right]\right)\right]\\ +\frac{\delta}{1+\delta+\delta^2}\left(\frac{p_0}{q^*}-\frac{p_0}{f(p_0)}\right)(c+m(0))\end{array}\right\}\\
&\geqslant W^F-\frac{\gamma(p^F)p^F}{p_0}\left\{\begin{array}{l}\frac{\delta+\delta^2}{1+\delta+\delta^2}\frac{p_0}{f(p_0)}\left[m(f(p_0))-\left(-\frac{c}{1+\delta}+\frac{\delta}{1+\delta}\left[\frac{f(p_0)}{f^{(2)}(p_0)}m(f^{(2)}(p_0))+\left(1-\frac{f(p_0)}{f^{(2)}(p_0)}\right)m(0)\right]\right)\right]\\ +\frac{\delta}{1+\delta+\delta^2}\left(\frac{p_0}{p^F}-\frac{p_0}{f(p_0)}\right)(c+m(0))\end{array}\right\}\\
&= W(\gamma)\geqslant W^N
\end{aligned}$$

Finally, we show that such perturbation in information disclosure weakly benefits the expert since we have the following

$$
\begin{aligned}
V(\gamma') &= \left[1-\gamma'(q^*)\frac{q^*}{p_0}\right]V^F+\gamma'(q^*)\frac{q^*}{p_0}\left\{\frac{\delta+\delta^2}{1+\delta+\delta^2}\left(1-\frac{p_0}{q^*}\right)m(0)+\frac{\delta^2}{1+\delta+\delta^2}\frac{p_0}{q^*}\left[\frac{q^*}{f^{(2)}(p_0)}m(f^{(2)}(p_0))+\left(1-\frac{q^*}{f^{(2)}(p_0)}\right)m(0)\right]\right\}\\
&= V^F+\gamma'(q^*)\frac{q^*}{p_0}\begin{Bmatrix}\frac{\delta^2}{1+\delta+\delta^2}\frac{p_0}{f(p_0)}\left[\frac{f(p_0)}{f^{(2)}(p_0)}m(f^{(2)}(p_0))+\left(1-\frac{f(p_0)}{f^{(2)}(p_0)}\right)m(0)-m(f(p_0))\right]-\frac{\delta}{1+\delta+\delta^2}\frac{p_0}{f(p_0)}m(f(p_0))\\ -\frac{\delta}{1+\delta+\delta^2}\left(\frac{p_0}{q^*}-\frac{p_0}{f(p_0)}\right)m(0)\end{Bmatrix}\\
&= V^F+\gamma(p^F)\frac{p^F}{p_0}\begin{Bmatrix}\frac{\delta^2}{1+\delta+\delta^2}\frac{p_0}{f(p_0)}\left[\frac{f(p_0)}{f^{(2)}(p_0)}m(f^{(2)}(p_0))+\left(1-\frac{f(p_0)}{f^{(2)}(p_0)}\right)m(0)-m(f(p_0))\right]-\frac{\delta}{1+\delta+\delta^2}\frac{p_0}{f(p_0)}m(f(p_0))\\ -\frac{\delta}{1+\delta+\delta^2}\left(\frac{p_0}{q^*}-\frac{p_0}{f(p_0)}\right)m(0)\end{Bmatrix}\\
&\geqslant V^F+\gamma(p^F)\frac{p^F}{p_0}\begin{Bmatrix}\frac{\delta^2}{1+\delta+\delta^2}\frac{p_0}{f(p_0)}\left[\frac{f(p_0)}{f^{(2)}(p_0)}m(f^{(2)}(p_0))+\left(1-\frac{f(p_0)}{f^{(2)}(p_0)}\right)m(0)-m(f(p_0))\right]-\frac{\delta}{1+\delta+\delta^2}\frac{p_0}{f(p_0)}m(f(p_0))\\ -\frac{\delta}{1+\delta+\delta^2}\left(\frac{p_0}{p^F}-\frac{p_0}{f(p_0)}\right)m(0)\end{Bmatrix}\\
&= V(\gamma)
\end{aligned}
$$

Any belief martingale γ with support $\{0, q^*, f(p_0)\}$ can be implemented by a convex combination of policy τ^I and full information disclosure policy. Since Bayesian plausibility implies that

$$
\begin{aligned}
&\gamma(0)0+\gamma(q^*)q^*+\gamma(f(p_0))f(p_0)=p_0\\
\Longrightarrow\ &\frac{\gamma(q^*)}{\frac{p_0}{q^*}}+\frac{\gamma(f(p_0))}{\frac{p_0}{f(p_0)}}=1,
\end{aligned}
$$

any policy with martingale γ can be implemented by τ^I with probability $\frac{\gamma(q^*)}{\frac{p_0}{q^*}}$, and full information disclosure policy with probability $\frac{\gamma(f(p_0))}{\frac{p_0}{f(p_0)}}$.

Denote V^H and W^H as the expert's and agent's payoff under disclosure policy τ^I. Once we have restricted to the set of belief martingales γ with support $\{0, q^*, f(p_0)\}$, the expert's payoff $V(\gamma)$ (the agent's payoff $W(\gamma)$) is a convex combination of V^H (W^H) and V^F (W^F). Then the optimal disclosure policy can be pinned down as a solution to the following optimization problem,

$$
\begin{aligned}
&\max_{\alpha\in[0,1]}\ \alpha V^H+(1-\alpha)V^F\\
&\quad s.t.\ \alpha W^H+(1-\alpha)W^F\geq W^N(\text{Plausibility})
\end{aligned}
\tag{16}
$$

Therefore if $V^F \geq V^H$, then it is optimal for the expert to fully disclose information. Otherwise, the expert prefers to maximize the weight on V^H or, in other words, maximize the weight of belief martingale on q^* subject to the constraint that the agent will be motivated to acquire information in the first period. In later case, the optimal policy, the optimal combination of τ^I and full information disclosure, is exactly the policy τ^U.

C.4. **Derivation of margin belief process.** If $dq_t = q_t(\rho q_t + c)dt$, then we have that

$$
\begin{aligned}
&\frac{dq_t}{q(\rho q_t + c)} = dt \Rightarrow \frac{dq_t}{q_t} - \frac{\rho}{\rho q_t + c} dq_t = cdt \\
&\Rightarrow d(\ln q_t - \ln(\rho q_t + c)) = dct \Rightarrow d \ln \frac{q_t}{\rho q_t + c} = dct \\
&\Rightarrow \ln \frac{q_t}{\rho q_t + c} - \ln \frac{q_0}{\rho q_0 + c} = ct \Rightarrow \frac{q_t}{\rho q_t + c} = e^{\ln \frac{q_0}{\rho q_0 + c} + ct} = \frac{q_0}{\rho q_0 + c} e^{ct} \\
&\Rightarrow q_t = \rho \frac{q_0}{\rho q_0 + c} e^{ct} q_t + c \frac{q_0}{\rho q_0 + c} e^{ct} \Rightarrow q_t = \frac{c \frac{q_0}{\rho q_0 + c} e^{ct}}{1 - \rho \frac{q_0}{\rho q_0 + c} e^{ct}} = \frac{c q_0 e^{ct}}{\rho q_0 (1 - e^{ct}) + c}
\end{aligned}
$$

C.5. **Proof of Lemma 2.** First, let me introduce some denotations. Denote a one-step-ahead process, which passes $p_{\hat{t}}$ at time $\hat{t}$, as $\mathcal{Q}^{\hat{t}}$ with a typical element at time t as $q_t^{\hat{t}}$.[13] Moreover, for any $\hat{t} \leq T(q^P, \mathcal{P})$, define a function $\gamma(t; \hat{t})$ by

$$\gamma(t; \hat{t}) := q_t^{\hat{t}} - p_t, \; \forall t \in [0, T(1; \mathcal{Q}^{\hat{t}})].$$

First, $\gamma(t; \hat{t})$ is strictly convex in t since

$$
\begin{aligned}
&\tfrac{d^2 \gamma(t; \hat{t})}{dt^2} = \tfrac{d \frac{dq_t^{\hat{t}}}{dt}}{dt} - \tfrac{d \frac{dp_t}{dt}}{dt} = \tfrac{d[q_t^{\hat{t}}(\rho q_t^{\hat{t}} + c)]}{dt} - \tfrac{d[\lambda p_t(1-p_t)]}{dt} = (c + 2\rho q_t^{\hat{t}}) \tfrac{dq_t^{\hat{t}}}{dt} + \lambda(2p_t - 1) \tfrac{dp_t}{dt} \\
&= q_t^{\hat{t}}(c + 2\rho q_t^{\hat{t}})(\rho q_t^{\hat{t}} + c) + \lambda^2 p_t (1 - p_t)(2p_t - 1) > 0.
\end{aligned}
$$

Second, I will show that $\frac{d\gamma(t;\hat{t})}{dt}\Big|_{t=\hat{t}} \leq 0$, where equality holds if and only if $\hat{t} = T(q^P, \mathcal{P})$. The argument follows from the following

$$\frac{d\gamma(t; \hat{t})}{dt}\bigg|_{t=\hat{t}} = \frac{dq_t^{\hat{t}}}{dt}\bigg|_{t=\hat{t}} - \frac{dp_t}{dt}\bigg|_{t=\hat{t}} = q_{\hat{t}}^{\hat{t}}(\rho q_{\hat{t}}^{\hat{t}} + c) - \lambda p_{\hat{t}}(1 - p_{\hat{t}}) = p_{\hat{t}}(\rho p_{\hat{t}} + c) - \lambda p_{\hat{t}}(1 - p_{\hat{t}}) \leq 0,$$

where the inequality holds if and only if $p_{\hat{t}} = q^P$. Third, since $\gamma(\hat{t}; \hat{t}) = 0$ and $\frac{d\gamma(t;\hat{t})}{dt}\Big|_{t=\hat{t}} \leq 0$, where equality holds if and only if $\hat{t} = T(q^P, \mathcal{P})$, the equation $\gamma(t; \hat{t}) = 0$ has exactly two solutions for any $\hat{t} < T(q^P, \mathcal{P})$ while it has a single solution if $\hat{t} = T(q^P, \mathcal{P})$. The first argument and the argument that $\theta(T(q^P, \mathcal{P}))$ naturally follows. Finally, for any $\hat{t} < \hat{t}' \leq T(q^P, \mathcal{P})$, one then have that $\frac{d\gamma(t;\hat{t}')}{dt}\Big|_{t=\hat{t}} < \frac{d\gamma(t;\hat{t}')}{dt}\Big|_{t=\hat{t}'} \leq 0$, which then implies that $\gamma(\hat{t}, \hat{t}') > 0$ or $q_{\hat{t}}^{\hat{t}'} > p_{\hat{t}} = q_{\hat{t}}^{\hat{t}}$. Therefore, we further have that $q_t^{\hat{t}'} > q_t^{\hat{t}}$ and $\gamma(t; \hat{t}') > \gamma(t; \hat{t})$ for any $t \leq T(1; \mathcal{Q}^{\hat{t}'})$. The argument that $\theta(\hat{t})$ is strictly decreasing in $\hat{t}$ then follows.

[13]Please note that the process $\mathcal{Q}(\hat{t})$ is a lower envelope of process $\mathcal{P}$ and $\mathcal{Q}^{\hat{t}}$;

C.6. **Proof of Corollary 1.** In the first step, we derive the derivative of hitting time $\theta(\hat{t})$. By the definition of hitting time function $\theta(\hat{t})$, the following equation holds

$$p_{\hat{t}} + \int_{\hat{t}}^{\theta(\hat{t})} \frac{dp_{\hat{t}}}{dt} dt = p_{\hat{t}} + \int_{\hat{t}}^{\theta(\hat{t})} \frac{dq_t^{\hat{t}}}{dt} dt \Leftrightarrow \int_{\hat{t}}^{\theta(\hat{t})} \left[\lambda p_t(1-p_t) - q_t^{\hat{t}}(\rho q_t^{\hat{t}} + c)\right] dt = 0 \tag{17}$$

Lemma 2 implies that one can take Implicit function theorem to derive $\theta'(\hat{t})$. Specifically, by taking derivative of both sides of Equation above with respect to $\hat{t}$, one then have that

$$\left[\lambda p_{\theta(\hat{t})}(1-p_{\theta(\hat{t})}) - q_{\theta(\hat{t})}^{\hat{t}}(\rho q_{\theta(\hat{t})}^{\hat{t}} + c)\right] \theta'(\hat{t}) - \left[\lambda p_{\hat{t}}(1-p_{\hat{t}}) - q_{\hat{t}}^{\hat{t}}(\rho q_{\hat{t}}^{\hat{t}} + c)\right] = 0.$$

Furthermore, one can then derive $\theta'(t)$ as

$$\theta'(\hat{t}) = \frac{\lambda p_{\hat{t}}(1-p_{\hat{t}}) - q_{\hat{t}}^{\hat{t}}(\rho q_{\hat{t}}^{\hat{t}} + c)}{\lambda p_{\theta(\hat{t})}(1-p_{\theta(\hat{t})}) - q_{\theta(\hat{t})}^{\hat{t}}(\rho q_{\theta(\hat{t})}^{\hat{t}} + c)} = \frac{p_{\hat{t}}\left[\lambda - c - (\lambda+\rho)p_{\hat{t}}\right]}{-p_{\theta(\hat{t})}\left[(\lambda+\rho)p_{\theta(\hat{t})} - (\lambda - c)\right]} = -\frac{p_{\hat{t}}(q^P - p_{\hat{t}})}{p_{\theta(\hat{t})}(p_{\theta(\hat{t})} - q^P)} \tag{18}$$

In the second step, we derive the partial derivative of term $\frac{p_0}{q_t(\hat{t})}$ with respect to $\hat{t}$, for any $t \in (\hat{t}, \theta(\hat{t}))$. At first, one can represent $\frac{p_0}{q_t(\hat{t})}$ as follows

$$\begin{aligned}
\frac{p_0}{q_t(\hat{t})} &= \frac{p_0}{\frac{c p_{\hat{t}} e^{c(t-\hat{t})}}{\rho p_{\hat{t}}(1-e^{c(t-\hat{t})})+c}} = \frac{p_0}{\frac{c\frac{p_0 e^{\lambda\hat{t}}}{p_0 e^{\lambda\hat{t}}+1-p_0} e^{c(t-\hat{t})}}{\rho\frac{p_0 e^{\lambda\hat{t}}}{p_0 e^{\lambda\hat{t}}+1-p_0}(1-e^{c(t-\hat{t})})+c}} \\
&= \frac{p_0\left[\rho\frac{p_0 e^{\lambda\hat{t}}}{p_0 e^{\lambda\hat{t}}+1-p_0}(1-e^{c(t-\hat{t})})+c\right]}{c\frac{p_0 e^{\lambda\hat{t}}}{p_0 e^{\lambda\hat{t}}+1-p_0} e^{c(t-\hat{t})}} = e^{-c(t-\hat{t})}\left[\frac{\rho+c}{c}p_0 + (1-p_0)e^{-\lambda\hat{t}}\right] - \frac{\rho p_0}{c}.
\end{aligned}$$

Therefore, one can further derive the partial derivative of term $\frac{p_0}{q_t(\hat{t})}$ with respect to $\hat{t}$ as

$$\begin{aligned}
&\frac{\partial\frac{p_0}{q_t(\hat{t})}}{\partial\hat{t}} = (\rho+c)p_0 e^{-ct}e^{-c\hat{t}} - (1-p_0)e^{-ct}(\lambda-c)e^{-(\lambda-c)\hat{t}} \\
&= p_0(\lambda-c)e^{-c(t-\hat{t})}\left[\frac{\rho+c}{\lambda-c} - \frac{1-p_0}{p_0}e^{-\lambda\hat{t}}\right] = p_0(\lambda-c)e^{-c(t-\hat{t})}\left[\frac{1-q^P}{q^P} - \frac{1-p_{\hat{t}}}{p_{\hat{t}}}\right]
\end{aligned} \tag{19}$$

Specifically, one can represent the expert's payoff upon the reporting policy, specified in Theorem 1 but with transition time $\hat{t}$, as

$$W(\hat{t}) = m(0)\int_0^{\theta(\hat{t})} e^{-\rho t}\frac{p_0}{q_t(\hat{t})}\frac{dq_t(\hat{t})}{q_t(\hat{t})dt}dt + e^{-\rho\theta(\hat{t})}\frac{p_0}{p_{\theta(\hat{t})}}m(p_{\theta(\hat{t})}) \tag{20}$$

$$= \rho m(0)\int_0^{\infty} e^{-\rho t}\left(1 - \frac{p_0}{\min\left\{q_t(\hat{t}), p_{\theta(\hat{t})}\right\}}\right)dt + e^{-\rho\theta(\hat{t})}\frac{p_0}{p_{\theta(\hat{t})}}m(p_{\theta(\hat{t})}) \tag{21}$$

By the definition of reporting process $\mathcal{Q}(\hat{t})$, the agent is indifferent between making decision at any moment during in time interval $[\hat{t}, \theta(\hat{t})]$, which therefore implies that

$$\underbrace{\int_{\hat{t}}^{\theta(\hat{t})} e^{-\rho t}\frac{p_0}{q_t(\hat{t})}\frac{dq_t(\hat{t})}{q_t(\hat{t})dt}dt + e^{-\rho\theta(\hat{t})}\frac{p_0}{p_{\theta(\hat{t})}}m(p_{\theta(\hat{t})})}_{\text{Discounted payoff if making decision at }\theta(\hat{t})} - \underbrace{c\int_{\hat{t}}^{\theta(\hat{t})} e^{-\rho t}\frac{p_0}{q_t(\hat{t})}dt}_{\text{Discounted research cost if funding until }\theta\hat{t})} = \underbrace{e^{-\rho\hat{t}}\frac{p_0}{p_{\hat{t}}}m(p_{\hat{t}})}_{\text{Discounted payoff if making decision at }\hat{t}}.$$

Please note that the agent only funds the expert if he has not received a message confirming the state being ω_1 and the probability of the event at time t is given by $\frac{p_0}{q_t(\hat{t})}$. Introduce the equation above into (20), one can derive another form of the expert's payoff as

$$W(\hat{t}) = \underbrace{m(0)\int_0^{\hat{t}} e^{-\rho t}\frac{p_0}{p_t}\frac{dp_t}{p_t dt}dt + e^{-\rho\hat{t}}\frac{p_0}{p_{\hat{t}}}m(p_{\hat{t}})}_{\text{Discounted payoff of early making decision at } \hat{t}} + \underbrace{c\int_{\hat{t}}^{\theta(\hat{t})} e^{-\rho t}\frac{p_0}{q_t(\hat{t})}dt}_{\text{Research grant occupation}} \tag{22}$$

Thirdly, by Equation (22), let me take a derivative of $W(\hat{t})$ with respect to $\hat{t}$ as

$$\begin{aligned}
W'(\hat{t}) &= e^{-\rho\hat{t}}\frac{p_0}{(p_{\hat{t}})^2}\left.\frac{dp_t}{dt}\right|_{t=\hat{t}} - \rho p_0 e^{-\rho\hat{t}} + c\left[e^{-\rho\theta(\hat{t})}\frac{p_0}{q_{\theta(\hat{t})}(\hat{t})}\theta'(\hat{t}) - e^{-\rho\hat{t}}\frac{p_0}{q_{\hat{t}}(\hat{t})}\right] + c\int_{\hat{t}}^{\theta(\hat{t})} e^{-\rho t}\frac{\partial\frac{p_0}{q_t(\hat{t})}}{\partial\hat{t}}dt \\
&= e^{-\rho\hat{t}}\frac{p_0}{(p_{\hat{t}})^2}\lambda p_{\hat{t}}(1-p_{\hat{t}}) - \rho p_0 e^{-\rho\hat{t}} + c\left[-e^{-\rho\theta(\hat{t})}\frac{p_0}{p_{\theta(\hat{t})}}\frac{p_{\hat{t}}(q^P-p_{\hat{t}})}{p_{\theta(\hat{t})}(p_{\theta(\hat{t})}-q^P)} - e^{-\rho\hat{t}}\frac{p_0}{p_{\hat{t}}}\right] + c\int_{\hat{t}}^{\theta(\hat{t})} e^{-\rho t}p_0(\lambda-c)e^{-c(t-\hat{t})}\left[\frac{1-q^P}{q^P} - \frac{1-p_{\hat{t}}}{p_{\hat{t}}}\right]dt \\
&= e^{-\rho\hat{t}}\frac{p_0}{p_{\hat{t}}}\lambda(1-p_{\hat{t}}) - \rho p_0 e^{-\rho\hat{t}} - ce^{-\rho\hat{t}}\frac{p_0}{p_{\hat{t}}} - ce^{-\rho\theta(\hat{t})}\frac{p_0}{p_{\theta(\hat{t})}}\frac{p_{\hat{t}}(q^P-p_{\hat{t}})}{p_{\theta(\hat{t})}(p_{\theta(\hat{t})}-q^P)} + cp_0(\lambda-c)e^{c\hat{t}}\left(\frac{1-q^P}{q^P} - \frac{1-p_{\hat{t}}}{p_{\hat{t}}}\right)\int_{\hat{t}}^{\theta(\hat{t})} e^{-(\rho+c)t}dt \\
&= e^{-\rho\hat{t}}\frac{p_0}{p_{\hat{t}}}\left[\lambda(1-p_{\hat{t}}) - \rho p_{\hat{t}} - c\right] - e^{-\rho\theta(\hat{t})}\frac{p_0}{p_{\hat{t}}}c\left(\frac{p_{\hat{t}}}{p_{\theta(\hat{t})}}\right)^2\frac{q^P-p_{\hat{t}}}{p_{\theta(\hat{t})}-q^P} + \frac{p_0}{p_{\hat{t}}}ce^{c\hat{t}}\left[(\lambda-c)\frac{1-q^P}{q^P}p_{\hat{t}} - (\lambda-c)(1-p_{\hat{t}})\right]\frac{e^{-(\rho+c)\hat{t}} - e^{-(\rho+c)\theta(\hat{t})}}{\rho+c} \\
&= e^{-\rho\hat{t}}\frac{p_0}{p_{\hat{t}}}\left[(\lambda-c) - (\lambda+\rho)p_{\hat{t}}\right] - e^{-\rho\hat{t}}\frac{p_0}{p_{\hat{t}}}(\lambda+\rho)(q^P-p_{\hat{t}})\left(\frac{p_{\hat{t}}}{p_{\theta(\hat{t})}}\right)^2\frac{ce^{-\rho(\theta(\hat{t})-\hat{t})}}{(\lambda+\rho)(p_{\theta(\hat{t})}-q^P)} - e^{-\rho\hat{t}}\frac{p_0}{p_{\hat{t}}}\left[(\lambda-c) - (\lambda+\rho)p_{\hat{t}}\right]\frac{c}{\rho+c}(1-e^{-(\rho+c)(\theta(\hat{t})-\hat{t})}) \\
&= e^{-\rho\hat{t}}\frac{p_0}{p_{\hat{t}}}(\lambda+\rho)(q^P-p_{\hat{t}})\left\{\frac{\rho}{\rho+c} - ce^{-(\rho+c)(\theta(\hat{t})-\hat{t})}\left[\left(\frac{p_{\hat{t}}}{p_{\theta(\hat{t})}}\right)^2\frac{e^{c(\theta(\hat{t})-\hat{t})}}{(\rho+\lambda)(p_{\theta(\hat{t})}-q^P)} - \frac{1}{\rho+c}\right]\right\}
\end{aligned}$$

The second equality above follows from the introduction of Equation (18) and (19).

Finally, I will characterize t^* and show its uniqueness. At first, I will show that the term $\left(\frac{p_{\hat{t}}}{p_{\theta(\hat{t})}}\right)^2\frac{e^{c(\theta(\hat{t})-\hat{t})}}{(\rho+\lambda)(p_{\theta(\hat{t})}-q^P)}$ is increasing in $\hat{t}$. One can rewrite the term as

$$\left(\frac{p_{\hat{t}}}{p_{\theta(\hat{t})}}\right)^2\frac{e^{c(\theta(\hat{t})-\hat{t})}}{(\rho+\lambda)(p_{\theta(\hat{t})}-q^P)} = \frac{p_{\hat{t}}}{p_{\theta(\hat{t})}}\frac{p_{\hat{t}}}{\frac{cp_{\hat{t}}e^{c(\theta(\hat{t})-\hat{t})}}{\rho p_{\hat{t}}[1-e^{c(\theta(\hat{t})-\hat{t})}]+c}}\frac{e^{c(\theta(\hat{t})-\hat{t})}}{(\rho+\lambda)(p_{\theta(\hat{t})}-q^P)} = \frac{p_{\hat{t}}}{p_{\theta(\hat{t})}}\frac{\rho p_{\hat{t}}[1-e^{c(\theta(\hat{t})-\hat{t})}]+c}{c(\rho+\lambda)(p_{\theta(\hat{t})}-q^P)}.$$

Since $\theta(\hat{t})$ is decreasing in $\hat{t}$ from Lemma 2, one then have that all the terms $\frac{p_{\hat{t}}}{p_{\theta(\hat{t})}}$, $\rho p_{\hat{t}}[1-e^{c(\theta(\hat{t})-\hat{t})}]+c$ and $\frac{1}{c(\rho+\lambda)(p_{\theta(\hat{t})}-q^P)}$ are increasing in $\hat{t}$, so is $\left(\frac{p_{\hat{t}}}{p_{\theta(\hat{t})}}\right)^2\frac{e^{c(\theta(\hat{t})-\hat{t})}}{(\rho+\lambda)(p_{\theta(\hat{t})}-q^P)}$. Secondly, if the following holds,

$$\left(\frac{p_{\hat{t}}}{p_{\theta(\hat{t})}}\right)^2\frac{e^{c(\theta(\hat{t})-\hat{t})}}{(\rho+\lambda)(p_{\theta(\hat{t})}-q^P)} \le \frac{1}{\rho+c},$$

then $W'(\hat{t}) > 0$. Moreover since

$$\lim_{\hat{t}\to T(q^P,\mathcal{P})}\left(\frac{p_{\hat{t}}}{p_{\theta(\hat{t})}}\right)^2\frac{e^{c(\theta(\hat{t})-\hat{t})}}{(\rho+\lambda)(p_{\theta(\hat{t})}-q^P)} = \infty,$$

and the continuity of $W(\cdot)$, there exists some $\hat{t}_1$ such that

$$\left(\frac{p_{\hat{t}_1}}{p_{\theta(\hat{t}_1)}}\right)^2\frac{e^{c(\theta(\hat{t}_1)-\hat{t}_1)}}{(\rho+\lambda)(p_{\theta(\hat{t}_1)}-q^P)} > \frac{1}{\rho+c}.$$

We therefore focus on the time interval $(\hat{t}_1, T(q^P, \mathcal{P})]$. Otherwise, we focus on the time interval $[0, T(q^P, \mathcal{P})]$. Finally, in this time interval, the term $ce^{-(\rho+c)(\theta(\hat{t})-\hat{t})}\left[\left(\frac{p_{\hat{t}}}{p_{\theta(\hat{t})}}\right)^2 \frac{e^{c(\theta(\hat{t})-\hat{t})}}{(\rho+\lambda)(p_{\theta(\hat{t})}-q^P)} - \frac{1}{\rho+c}\right]$ is always increasing in $\hat{t}$. Accordingly, there exists a unique t^* to maximize $W(\cdot)$. Moreover, if $W'(0) \leq 0$, then $t^* = 0$; Otherwise, t^* is a solution to the equation

$$\frac{\rho}{\rho+c} = ce^{-(\rho+c)(\theta(\hat{t})-\hat{t})}\left[\left(\frac{p_{\hat{t}}}{p_{\theta(\hat{t})}}\right)^2 \frac{e^{c(\theta(\hat{t})-\hat{t})}}{(\rho+\lambda)(p_{\theta(\hat{t})}-q^P)} - \frac{1}{\rho+c}\right].$$

C.7. **Proof of Corollary 2.** One can further simplify the aggregate information value $s(t;t^*)$ as

$$s(t;t^*) = \frac{p_0}{q_t(t^*)} - \frac{p_0}{p_t}.$$

For any $t \in [0, t^*] \cup [\theta(t^*), \infty)$, since $q_t(t^*) = p_t$, then $s(t, t^*) \equiv 0$. For any $t \in (t^*, \theta(t^*))$, $q_t(t^*) = q_t^{t^*}$. Then let me take first order derivative of $s(t;t^*)$ with respect to t as follows

$$\begin{aligned}\frac{ds(t;t^*)}{dt} &= -\frac{p_0}{(q_t^{t^*})^2}\frac{dq_t^{t^*}}{dt} + \frac{p_0}{p_t^2}\frac{dp_t}{dt} = \frac{p_0}{p_t^2}\lambda p_t(1-p_t) - \frac{p_0}{(q_t^{t^*})^2} q_t^{t^*}(\rho q_t^{t^*} + c) = \frac{\lambda}{p_t} - \frac{c}{q_t^{t^*}} - (\lambda+\rho) \\ &= \frac{\lambda}{\frac{p_{t^*}e^{\lambda(t-t^*)}}{p_{t^*}e^{\lambda(t-t^*)}+(1-p_{t^*})}} - \frac{c}{\frac{cp_{t^*}e^{c(t-t^*)}}{\rho p_{t^*}\left[1-e^{c(t-t^*)}\right]+c}} - (\lambda+\rho) = \frac{\rho p_{t^*}+c}{p_{t^*}}e^{-\lambda(t-t^*)}\left[\frac{\lambda(1-p_{t^*})}{\rho p_{t^*}+c} - e^{(\lambda-c)(t-t^*)}\right]\end{aligned}$$

As a result, the turning point t^S is given by

$$\tilde{t}_1 = t^* + \frac{1}{\lambda - c}\ln\frac{\lambda(1-p_{t^*})}{\rho p_{t^*}+c}.$$

Moreover, for any $t < t^S$, $s(t;t^*)$ is increasing in t and vice versa. Finally, As we have shown in Lemma 2, $\theta(\hat{t}) \geq T(q^P, \mathcal{P})$ with equality holding if and only if $\hat{t} = T(q^P, \mathcal{P})$. Therefore, we have that $p_{\theta(t^*)} = q_{\theta(t^*)}(t^*) = q_{\theta(t^*)}^{t^*} \geq q^P$. One then have that

$$\left.\frac{ds(t;t^*)}{dt}\right|_{t=\theta(t^*)} = \frac{\lambda}{p_{\theta(t^*)}} - \frac{c}{q_{\theta(t^*)}^{t^*}} - (\lambda+\rho) = \frac{\lambda - c}{p_{\theta(t^*)}} - (\lambda+\rho) \leq \frac{\lambda-c}{q^P} - (\lambda+\rho) = 0,$$

which therefore implies that $\theta(t^*) \geq t^S$, with equality holding if and only if $t^* = T(q^P, \mathcal{P})$.

C.8. **Proof of Corollary 3.** One can easily verify the reporting strategy of truthfully report with delay $D(t;t^*)$ exactly induces a belief martingale which either jumps to belief 0 or moves along the reporting process $\mathcal{Q}(t^*)$ with transition time t^*. Besides, upon not receiving a single message confirming the state being ω_1, the agent will not cut funding until the hitting time $\theta(t^*)$. Therefore, the expert, under this reporting strategy, derives what she can derive under optimal disclosure policy.

Moreover, for any $t \in (t^*, \theta(t^*))$, let me take derivative of both sides of Equation (7) with respect to t as

$$\frac{dp_{t-D(t;t^*)}}{dt} = \frac{dq_t(t^*)}{dt} \Rightarrow \frac{dp_{t'}}{dt'}\Big|_{t'=t-D(t;t^*)}(1 - \frac{dD(t;t^*)}{dt}) = \frac{dq_t^{t^*}}{dt}$$
$$\Rightarrow \frac{dD(t;t^*)}{dt} = 1 - \frac{\frac{dq_t^{t^*}}{dt}}{\frac{dp_{t'}}{dt'}\Big|_{t'=t-D(t;t^*)}} = 1 - \frac{q_t^{t^*}(\rho q_t^{t^*}+c)}{\lambda p_{t-D(t;t^*)}(1-p_{t-D(t;t^*)})} = 1 - \frac{q_t^{t^*}(\rho q_t^{t^*}+c)}{\lambda q_t^{t^*}(1-q_t^{t^*})} = \frac{\lambda+\rho}{\lambda(1-q_t^{t^*})}(q^P - q_t^{t^*})$$

As we have shown that $\theta(t^*) \geq T(q^P, \mathcal{P})$ in Lemma 2, one then have that $t^* \leq T(q^P, \mathcal{Q}(t^*)) \leq \theta(t^*)$. Therefore, the turning point t^D is given by $T(q^P, \mathcal{Q}(t^*))$. Besides, $\frac{dD(t;t^*)}{dt} > 0$ for any $t \in (t^*, t^D)$ while $\frac{dD(t;t^*)}{dt} < 0$ for any $t \in (t^D, \theta(t^*))$.

C.9. **Proof of Theorem 1.** Our proof covers the case where $p_0 \geq \frac{1}{2}$. However, it can be easily generated to the case where $p_0 \in [\underline{p}_0, \frac{1}{2})$ by replacing time 0 with time $\underline{t}(p_0)$.

C.9.1. *Stopping belief distribution and its Bayesian plausibility.* In the *first* step, we argue that with probability 1, the agent will finally stop research funding and make decision, given any dynamic disclosure policy.

We prove Lemma 3 by proving Lemma 9, 10 and 11.

Lemma 9. *For any dynamic disclosure policy* (π, M)*, the agent will stop research funding and make decision with probability* 1.

Proof. Suppose with some probability $\alpha > 0$, the agent chooses to keep funding forever. Let me define

$$\alpha_t := Pr(\text{the agent continues funding research at time } t).$$

By definition, α_t is monotonically decreasing and therefore converges to α. As a result, for any ϵ, there exists some $T > 0$ such that $\alpha_t - \alpha < \epsilon$, for any $t > T$. Therefore at any time $t > T$, the agent's continuation payoff is bounded above by

$$-\frac{\alpha}{\alpha_t}\int_0^\infty e^{-\rho t}cdt + \frac{\alpha_t - \alpha}{\alpha_t}max_p m(p) \leq -\frac{\alpha}{\alpha+\epsilon}\int_0^\infty e^{-\rho t}cdt + \frac{\epsilon}{\alpha}$$

By making ϵ extremely small, the terms above is below zero, which leads to contradiction, as the agent can at least secure payoff $m(p_t) \geq \frac{1}{2}$ if stopping research funding and making decision right away. □

For any dynamic disclosure policy (π, M), denote $\tau_t^{(\pi,M)}(p)$ as the probability that the agent chooses to stop funding and make decision exactly at time t and if he believes that the state ω_2 is with probability p. Then, denote $\tau^{(\pi,M)}(p) := \int_{t=0}^\infty \tau_t^{(\pi,\mathcal{M})}dt$ as the aggregate (across time) probability that the agent chooses to stop funding and make decision given his belief being p. Lemma 9 then implies that $\tau^{(\pi,M)}$ is a probability distribution over $\Delta\{\omega_1,\omega_2\}$. We therefore name such $\tau^{(\pi,M)}$ as **stopping belief distribution** induced by dynamic disclosure policy (π, M). Moreover, if the agent chooses to stop research funding and make decision, when he believes that the state is ω_2 with probability p, the his belief will stay at p since then. This is due to the fact that there will be no information acquisition since then and belief will no longer be updated. In this sense, it is similar to static Bayesian persuasion (c.f. Kamenica and Gentzkow [2011]) which only involves one step information disclosure. As a result, we have **Bayesian plausibility** constraint,

$$\int_0^1 p\tau^{(\pi,M)}(p)dp = p_0,\ \forall(\pi, M).$$

Finally, whenever there is no confusion, let me simplify $\tau^{(\pi,M)}$ as τ^{π}.

Example 1. *In the three period model introduced in section 3, let us consider the uncertainty increase dynamic disclosure policy. Under this dynamic disclosure policy, the belief martingale, starting from prior p_0, jumps to belief 0 with probability $\frac{p_0}{\underline{p}}$, to belief with probability $f(p_0)$ with probability $\frac{p_0}{\underline{p}}\frac{\underline{p}-q^*}{f(p_0)-q^*}$ and to belief q^* with probability $\frac{p_0}{\underline{p}}\frac{f(p_0)-\underline{p}}{f(p_0)-q^*}$. Only if the first two beliefs are reached does the agent choose to stop funding and make decision immediately, which implies that $\tau_1(0) = \frac{p_0}{\underline{p}}$ and $\tau_1(f(p_0)) = \frac{p_0}{\underline{p}}\frac{\underline{p}-q^*}{f(p_0)-q^*}$; Otherwise, the agent chooses to fund for another period and the belief martingale, resuming at belief q^*, jumps to belief 0 with probability $\frac{f^{(2)}(p_0)-q^*}{q^*}$ and to belief $f^{(2)}(p_0)$ with probability $\frac{q^*}{f^{(2)}(p_0)}$. Then, $\tau_2(0) = \frac{p_0}{\underline{p}}\frac{f(p_0)-\underline{p}}{f(p_0)-q^*}\frac{f^{(2)}(p_0)-q^*}{q^*}$ and $\tau_2(f^{(2)}(p_0)) = \frac{p_0}{\underline{p}}\frac{f(p_0)-\underline{p}}{f(p_0)-q^*}\frac{q^*}{f^{(2)}(p_0)}$. One can then derive that*

$$\begin{cases} \tau(0) = \tau_1(0) + \tau_2(0) = \frac{p_0}{\underline{p}} + \frac{p_0}{\underline{p}}\frac{f(p_0)-\underline{p}}{f(p_0)-q^*}\frac{f^{(2)}(p_0)-q^*}{q^*} \\ \tau(f(p_0)) = \tau_1(f(p_0)) = \frac{p_0}{\underline{p}}\frac{\underline{p}-q^*}{f(p_0)-q^*} \\ \tau(f^{(2)}(p_0)) = \tau_2(f^{(2)}(p_0)) = \frac{p_0}{\underline{p}}\frac{f(p_0)-\underline{p}}{f(p_0)-q^*}\frac{q^*}{f^{(2)}(p_0)} \\ \tau(p) = 0,\ \textit{otherwise} \end{cases}.$$

One can then show that

$$\begin{cases} \int_0^1 \tau(p)dp = \tau(0) + \tau(f(p_0)) + \tau(f^{(2)}(p_0)) = 1 \\ \int_0^1 \tau(p)pdp = \tau(0)0 + \tau(f(p_0))f(p_0) + \tau(f^{(2)}(p_0))f^{(2)}(p_0) = p_0 \end{cases}.$$

C.9.2. *Responsiveness.* Let me focus on a subclass of dynamic disclosure policy under which the expert has reported all her privately observed research progress whenever the agent chooses to stop research funding. We define such property as follows

Definition 3. *A dynamic disclosure policy is* ***responsive*** *if the agent chooses to stop research funding and make decision right away only when his belief is consistent with the expert's.*

Lemma 10. *It is without loss of optimality to focus on the set of* ***responsive*** *dynamic disclosure policy.*

Proof. Suppose that the expert's belief at some time is given by q. Moreover upon receiving some message, the agent forms a posterior $p < q$ and chooses to stop research funding and make decision right away. Then he believes that his continuation payoff will be weakly lower than $m(p)$ if he continues to fund research and put off decision making. Then let me consider a perturbed version of dynamic disclosure policy,

(1) The expert replaces the message, which leads that the agent quit research funding and making decision right away while forming posterior p, by revealing whether the state is of ω_1 (and therefore forms a posterior 0) or the state is ω_2 with probability q; Furthermore, if the message which means that the state is ω_2 is realized, the expert commits to conceal all the information in the future if she gets funded.
(2) keeping all else the same.

If the agent receives the message which means that the state is exactly ω_1, then he will make decision, by choosing policy a_1, right away. Otherwise if the agent receives the message which means that the state is ω_2 with probability $q \geq p_0 \geq \frac{1}{2}$, then he will also make decision, by choosing policy a_2, right away since he knows that there will be no information disclosure in the future even if he chooses to fund expert to conduct research in the future. As a result, his expected payoff is given by $\frac{p}{q}m(q) + (1 - \frac{p}{q})m(0) \geq m(p)$. Put it in other words, the perturbed version of dynamic disclosure policy, compared to the original one, improves social welfare through increasing the accuracy of agent's choice between candidate policy without delay. Moreover, such perturbation also induces the agent to be more willing to fund research and delay decision making in previous periods as it weakly increases the continuation payoff for the agent. Put it in other words, such perturbation relaxes the incentive compatibility constraints in previous periods. Let me define such dynamic disclosure policy as responsive dynamic disclosure policy. □

The definition of **responsive** dynamic disclosure policy (π, M) implies that the support of its stopping belief distribution τ^π should be either 0 or weakly above p_0. Or mathematically we have that

$$\text{supp}(\tau^\pi) \subseteq \{0\} \cup [p_0, 1]\,.$$

Our *next* step argues that one can even further restrict its support to $\{0\} \cup [q^P, q^R]$ when looking for optimal dynamic disclosure policy.

Lemma 11. *It is without loss of optimality to further focus on a subset of responsive dynamic disclosure policies with stopping belief distribution τ, whose support lies in $\{0\} \cup [q^P, q^R]$.*

Proof. Let us consider a responsive dynamic disclosure policy (π, M) with $\int_{p\in(q^R,1]} \tau^\pi(p)dp > 0$, then the probability that the agent keeps research funding at $T(q^R, \mathcal{P})$ is strictly positive. Let us consider a perturbation (π', M') of the original dynamic disclosure policy,

(1) For any history node such that the the agent keeps funding research prior to and at $T(q^R, \mathcal{P})$, the expert replaces the message sent at the end of $T(q^R, \mathcal{P})$ by revealing all what she has previously learnt at $T(q^R, \mathcal{P})$, inducing a belief martingale, resuming from some belief $p \in (0, q^R)$, jumping either to belief 0 or belief q^R instantly. Furthermore, she commits to disclose all what she has newly learnt in time if she is funded to conduct information acquisition further;
(2) Keeping all else the same.

Contingent on the event that the agent keeps funding research prior to and at $T(q^R, \mathcal{P})$, suppose that the agent blindly follows the expert's recommendations since then. Lemma 1 implies that the expert will recommends that the agent stop funding and make decision right away. We then argue that this recommendation is incentive compatible (when the agent behaves strategically), since the preference of the agent is perfectly aligned with the one of the expert except that the former has to bear the research cost. Finally, the agent's continuation payoff upon history nodes where the research has been funded prior to and at $T(q^R, \mathcal{P})$ reaches maximum as the expert conducts full information disclosure since then. Therefore, the agent still keeps funding research along these history nodes under the perturbed responsive dynamic disclosure policy (π', M'). Accordingly, such perturbation strictly improves the expert's payoff and restricts the support of π' to $\{0\} \cup [p_0, q^R]$. Similarly, one can show that $\int_{p\in[p_0,p^P)} \tau(p)dp = 0$ at optimum. □

C.9.3. *Directness and revelation principle.* Let us consider a a sub-class responsive dynamic disclosure policies (π, M) such that the message space $M := \{F, S_1, S_2\}$, where F means keeping research funding and S_1 (S_2) means stoping funding and choosing policy a_1 (a_2). We name such property as **directness**.

Definition 4. *A responsive dynamic disclosure policy is* ***direct*** *if at any time t, there exists at most one message upon which the agent chooses to continue to fund research.*

In the *forth* step, we show that revelation principle holds in our context[14], i.e., restriction on the class of direct and responsive dynamic disclosure policies is without loss of generality.

Lemma 12. *It is without loss of generality to focus on the class of direct and responsive dynamic disclosure policies.*

Proof. Given any dynamic disclosure policy (π, M), let me write $\lambda^{(\pi,M)}(t)$ as the *accumulating probability* that the agent believes that the state is ω_1 up to date t. Whenever there is no confusion, let me simplify it at $\lambda^{\pi}(t)$. Please that $\lambda(t)$ is monotonically increasing and converges to $\tau(0)$. Let me name it as **weight accumulation speed** at posterior 0. At first, different responsive dynamic disclosure policies end up with the same payoff to both the expert and agent if they have the same weight accumulation speed λ and stopping posterior distribution τ. Specifically, given (λ, τ), the expert's payoff is given by

$$V(\lambda, \tau) = \rho \int_0^{\infty} e^{-\rho t}\lambda(t)m(0)dt + \int_{p\in[q^P,q^R]} e^{-\rho T(p,P)}\tau(p)m(p)dp,$$

and the agent's payoff is given by

$$U(\lambda, \tau) = V(\lambda, \tau) - c\int_0^{\infty} e^{-\rho t}\left[1 - \lambda(t) - \int_{p\in[q^P,p_t]} \tau(p)dp\right] dt.$$

Secondly, for any responsive dynamic disclosure policy π with weight accumulation speed and stopping posterior distribution $(\lambda^{\pi}, \tau^{\pi})$, consider the following direct disclosure policy. At $t < T(q^P, \mathcal{P})$, the expert sends two messages, one is choosing policy a_1, and the other is keep funding research. The accumulating probability that the first message is sent is given by $\lambda^{\pi}(t)$ while the instant probability that the second message is sent is given by $1-\lambda^{\pi}(t)$. Upon receiving the second message, the agent forms a posterior $\frac{p_0}{1-\lambda^{\pi}(t)}$. At $t \in [T(q^P, \mathcal{P}), T(q^R, \mathcal{P})]$, the expert sends three messages, the first one is choosing policy a^1, the second one is keeping funding research and the third one is choosing policy a_2. The accumulating probability that the first message is sent is given by $\lambda^{\pi}(t)$; The instant probability that the second message is sent is given by $1 - \lambda^{\pi}(t) - \int_{p\in[q^P,p_t]} \tau^{\pi}(p)dp$; Finally the accumulating probability that the third message is sent is given by $\int_{p\in[q^P,p_t]} \tau^{\pi}(p)dp$. Whenever it is with strictly positive probability that the second message is sent, the

[14]Revelation principle in dynamic information design was introduced by Myerson [1986] and it has been recently discussed by Doval and Ely [2018], Sugaya and Wolitzky [2019] and Makris and Renou [2018]. We show that revelation principle still holds in our setting, which is a stopping game with both continuous time and state at each period evolving based on actions taken by the agent.

agent forms a posterior

$$\hat{p}_t(\lambda,\tau) = \frac{p_0 - \int_{p\in[q^P,p_t]} \tau^\pi(p)pdp}{1-\lambda^\pi(t) - \int_{p\in[q^P,p_t]} \tau^\pi(p)dp} \tag{23}$$

Please note that if $t \in [0, T(q^P, \mathcal{P}))$, then $p_t < q^P$ and the interval $[q^P, p_t]$ is an empty set. Further it implies that

$$\int_{p\in[q^P,p_t]} \tau^\pi(p)pdp = \int_{p\in[q^P,p_t]} \tau^\pi(p)dp = 0.$$

Therefore Equation (23) is a unified formulation for the posterior, at which the agent continues to fund research. Besides the continuation payoff $V_t(\lambda,\tau)$ to the expert, if the agent receives the second message at time t, is given by

$$V_t(\lambda,\tau) = \frac{\rho\int_t^\infty e^{-\rho(t'-t)}[\lambda(t')-\lambda(t)]m(0)dt' + \int_{p\in(p_t,q^R]} e^{-\rho[T(p,P)-t]}\tau(p)m(p)dp}{1-\lambda(t) - \int_{p\in[q^P,p_t]} \tau(p)dp} \tag{24}$$

while the continuation payoff $U_t(\lambda,\tau)$ to the agent is given by

$$U_t(\lambda,\tau) = V_t(\lambda,\tau) - \frac{\int_t^\infty e^{-\rho(t'-t)}[1-\lambda(t') - \int_{p\in[q^P,p_{t'}]} \tau(p)dp]dt'}{1-\lambda(t) - \int_{p\in[q^P,p_t]} \tau(p)dp}c \tag{25}$$

We then argue that the direct dynamic disclosure policy is incentive compatible. By construction, the direct disclosure policy π^D has weight accumulating speed at posterior 0 as λ^π and stopping posterior distribution τ^π. Based on our previous argument, the direct dynamic disclosure policy π^D has the same payoff as the original dynamic disclosure policy π. Moreover at any t, q^t is a garbling of all posteriors implied by messages, upon which the agent continues to fund research at time t in the original disclosure policy. As a result, the agent will still continue to fund research at the garbling of beliefs implied by these messages, which therefore implies the incentive compatibility of the direct disclosure policy. □

C.9.4. *Fastest weight accumulation speed* λ. In the *fifth* step, we are going to study the fastest weight accumulation speed, given some fixed stopping belief distribution τ^π, among all direct and responsive dynamic disclosure policies (π, M). Basically, for any stopping belief distribution τ such that Bayesian plausibility constraint holds, we construct an induced belief martingale and show that the underlying weight accumulating speed λ reaches its maximum. Furthermore, construction of the induced belief martingale is divided into two parts, before and after time $T(q^P;\mathcal{P})$, and conducted in flashback.

At first, let me assume that $q_{T(q^R,\mathcal{P})}(0) \leq q^R$. Lemma 2 then implies that for any belief $p \in [q^P, q^R]$, there exists some transition time $\varphi(p)$ such that

$$\theta(\varphi(p)) = T(p;\mathcal{P}).$$

To simplify notation, let me define belief $\eta(p) := q_{T(q^P;\mathcal{P})}(\varphi(p))$. Mathematically, one can solve for $\eta(p)$ as

$$\frac{c\eta(p)e^{c[T(p,\mathcal{P})-T(q^P,\mathcal{P})]}}{\rho\eta(p)(1-e^{c[T(p,\mathcal{P})-T(q^P,\mathcal{P})]})+c} = p \implies \eta(p) = \frac{cp}{e^{c[T(p,\mathcal{P})-T(q^P,\mathcal{P})]}(c+\rho p)-\rho p}.$$

One can easily verify that $\eta(q^P) = q^P$. Besides, Lemma 2 implies that $\eta(p)$ is decreasing in p.

Then we recover the second part (i.e. after time $T(q^P;\mathcal{P})$) of the belief martingale which implements the stopping belief distribution τ. Suppose at time $T(q^P;\mathcal{P})$, the belief martingale resumes at belief $\eta(p)$ with some probability $\gamma(p)$, which will be specified later, for any $p \in [q^P, q^R]$; With the rest probability $1 - \int_{q^P}^{q^R} \gamma(p)dp$, it resumes from belief 0. If the belief martingale ends up resuming at belief $\eta(p)$, then it either jumps to belief 0 or follows the reporting process $\mathcal{Q}(\varphi(p))$. Besides, whenever the belief martingale jumps to belief 0, it then stays there forever. Under this belief martingale, the agent, contingent on the fact that belief martingale resumes at $\eta(p)$ at time $T(q^P;\mathcal{P})$, will stop funding and make decision at time $T(p;\mathcal{P})$ if the belief martingale has not jumped to belief 0 until then[15]. This therefore implements a stopping belief distribution $\tau(\cdot;\gamma)$ where

$$\tau(p;\gamma) = \begin{cases} \frac{\gamma(p)\eta(p)}{p}, & \text{if } p \in [q^P, q^R] \\ 1 - \int_{p\in[q^P,q^R]} \frac{\gamma(p)\eta(p)}{p}dp, & \text{if } p = 0 \\ 0, & \text{otherwise} \end{cases}$$

Consequently, in order to implement a stopping belief martingale τ, one should then set

$$\gamma(p;\tau) = \frac{\tau(p)p}{\eta(p)}, \ \forall p \in [q^P, q^R].$$

Next, We recover the first part (i.e. before time $T(q^P;\mathcal{P})$) of the belief martingale. Let me define $\hat{\eta}(\tau)$ as the weighted average of $\eta(p)$ or mathematically,

$$\hat{\eta}(\tau) = \int_{p\in[q^P,q^R]} \frac{\gamma(p;\tau)}{\int_{p\in[q^P,q^R]}\gamma(p;\tau)dp}\eta(p)dp.$$

[15]Whenever the belief martingale jumps to belief 0, the agent will stop funding and make decision immediately;

Then let me define the transition time $\hat{\varphi}(\tau)$ with which the reporting process $\mathcal{Q}(\hat{\varphi}(\tau))$ transits to belief $\hat{\eta}(\tau)$ at time $T(q^P)$ or mathematically,

$$q_{T(q^P;\mathcal{P})}(\hat{\varphi}(\tau)) = \hat{\eta}(\tau).$$

Then the first part of belief martingale is specified as, starting from belief p_0, it either jumps to belief 0 (and then stays therefore forever) or transits based on reporting process $\mathcal{Q}(\hat{\varphi}(\tau))$.

Finally, let me combine the first and the second part of the belief martingale. If the belief martingale has transited to belief $\hat{\eta}(\tau)$ at time $T(q^P;\mathcal{P})$, then it jumps to belief $\eta(p)$ with probability $\frac{\tau(p)p}{\eta(p)}$ instantly, for any $p \in [q^P, q^R]$. One can then derive the underlying weight accumulation speed $\lambda^*(\cdot;\tau)$ as

$$\lambda^*(t;\tau) = \begin{cases} 1 - \frac{p_0}{q_t(\hat{\varphi}(\tau))}, & \text{if } t \in [0, T(q^P;\mathcal{P})) \\ 1 - \int_{p\in[q^P,q^R]} \frac{\tau(p)p}{\min\{q_t(\varphi(p)),p\}} dp, & \text{if } t \in [T(q^P;\mathcal{P}), \infty) \end{cases}$$

In the following Lemma, we show that $\lambda^*(\cdot;\tau)$ reaches the upper-bound of weight accumulation speed. We rewrite Lemma 4 as the following lemma.

Lemma 13. *For any pair of (λ, τ), which can be implemented by some direct and responsive dynamic disclosure policy, $\lambda^*(\cdot;\tau)$ is the fastest or*

$$\lambda(t) \leq \lambda^*(t;\tau), \ \forall t \in [0, \infty).$$

Proof. Let me define ϑ as follows

$$\vartheta := \inf\{t : \lambda(t') \leq \lambda^*(t';\tau), \ \forall t' \geq t\}.$$

Since $\lambda(T(q^R, \mathcal{P})) = \tau(0) = \lambda^*(T(q^R, \mathcal{P}), \tau)$, we have that $\vartheta \leq T(q^P, P)$. Moreover if $\vartheta \leq \hat{\varphi}(\tau)$, then $\vartheta = 0$. This is due to the fact that

$$\lambda^*(t;\tau) = \int_{s=0}^{t} \lambda(1 - p_s) ds, \forall t \in t^*(\tau)$$

which gives the upper-bound of weight accumulating speed. Our remaining task is to show that $\vartheta = 0$. Suppose the argument is not true, then $\vartheta > \hat{\varphi}(\tau)$. In the first step, we show that $\lambda(\vartheta) \leq \lambda^*(\vartheta;\tau)$. Suppose this is not true, then we argue that $\lambda(\vartheta) < \tau(0)$. Since $\lambda(\vartheta) \leq \tau(0)$, if the argument is wrong, then $\lambda(\vartheta) = \tau(0)$. Since both disclosure policies are responsive, we have that $\lambda^*(\vartheta;\tau) = \tau(0)$, which then leads to contradiction. At first, we consider the scenario in which $\vartheta \geq T(q^P, P)$. Moreover, the continuation payoff at ϑ is given as follows (it is induced from Equation (25)),

$$
\begin{aligned}
U_\vartheta(\lambda,\tau) &= \frac{\left[\begin{array}{c}\rho\int_\vartheta^\infty e^{-\rho(t-\vartheta)}[\lambda(t)-\lambda(\vartheta)]m(0)dt+\int_{p\in(p_\vartheta,q^R]}e^{-\rho(T(p;\mathcal{P})-\vartheta)}\tau(p)m(p)dp\\ -c\int_\vartheta^\infty e^{-\rho(t-\vartheta)}[1-\lambda(t)-\int_{p\in[q^P,p_t]}\tau(p)dp]dt\end{array}\right]}{1-\lambda(\vartheta)-\int_{p\in[q^P,p_\vartheta]}\tau(p)dp}\\
&\leqslant \frac{\left[\begin{array}{c}\rho\int_\vartheta^\infty e^{-\rho(t-\vartheta)}[\lambda^*(t;\tau)-\lambda(\vartheta)]m(0)dt+\int_{p\in(p_\vartheta,q^R]}e^{-\rho(T(p;\mathcal{P})-\vartheta)}\tau(p)m(p)dp\\ -c\int_\vartheta^\infty e^{-\rho(t-\vartheta)}[1-\lambda^*(t;\tau)-\int_{p\in[q^P,p_t]}\tau(p)dp]dt\end{array}\right]}{1-\lambda(\vartheta)-\int_{p\in[q^P,p_\vartheta]}\tau(p)dp}\\
&< \frac{\left[\begin{array}{c}\rho\int_\vartheta^\infty e^{-\rho(t-\vartheta)}[\lambda^*(t;\tau)-\lambda^*(\vartheta;\tau)]m(0)dt+\int_{p\in(p_\vartheta,q^R]}e^{-\rho(T(p;\mathcal{P})-\vartheta)}\tau(p)m(p)dp\\ -c\int_\vartheta^\infty e^{-\rho(t-\vartheta)}[1-\lambda^*(t;\tau)-\int_{p\in[q^P,p_t]}\tau(p)dp]dt\end{array}\right]}{1-\lambda(\vartheta)-\int_{p\in[q^P,p_\vartheta]}\tau(p)dp}\\
&= \frac{\int_{p\in(p_\vartheta,q^R]}\frac{\tau(p)p}{q_\vartheta(\varphi(p))}m(q_\vartheta(\varphi(p)))dp}{1-\lambda(\vartheta)-\int_{p\in[q^P,p_\vartheta]}\tau(p)dp} = \frac{\int_{p\in(p_\vartheta,q^R]}\frac{\tau(p)p}{q_\vartheta(\varphi(p))}q_\vartheta(\varphi(p))dp}{1-\lambda(\vartheta)-\int_{p\in[q^P,p_\vartheta]}\tau(p)dp}\\
&= \frac{p_0-\int_{p\in[q^P,p_\vartheta]}\tau(p)dp}{1-\lambda(\vartheta)-\int_{p\in[q^P,p_\vartheta]}\tau(p)dp} = m(\hat{p}_\vartheta(\lambda,\tau))
\end{aligned}
$$

The first inequality follows from the definition of ϑ, which implies that $\lambda(t) \le \lambda^*(t;\tau)$ for any $t > \vartheta$. The second inequality (the only strict inequality) follows from the assumption that $\lambda(\vartheta) > \lambda^*(\vartheta;\tau)$. The second equality follows from three facts. At first, the numerator of the term at the third line in the equation above is exactly the expected payoff to the agent contingent on the fact he keeps funding research until then. Secondly, at time t, the belief martingale assigns probability $\frac{\tau(p)p}{q_\vartheta(\varphi(p))}$ on belief $q_\vartheta(\varphi(p))$. Thirdly, given our constructed disclosure policy, the continuation payoff to the agent, upon receiving message which implies posterior $q_\vartheta(\varphi(p))$, is given by $m(q_\vartheta(\varphi(p)))$. The forth equality follows from Bayesian plausibility constraint. The last equality follows from the definition of $\hat{p}_t(\lambda,\tau)$, which is implied by Equation (23). At time ϑ, if the agent has not yet received a message confirming the state being ω_1 and forms belief $\hat{p}_t(\lambda,\tau)$, he, therefore, prefers to stop funding, making decision right away, rather than keep funding research, putting off decision making in the future. This then leads to the failure of implementation of stopping belief distribution τ and therefore contradiction.

Similarly, in the case where $\vartheta \in (\hat{\varphi}(\tau), T(q^P, P))$, the continuation payoff at ϑ, given direct and responsive disclosure policy (λ,τ), is

$$
\begin{aligned}
U_\vartheta(\lambda,\tau) &= \frac{\begin{bmatrix} \rho\displaystyle\int_\vartheta^\infty e^{-\rho(t-\vartheta)}[\lambda(t)-\lambda(\vartheta)]m(0)dt + \int_{p\in[q^P,q^P]} e^{-\rho(T(p;\mathcal{P})-\vartheta)}\tau(p)m(p)dp \\ -c\displaystyle\int_\vartheta^\infty e^{-\rho(t-\vartheta)}[1-\lambda(t)-\int_{p\in[q^P,p_t]}\tau(p)dp]dt \end{bmatrix}}{1-\lambda(\vartheta)} \\
&\le \frac{\begin{bmatrix} \rho\displaystyle\int_\vartheta^\infty e^{-\rho(t-\vartheta)}[\lambda^*(t;\tau)-\lambda(\vartheta)]m(0)dt + \int_{p\in[q^P,q^P]} e^{-\rho(T(p;\mathcal{P})-\vartheta)}\tau(p)m(p)dp \\ -c\displaystyle\int_\vartheta^\infty e^{-\rho(t-\vartheta)}[1-\lambda^*(t;\tau)-\int_{p\in[q^P,p_t]}\tau(p)dp]dt \end{bmatrix}}{1-\lambda(\vartheta)} \\
&< \frac{\begin{bmatrix} \rho\displaystyle\int_\vartheta^\infty e^{-\rho(t-\vartheta)}[\lambda^*(t;\tau)-\lambda^*(\vartheta;\tau)]m(0)dt + \int_{p\in[q^P,q^P]} e^{-\rho(T(p;\mathcal{P})-\vartheta)}\tau(p)m(p)dp \\ -c\displaystyle\int_\vartheta^\infty e^{-\rho(t-\vartheta)}[1-\lambda^*(t;\tau)-\int_{p\in[q^P,p_t]}\tau(p)dp]dt \end{bmatrix}}{1-\lambda(\vartheta)} \\
&= \frac{\frac{p_0}{q_\vartheta(\hat\varphi(\tau))}m(q_\vartheta(\hat\varphi(\tau)))}{1-\lambda(\vartheta)} = \frac{\frac{p_0}{q_\vartheta(\hat\varphi(\tau))}q_\vartheta(\hat\varphi(\tau))}{1-\lambda(\vartheta)} = m(\hat p_\vartheta(\lambda,\tau))
\end{aligned}
$$

The second equality follows from the fact that, given our constructed belief martingale, at $\vartheta \in (\hat\varphi(\tau), T(q^P, P))$, the agent assigns probability $\frac{\tau(p)p}{q_\vartheta(\varphi(p))}$ on belief $q_\vartheta(\varphi(p))$. Moreover, given this belief, the agent chooses to continue to fund research and gain continuation payoff $m(q_\vartheta(\varphi(p)))$. The last equality follows from the definition of $\hat p_\vartheta(\lambda,\tau)$. A contradiction, again, has been reached. All in all, it can not be true that $\lambda(\vartheta) > \lambda^*(\vartheta;\tau)$, otherwise stopping belief distribution τ can not be implemented.

If $\lambda(\vartheta) \le \lambda^*(\vartheta;\tau)$ and $\vartheta > 0$, then for any $\varepsilon > 0$, there exists some $t \in (\vartheta-\varepsilon, \vartheta)$ such that $\lambda(t) > \lambda^*(t;\tau)$. At first, we focus on the case in which $\vartheta > T(q^P, P)$ (The case in which $\vartheta \in (t^*(\tau), T(q^P, P)]$ is similar and therefore ignored) we pick ε small enough such that

$$
\begin{cases} \dfrac{m(0)}{[c+\rho m(0)]\varepsilon} > 1 \\ \vartheta - \varepsilon \geqslant T(q^P, P) \end{cases}
$$

Therefore the continuation payoff at ϑ, given direct and responsive disclosure policy (λ,τ), is given as follows

$$U_t(\lambda,\tau)=\frac{\left[\begin{array}{c}\rho\int_t^\infty e^{-\rho(s-t)}[\lambda(s)-\lambda(t)]m(0)ds+\int_{p\in(p_t,q^R]}e^{-\rho(T(p;\mathcal{P})-t)}\tau(p)m(p)dp\\ -c\int_t^\infty e^{-\rho(s-t)}[1-\lambda(s)-\int_{p\in[q^P,p_s]}\tau(p)dp]ds\end{array}\right]}{1-\lambda(t)-\int_{p\in[q^P,p_t]}\tau(p)dp}$$

$$\leq\frac{\left[\begin{array}{c}\rho\int_t^\infty e^{-\rho(s-t)}[\lambda^*(s;\tau)-\lambda^*(t;\tau)]m(0)ds+\int_{p\in(p_t,q^R]}e^{-\rho(T(p;\mathcal{P})-t)}\tau(p)m(p)dp\\ -c\int_t^\infty e^{-\rho(s-t)}[1-\lambda^*(s;\tau)-\int_{p\in[q^P,p_s]}\tau(p)dp]ds\end{array}\right]}{1-\lambda(t)-\int_{p\in[q^P,p_t]}\tau(p)dp}$$

$$+\frac{[c+\rho m(0)]\int_t^\vartheta e^{-\rho(s-t)}[\lambda(s)-\lambda^*(s;\tau)]ds-[\lambda(t)-\lambda^*(t;\tau)]m(0)}{1-\lambda(t)-\int_{p\in[q^P,p_t]}\tau(p)dp}$$

$$=m(\hat{p}_t(\lambda,\tau))+\frac{[c+\rho m(0)]\int_t^\vartheta e^{-\rho(s-t)}[\lambda(s)-\lambda^*(s;\tau)]ds-[\lambda(t)-\lambda^*(t;\tau)]m(0)}{1-\lambda(t)-\int_{p\in[q^P,p_t]}\tau(p)dp}$$

The first inequality follows from the following equation

$$\begin{aligned}
&\rho\int_t^\infty e^{-\rho(s-t)}[\lambda(s)-\lambda(t)]m(0)ds\\
&=\rho\int_t^\vartheta e^{-\rho(s-t)}[\lambda(s)-\lambda(t)]m(0)ds+\rho\int_\vartheta^\infty e^{-\rho(s-t)}[\lambda(s)-\lambda(t)]m(0)ds\\
&\leq\rho\int_t^\vartheta e^{-\rho(s-t)}[\lambda(s)-\lambda(t)]m(0)ds+\rho\int_\vartheta^\infty e^{-\rho(s-t)}[\lambda^*(s;\tau)-\lambda(t)]m(0)ds\\
&=\rho\int_t^\vartheta e^{-\rho(s-t)}[\lambda^*(s;\tau)-\lambda(t)]m(0)ds+\rho\int_\vartheta^\infty e^{-\rho(s-t)}[\lambda^*(s;\tau)-\lambda(t)]m(0)ds+\rho\int_t^\vartheta e^{-\rho(s-t)}[\lambda(s)-\lambda^*(s;\tau)]m(0)ds\\
&=\rho\int_t^\infty e^{-\rho(s-t)}[\lambda^*(s;\tau)-\lambda(t)]m(0)ds+\rho\int_t^\vartheta e^{-\rho(s-t)}[\lambda(s)-\lambda^*(s;\tau)]m(0)ds\\
&=\rho\int_t^\infty e^{-\rho(s-t)}[\lambda^*(s;\tau)-\lambda^*(t;\tau)]m(0)ds-\rho\int_t^\infty e^{-\rho(s-t)}ds[\lambda(t)-\lambda^*(t;\tau)]m(0)+\rho\int_t^\vartheta e^{-\rho(s-t)}[\lambda(s)-\lambda^*(s;\tau)]m(0)ds\\
&=\rho\int_t^\infty e^{-\rho(s-t)}[\lambda^*(s;\tau)-\lambda^*(t;\tau)]m(0)ds-[\lambda(t)-\lambda^*(t;\tau)]m(0)+\rho\int_t^\vartheta e^{-\rho(s-t)}[\lambda(s)-\lambda^*(s;\tau)]m(0)ds
\end{aligned}$$

Similarly, we also have that

$$\begin{aligned}
&\int_{s>t}^\infty e^{-\rho(s-t)}[1-\lambda(s)-\int_{p\in[q^P,p_s]}\tau(p)dp]ds\\
&\geqslant\int_t^\infty e^{-\rho(s-t)}[1-\lambda^*(s;\tau)-\int_{p\in p\in[q^P,p_s]}\tau(p)dp]ds-\int_{s\in(t,\vartheta)}e^{-\rho(s-t)}[\lambda(s)-\lambda^*(s;\tau)]ds
\end{aligned}$$

The second equality follows from the previous argument and therefore ignored. Henceforth, to satisfy incentive compatibility constraint, it is required that

$$[c+\rho m(0)]\int_t^\vartheta e^{-\rho(s-t)}[\lambda(s)-\lambda^*(s;\tau)]ds-[\lambda(t)-\lambda^*(t;\tau)]m(0)\geq 0.$$

Therefore it requires that there exists some $t_1 \in (t, \vartheta)$ such that

$$[c + \rho m(0)]e^{-\rho(t_1 - t)}[\lambda(t_1) - \lambda^*(t_1; \tau)](\vartheta - t) \geqslant m(0)[\lambda(t) - \lambda^*(t; \tau)]$$

Moreover since we have that

$$[c + \rho m(0)]e^{-\rho(t_1 - t)}[\lambda(t_1) - \lambda^*(t_1; \tau)](\vartheta - t) \leqslant [c + \rho m(0)][\lambda(t_1) - \lambda^*(t_1; \tau)]\varepsilon,$$

combining these two inequalities, we then have that

$$[\lambda(t_1) - \lambda^*(t_1; \tau)] \geq \frac{m(0)}{[c + \rho m(0)]\varepsilon}[\lambda(t) - \lambda^*(t; \tau)].$$

Since the term $\frac{m(0)}{[c+\rho m(0)]\varepsilon}$ is strictly larger than 1, if we repeat this process by some finite n times, we can find some $t_n \in (t_{n-1}, \vartheta)$ such that

$$[\lambda(t_n) - \lambda^*(t_n; \tau)] \geq \left[\frac{m(0)}{[c + \rho m(0)]\varepsilon}\right]^n [\lambda(t) - \lambda^*(t; \tau)] > \lambda^*(\vartheta; \tau) - \lambda^*(t; \tau),$$

which therefore achieves contradiction since $\lambda(t)$ is monotonically increasing and $\lambda(\vartheta) \leq \lambda^*(\vartheta; \tau)$. □

To end this subsection, we relax the assumption adopted at the beginning of this subsection by considering that

$$q_{T(q^P;\mathcal{P})}(0) > \eta(q^R).$$

Our next corollary shows that for any τ^π which can be implemented by some direct and responsive dynamic disclosure policy (π, M), then there exists some $t \in [0, T(q^P; \mathcal{P})]$ such that $q_{T(q^P;\mathcal{P})}(t) = \hat{\eta}(\tau)$.

Corollary 4. *For any direct and responsive dynamic disclosure policy (π, M) with implemented (λ^π, τ^π), then*

$$q_{T(q^P;\mathcal{P})}(0) \leq \hat{\eta}(\tau^\pi).$$

Proof. Suppose on the contrary, we have that $q_{T(q^P;\mathcal{P})}(0) > \hat{\eta}(\tau^\pi)$ for some (λ^π, τ^π). Then there exists $\zeta \in (0, 1)$ such that

$$\zeta q^P + (1 - \zeta)\hat{\eta}(\tau^\pi) = q_{T(q^P;\mathcal{P})}(0).$$

Then let me denote (π^F, M^F) as the full information disclosure policy. Moreover, let me define the following disclosure policy (π', M') as, at time zero, the expert, commits to dynamic disclosure policy (π^F, M^F) with probability ζ and to policy (π, M) with the rest probability $1-\zeta$. One then have that

$$U_0(\lambda^{\pi'},\tau^{\pi'}) = \zeta U_0(\lambda^{\pi^F},\tau^{\pi^F}) + (1-\zeta)U_0(\lambda^{\pi},\tau^{\pi}) > U_0(\lambda^{\pi},\tau^{\pi}).$$

Moreover, Lemma 13 implies that

$$\lambda^*(t;\tau^{\pi'}) \geq \lambda^{\pi'}(t), \forall t \in [0,\infty).$$

Therefore, we have that

$$U_0(\lambda^*(t;\tau^{\pi'}),\tau^{\pi'}) \geq U_0(\lambda^{\pi'},\tau^{\pi'}) > U_0(\lambda^{\pi},\tau^{\pi}).$$

Moreover, by the definition of ζ, one then have that $\hat{\varphi}(\tau^{\pi'}) = 0$. Therefore, we also have that

$$m(p_0) = U_0(\lambda^*(t;\tau^{\pi'}),\tau^{\pi'}) > U_0(\lambda^{\pi},\tau^{\pi}).$$

We have therefore reached a contradiction as $(\lambda^{\pi},\tau^{\pi})$ can not be implemented. □

C.9.5. *Optimal stopping belief distribution* τ. Our remaining task, after the procedures of refinery in previous subsection, is to pin down the optimal stopping belief distribution τ. To do this, let me remind you how we combine these two parts of belief martingales in the previous subsection. To implement some stopping belief distribution τ, at time $T(q^P;\mathcal{P})$, the expert is supposed to conduct a splitting of $\hat{\eta}(\tau)$ among interval $[q^P,\eta(q^R)]$, such that the induced belief martingale jumps to belief $\eta(p)$ with probability $\hat{\gamma}(p;\tau) = \frac{\gamma(p;\tau)}{\int_{p\in[q^P,q^R]}\gamma(p;\tau)dp}$, for any $p \in [q^P,q^R]$.

For any belief $\mu \in [q^P,\eta(q^R)]$, let me define $\tilde{\mathcal{Q}}^{\mu}$ as a margin belief process with starting belief μ. Specifically,

$$\tilde{q}_t^{\mu} = \frac{c\mu e^{ct}}{\rho\mu[1-e^{ct}]+c} \wedge q_t \leq 1.$$

Then let me define $\tilde{\theta}(q)$ as the time when research progress process $\mathcal{P}$ with starting belief q^P intersects with the margin belief process $\tilde{Q}^{\mu}$. Mathematically,

$$\frac{c\mu e^{c\tilde{\theta}(\mu)}}{\rho\mu[1-e^{c\tilde{\theta}(\mu)}]+c} = \frac{q^P e^{\lambda\tilde{\theta}(\mu)}}{q^P(e^{\lambda\tilde{\theta}(\mu)}-1)+1} \tag{26}$$

One can turn to Lemma 2 to show that $\tilde{\theta}(\cdot)$ is well defined and strictly decreasing in μ. One can simply verify that $\tilde{\theta}(q^R) = 0$. Finally, let me define value function $\tilde{V}(\mu)$ as

$$\tilde{V}(\mu) = m(\mu) + c\int_0^{\tilde{\theta}(\mu)} e^{-\rho t}\frac{\mu}{\tilde{q}_t^{\mu}}dt.$$

Kamenica and Gentzkow [2011] implies that, fix some transition time $\hat{t} \in [0,T(q^P;\mathcal{P})]$, the optimal stopping belief distribution $\tau^*(\cdot;\hat{t})$ can be derived through concavification of

$\tilde{V}(\cdot)$ within interval $[q^P, \eta(q^R)]$ at belief $p_{T(q^P;\mathcal{P})}(\hat{t})$. Our next Lemma shows that $\tilde{V}(\cdot)$ is concave within interval $[q^P, \eta(q^R)]$, which, as an immediate corollary, implies that

$$\tau^*(p;\hat{t}) = \begin{cases} \frac{p_0}{p_{\theta(\hat{t})}(\hat{t})}, & \text{if } p = p_{\theta(\hat{t})}(\hat{t}) \\ 1 - \frac{p_0}{p_{\theta(\hat{t})}(\hat{t})}, & \text{if } p = 0 \\ 0, & \text{otherwise} \end{cases}$$

We rewrite Lemma 5 as the following lemma.

Lemma 14. *$\tilde{V}(\mu)$ is concave in μ for any $\mu \in [q^P, \eta(q^R)]$.*

Proof. In the first step, we simplify $\tilde{V}(\mu)$ and show that $\tilde{V}(\mu)$ is concave in μ if and only if function

$$W(\mu) := -\mu e^{-(\lambda+\rho)\tilde{\theta}(\mu)}$$

is concave in μ. One can rewrite $\tilde{V}(\mu)$ as follows

$$\begin{aligned}
&\tilde{V}(\mu) = m(\mu) + c\int_0^{\tilde{\theta}(\mu)} e^{-\rho t}\tfrac{\mu}{\tilde{q}_t^{\mu}}dt = \mu + c\int_0^{\tilde{\theta}(\mu)} e^{-\rho t}\tfrac{\mu}{\frac{c\mu e^{ct}}{\rho\mu(1-e^{ct})+c}}dt \\
&= \mu + \int_0^{\tilde{\theta}(\mu)} e^{-\rho t}\tfrac{\rho\mu(1-e^{ct})+c}{e^{ct}}dt = \mu + \int_0^{\tilde{\theta}(\mu)} e^{-(\rho+c)t}(\rho\mu+c)dt - \rho\mu\int_0^{\tilde{\theta}(\mu)} e^{-\rho t}dt \\
&= \mu + \tfrac{\rho\mu+c}{\rho+c}(1-e^{-(\rho+c)\tilde{\theta}(\mu)}) - \mu(1-e^{-\rho\tilde{\theta}(\mu)}) = \tfrac{\rho\mu+c}{\rho+c} + e^{-\rho\tilde{\theta}(\mu)}(\mu - \tfrac{\rho\mu+c}{\rho+c}e^{-c\tilde{\theta}(\mu)}) \\
&= \tfrac{\rho\mu+c}{\rho+c} + e^{-\rho\tilde{\theta}(\mu)}\tfrac{c\mu+\rho\mu-(\rho\mu+c)e^{-c\tilde{\theta}(\mu)}}{\rho+c}
\end{aligned}$$

By the definition of $\theta(q)$, which is defined in Equation (26), one then have that

$$(\rho\mu+c)e^{-c\tilde{\theta}(\mu)} - \rho\mu = \frac{c\mu[q^P + (1-q^P)e^{-\lambda\tilde{\theta}(\mu)}]}{q^P}.$$

Therefore, one can further simplify $\tilde{V}(\mu)$ as

$$\tilde{V}(\mu) = \frac{\rho\mu+c}{\rho+c} + e^{-\rho\tilde{\theta}(\mu)}\frac{c\mu - \frac{c\mu[q^P+(1-q^P)e^{-\lambda\tilde{\theta}(\mu)}]}{q^P}}{\rho+c} = \frac{\rho\mu+c}{\rho+c} - \frac{1-q^P}{q^P}\frac{c}{\rho+c}\mu e^{-(\lambda+\rho)\tilde{\theta}(\mu)} = \frac{\rho\mu+c}{\rho+c} + \frac{1-q^P}{q^P}\frac{c}{\rho+c}W(\mu).$$

Therefore $\tilde{V}''(\mu)$ is negative in $(q^P, \eta(q^R))$ if and only if $W''(\mu)$ is negative in $(q^P, \eta(q^R))$.

Then let me take a derivative of $W(\cdot)$ as

$$W'(\mu) = -e^{-(\lambda+\rho)\tilde{\theta}(\mu)} + (\lambda+\rho)\mu e^{-(\lambda+\rho)\tilde{\theta}(\mu)}\tilde{\theta}'(\mu) = e^{-(\lambda+\rho)\tilde{\theta}(\mu)}[(\lambda+\rho)\mu\tilde{\theta}'(\mu) - 1].$$

Further, let me take a derivative of $W'(\cdot)$ as

$$\begin{aligned}
&W''(\mu) = -e^{-(\lambda+\rho)\tilde{\theta}(\mu)}[(\lambda+\rho)\mu\tilde{\theta}'(\mu) - 1](\lambda+\rho)\tilde{\theta}'(\mu) + e^{-(\lambda+\rho)\tilde{\theta}(\mu)}(\lambda+\rho)[\tilde{\theta}'(\mu) + \mu\tilde{\theta}''(\mu)] \\
&= e^{-(\lambda+\rho)\tilde{\theta}(\mu)}(\lambda+\rho)[2\tilde{\theta}'(\mu) + \mu\tilde{\theta}''(\mu) - (\lambda+\rho)(\tilde{\theta}'(\mu))^2]
\end{aligned}$$

Our remaining task is to show that the term

$$2\tilde{\theta}'(\mu)+\mu\tilde{\theta}''(\mu)-(\lambda+\rho)(\tilde{\theta}'(\mu))^2<0 \tag{27}$$

In the second step, we show that the term

$$2\tilde{\theta}'(\mu)+\mu\theta''(\mu)+\mu\frac{c(c+\rho)+(\lambda-c)^2\frac{1-q^P}{q^P}e^{-\lambda\tilde{\theta}(\mu)}}{(c+\rho)-\frac{1-q^P}{q^P}(\lambda-c)e^{-\lambda\tilde{\theta}(\mu)}}(\tilde{\theta}'(\mu))^2=0,$$

for any $\mu\in(q^P,\eta(q^R))$. Then Equation (27) immediately follows since $\theta'(\mu)$ is strictly negative from Lemma 2. From Equation (26) implies that

$$\frac{c\mu e^{c\tilde{\theta}(\mu)}}{\rho\mu(1-e^{c\tilde{\theta}(\mu)})+c}=\frac{q^Pe^{\lambda\tilde{\theta}(\mu)}}{q^P(e^{\lambda\tilde{\theta}(\mu)}-1)+1}\Rightarrow\frac{c}{(\rho+\frac{c}{\mu})e^{-c\tilde{\theta}(\mu)}-\rho}=\frac{q^P}{q^P+(1-q^P)e^{-\lambda\tilde{\theta}(\mu)}}$$
$$\Rightarrow c[q^P+(1-q^P)e^{-\lambda\tilde{\theta}(\mu)}]=q^P[(\rho+\tfrac{c}{\mu})e^{-c\tilde{\theta}(\mu)}-\rho]\Rightarrow\rho+\tfrac{c}{\mu}=(c+\rho)e^{c\tilde{\theta}(\mu)}+\tfrac{c(1-q^P)}{q^P}e^{-(\lambda-c)\tilde{\theta}(\mu)}$$

Let me take derivative of μ on both sides of the equation above, one then have that

$$\begin{aligned}&-\tfrac{c}{\mu^2}=(c+\rho)e^{c\tilde{\theta}(\mu)}c\tilde{\theta}'(\mu)+\tfrac{c(1-q^P)}{q^P}e^{-(\lambda-c)\tilde{\theta}(\mu)}(-1)(\lambda-c)\tilde{\theta}'(\mu)\\&\Rightarrow-\tfrac{1}{\mu^2}=e^{c\tilde{\theta}(\mu)}\tilde{\theta}'(\mu)[(c+\rho)-\tfrac{1-q^P}{q^P}(\lambda-c)e^{-\lambda\tilde{\theta}(\mu)}]\end{aligned} \tag{28}$$

One can easily verify that $\tilde{\theta}'(\mu)$ is strictly negative in interval $(q^P,\eta(q^R))$ from Equation (28) since

$$\frac{1-q^P}{q^P}e^{-\lambda\tilde{\theta}(q)}<\frac{1-q^P}{q^P}=\frac{1-\frac{\lambda-c}{\lambda+\rho}}{\frac{\lambda-c}{\lambda+\rho}}=\frac{c+\rho}{\lambda-c},$$

for any $\mu\in(q^P,\eta(q^R))$. Let me further take derivative of μ on both sides of the Equation (28), one then have that

$$\begin{aligned}&\tfrac{2}{\mu^3}=e^{c\tilde{\theta}(\mu)}c(\tilde{\theta}'(\mu))^2[(c+\rho)-\tfrac{1-q^P}{q^P}(\lambda-c)e^{-\lambda\tilde{\theta}(\mu)}]+e^{c\tilde{\theta}(\mu)}\tilde{\theta}''(\mu)[(c+\rho)-\tfrac{1-q^P}{q^P}(\lambda-c)e^{-\lambda\tilde{\theta}(\mu)}]\\&+e^{c\tilde{\theta}(\mu)}\tfrac{1-q^P}{q^P}\lambda(\lambda-c)e^{-\lambda\tilde{\theta}(\mu)}(\tilde{\theta}'(\mu))^2\\&=e^{c\tilde{\theta}(\mu)}[(c+\rho)-\tfrac{1-q^P}{q^P}(\lambda-c)e^{-\lambda\tilde{\theta}(\mu)}]\left[\tilde{\theta}(\mu)+\frac{c(c+\rho)+(\lambda-c)^2\frac{1-q^P}{q^P}e^{-\lambda\tilde{\theta}(\mu)}}{(c+\rho)-\frac{1-q^P}{q^P}(\lambda-c)e^{-\lambda\tilde{\theta}(\mu)}}\right]\end{aligned}$$

Moreover Equation (28) implies that

$$\frac{2}{\mu^3}=(-\frac{2}{\mu})\times(-\frac{1}{\mu^2})=-\frac{2}{\mu}e^{c\tilde{\theta}(\mu)}\tilde{\theta}'(\mu)[(c+\rho)-\frac{1-q^P}{q^P}(\lambda-c)e^{-\lambda\tilde{\theta}(\mu)}].$$

Accordingly, we have that

$$\begin{aligned}&e^{c\tilde{\theta}(\mu)}[(c+\rho)-\tfrac{1-q^P}{q^P}(\lambda-c)e^{-\lambda\tilde{\theta}(\mu)}]\left[\tilde{\theta}''(\mu)+\frac{c(c+\rho)+(\lambda-c)^2\frac{1-q^P}{q^P}e^{-\lambda\tilde{\theta}(\mu)}}{(c+\rho)-\frac{1-q^P}{q^P}(\lambda-c)e^{-\lambda\tilde{\theta}(\mu)}}\right]\\&=-\tfrac{2}{q}e^{c\tilde{\theta}(\mu)}\tilde{\theta}'(\mu)[(c+\rho)-\tfrac{1-q^P}{q^P}(\lambda-c)e^{-\lambda\tilde{\theta}(\mu)}]\\&\Rightarrow\frac{e^{c\tilde{\theta}(\mu)}[(c+\rho)-\frac{1-q^P}{q^P}(\lambda-c)e^{-\lambda\tilde{\theta}(\mu)}]}{q}\left[2\tilde{\theta}'(\mu)+\mu\tilde{\theta}''(\mu)+\mu\frac{c(c+\rho)+(\lambda-c)^2\frac{1-q^P}{q^P}e^{-\lambda\tilde{\theta}(\mu)}}{(c+\rho)-\frac{1-q^P}{q^P}(\lambda-c)e^{-\lambda\tilde{\theta}(\mu)}}(\tilde{\theta}'(\mu))^2\right]=0\end{aligned}$$

The argument has therefore been shown. $\square$

C.10. **Derivation of one-step-ahead process in Section 4.5.** Equation (8) implies that

$$\begin{aligned}
&\frac{dq_t}{q_t}[m(0)-u(a_1;0)] = [\rho m(q_t)+c]dt \\
\Leftrightarrow &\frac{dq_t}{q_t[\rho m(q_t)+c]} = \frac{dt}{m(0)-u(a_1;0)} \\
\Leftrightarrow &\frac{dq_t}{q_t\left[\rho[u(a_1;1)-u(a_1;0)]q_t+\rho u(a_1;0)+c\right]} = \frac{dt}{m(0)-u(a_1;0)} \\
\Leftrightarrow &\frac{dq_t}{q_t} - \frac{\rho[u(a_1;1)-u(a_1;0)]dq_t}{\rho[u(a_1;1)-u(a_1;0)]q_t+\rho u(a_1;0)+c} = \frac{[\rho u(a_1;0)+c]dt}{m(0)-u(a_1;0)} \\
\Leftrightarrow &d\ln q_t - d\ln\left[\rho[u(a_1;1)-u(a_1;0)]q_t+\rho u(a_1;0)+c\right] = \frac{[\rho u(a_1;0)+c]dt}{m(0)-u(a_1;0)} \\
\Leftrightarrow &d\ln\left[\frac{q_t}{\rho[u(a_1;1)-u(a_1;0)]q_t+\rho u(a_1;0)+c}\right] = \frac{[\rho u(a_1;0)+c]dt}{m(0)-u(a_1;0)} \\
\Leftrightarrow &\frac{q_t}{\rho[u(a_1;1)-u(a_1;0)]q_t+\rho u(a_1;0)+c} = \frac{q_0}{\rho[u(a_1;1)-u(a_1;0)]q_0+\rho u(a_1;0)+c}\exp\left(\frac{[\rho u(a_1;0)+c]}{m(0)-u(a_1;0)}t\right) \\
\Leftrightarrow &q_t = \frac{[\rho u(a_1;0)+c]\exp\left(\frac{[\rho u(a_1;0)+c]}{m(0)-u(a_1;0)}t\right)}{\rho[u(a_1;1)-u(a_1;0)]\left[\exp\left(\frac{[\rho u(a_1;0)+c]}{m(0)-u(a_1;0)}t\right)-1\right]+\frac{\rho u(a_1;0)+c}{q_0}}.
\end{aligned}$$

C.11. **Proof of Proposition 2 and 3.** All results prior Lemma 5 holds. We then show that if $u(a_2;\omega_2)\geq u(a_1;\omega_1)$, Lemma 14 holds.

We at first consider the definition of $\tilde{\theta}(\mu)$, given as follows

$$\frac{q^P e^{\lambda\tilde{\theta}(\mu)}}{q^P e^{\lambda\tilde{\theta}(\mu)}+(1-q^P)} = \frac{\mu[\rho u(a_2;\omega_1)+c]\exp\left(\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}\tilde{\theta}(\mu)\right)}{\{\rho\mu[u(a_2;\omega_2)-u(a_2;\omega_1)]+\rho u(a_2;\omega_1)+c\}-\rho\mu[u(a_2;\omega_2)-u(a_2;\omega_1)]\exp\left(\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}t\right)},$$

which then implies that

$$\frac{1-q^P}{q^P}e^{-\lambda\tilde{\theta}(\mu)}+1+\frac{\rho[u(a_2;\omega_2)-u(a_2;\omega_1)]}{\rho u(a_2;\omega_1)+c} = \frac{\rho\mu[u(a_2;\omega_2)-u(a_2;\omega_1)]+\rho u(a_2;\omega_1)+c}{\mu[\rho u(a_2;\omega_1)+c]}\exp\left(-\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}\tilde{\theta}(\mu)\right) \quad (29)$$

The expert's continuation payoff is given by

$$
\begin{aligned}
\tilde{V}(\mu) &= m(\mu) + c\int_0^{\tilde{\theta}(\mu)} e^{-\rho t}\frac{\mu}{\tilde{q}_t^{\mu}}dt \\
&= m(\mu) + c\int_0^{\tilde{\theta}(\mu)} e^{-\rho t}\frac{\mu}{\frac{[\rho u(a_2;\omega_1)+c]\exp\left(\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}t\right)}{\rho[u(a_2;\omega_2)-u(a_2;\omega_1)]\left[1-\exp\left(\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}t\right)\right]+\frac{\rho u(a_2;\omega_1)+c}{\mu}}}dt \\
&= m(\mu) + c\int_0^{\tilde{\theta}(\mu)} e^{-\rho t}\left(\frac{\rho\mu[u(a_2;\omega_2)-u(a_2;\omega_1)]+\rho u(a_2;\omega_1)+c}{\rho u(a_2;\omega_1)+c}\exp\left(-\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}t\right) - \frac{\rho\mu[u(a_2;\omega_2)-u(a_2;\omega_1)]}{\rho u(a_2;\omega_1)+c}\right)dt \\
&= m(\mu) - c\frac{\rho\mu[u(a_2;\omega_2)-u(a_2;\omega_1)]}{\rho u(a_2;\omega_1)+c}\int_0^{\tilde{\theta}(\mu)} e^{-\rho t}dt + c\frac{\rho\mu[u(a_2;\omega_2)-u(a_2;\omega_1)]+\rho u(a_2;\omega_1)+c}{\rho u(a_2;\omega_1)+c}\int_0^{\tilde{\theta}(\mu)}\exp\left(-\left(\rho+\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}\right)\right. \\
&= \begin{aligned} & m(\mu) - c\frac{\rho\mu[u(a_2;\omega_2)-u(a_2;\omega_1)]}{\rho u(a_2;\omega_1)+c}\frac{1}{\rho}(1-e^{-\rho\tilde{\theta}(\mu)}) \\ & + c\frac{\rho\mu[u(a_2;\omega_2)-u(a_2;\omega_1)]+\rho u(a_2;\omega_1)+c}{\rho u(a_2;\omega_1)+c}\frac{u(a_1;\omega_1)-u(a_2;\omega_1)}{\rho[u(a_1;\omega_1)-u(a_2;\omega_1)]+\rho u(a_2;\omega_1)+c}\left[1-\exp\left(-\left(\rho+\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}\right)\tilde{\theta}(\mu)\right)\right]\end{aligned} \\
&= \begin{aligned} & f(\mu) + c\frac{\mu[u(a_2;\omega_2)-u(a_2;\omega_1)]}{\rho u(a_2;\omega_1)+c}e^{-\rho\tilde{\theta}(\mu)} \\ & - c\frac{u(a_1;\omega_1)-u(a_2;\omega_1)}{\rho[u(a_1;\omega_1)-u(a_2;\omega_1)]+\rho u(a_2;\omega_1)+c}\frac{\rho\mu[u(a_2;\omega_2)-u(a_2;\omega_1)]+\rho u(a_2;\omega_1)+c}{\rho u(a_2;\omega_1)+c}\exp\left(-\left(\rho+\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}\right)\tilde{\theta}(\mu)\right)\end{aligned},
\end{aligned}
$$

for some linear function $f(\cdot)$.

Then we insert Equation (29) into the formula of $\tilde{V}(\mu)$, one then have that

$$
\begin{aligned}
\tilde{V}(\mu) &= f(\mu) + e^{-\rho\tilde{\theta}(\mu)}\left\{c\frac{\mu[u(a_2;\omega_2)-u(a_2;\omega_1)]}{\rho u(a_2;\omega_1)+c} - c\mu\frac{u(a_1;\omega_1)-u(a_2;\omega_1)}{\rho[u(a_1;\omega_1)-u(a_2;\omega_1)]+\rho u(a_2;\omega_1)+c}\left[\frac{1-q^P}{q^P}e^{-\lambda\tilde{\theta}(\mu)}+1+\frac{\rho[u(a_2;\omega_2)-u(a_2;\omega_1)]}{\rho u(a_2;\omega_1)+c}\right]\right\} \\
&= f(\mu) - c\frac{u(a_1;\omega_1)-u(a_2;\omega_1)}{\rho[u(a_1;\omega_1)-u(a_2;\omega_1)]+\rho u(a_2;\omega_1)+c}\left[\frac{1-q^P}{q^P}\mu e^{-(\lambda+\rho)\tilde{\theta}(\mu)} + \frac{\rho[u(a_2;\omega_2)-u(a_1;\omega_1)]}{\rho u(a_2;\omega_1)+c}\mu e^{-\rho\tilde{\theta}(\mu)}\right].
\end{aligned}
$$

Equation (29) implies that

$$
\begin{aligned}
&\frac{1-q^P}{q^P}\exp\left(\left[\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}-\lambda\right]\tilde{\theta}(\mu)\right) + \frac{\rho u(a_2;\omega_1)+c+\rho[u(a_2;\omega_2)-u(a_2;\omega_1)]}{\rho u(a_2;\omega_1)+c}\exp\left(\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}\tilde{\theta}(\mu)\right) \\
=&\frac{\rho[u(a_2;\omega_2)-u(a_2;\omega_1)]}{\rho u(a_2;\omega_1)+c}+\frac{1}{\mu}
\end{aligned} \tag{30}
$$

Take derivative of Equation (30) with respect to μ on both sides, we then have that

$$
\begin{aligned}
-\frac{1}{\mu^2} &= -\frac{1-q^P}{q^P}\exp\left(\left[\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}-\lambda\right]\tilde{\theta}(\mu)\right)\left[\lambda-\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}\right]\tilde{\theta}'(\mu) \\
&\quad + \frac{\rho u(a_2;\omega_1)+c+\rho[u(a_2;\omega_2)-u(a_2;\omega_1)]}{\rho u(a_2;\omega_1)+c}\exp\left(\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}\tilde{\theta}(\mu)\right)\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}\tilde{\theta}'(\mu) \\
&=\tilde{\theta}'(\mu)\exp\left(\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}\tilde{\theta}(\mu)\right)\left[\frac{\rho u(a_2;\omega_1)+c+\rho[u(a_2;\omega_2)-u(a_2;\omega_1)]}{u(a_1;\omega_1)-u(a_2;\omega_1)} - \frac{1-q^P}{q^P}\left[\lambda-\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}\right]e^{-\lambda\tilde{\theta}(\mu)}\right]
\end{aligned} \tag{31}
$$

Take derivative of Equation (31) with respect to μ on both sides, we then have that

$$\begin{aligned}\frac{2}{\mu^3} =&\tilde{\theta}''(\mu)\exp\left(\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}\tilde{\theta}(\mu)\right)g(\mu)\\ &+\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}\left(\tilde{\theta}'(\mu)\right)^2\exp\left(\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}\tilde{\theta}(\mu)\right)g(\mu)\\ &+\lambda\exp\left(\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}\tilde{\theta}(\mu)\right)\left(\tilde{\theta}'(\mu)\right)^2\frac{1-q^P}{q^P}\left[\lambda-\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}\right]e^{-\lambda\tilde{\theta}(\mu)}\end{aligned} \tag{32}$$

where

$$g(\mu)=\frac{\rho u(a_2;\omega_1)+c+\rho[u(a_2;\omega_2)-u(a_2;\omega_1)]}{u(a_1;\omega_1)-u(a_2;\omega_1)}-\frac{1-q^P}{q^P}\left[\lambda-\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}\right]e^{-\lambda\tilde{\theta}(\mu)}$$

Moreover, Equation (30) implies that

$$\begin{aligned}\frac{2}{\mu^3}=&-\frac{2}{\mu}\times-\frac{1}{\mu^2}\\ =&-\frac{2}{\mu}\tilde{\theta}'(\mu)\exp\left(\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}\tilde{\theta}(\mu)\right)g(\mu)\end{aligned} \tag{33}$$

Combining Equation (32) and (33), one then have that

$$\begin{aligned}0 =&\tilde{\theta}''(\mu)\exp\left(\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}\tilde{\theta}(\mu)\right)g(\mu)\\ &+\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}\left(\tilde{\theta}'(\mu)\right)^2\exp\left(\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}\tilde{\theta}(\mu)\right)g(\mu)\\ &+\lambda\exp\left(\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}\tilde{\theta}(\mu)\right)\left(\tilde{\theta}'(\mu)\right)^2\frac{1-q^P}{q^P}\left[\lambda-\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}\right]e^{-\lambda\tilde{\theta}(\mu)}\\ &+\frac{2}{\mu}\tilde{\theta}'(\mu)\exp\left(\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}\tilde{\theta}(\mu)\right)g(\mu),\end{aligned}$$

which then implies that

$$\mu\tilde{\theta}''(\mu)+2\tilde{\theta}'(\mu)+\left[\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}+\frac{\lambda\frac{1-q^P}{q^P}\left[\lambda-\frac{\rho u(a_2;\omega_1)+c}{u(a_1;\omega_1)-u(a_2;\omega_1)}\right]e^{-\lambda\tilde{\theta}(\mu)}}{g(\mu)}\right]\mu\left(\tilde{\theta}'(\mu)\right)^2=0 \tag{34}$$

We then show that the function $-\mu e^{-\alpha\mu}$ is concave in μ for any constant $\alpha>0$. Take the first derivative of $-\mu e^{-\alpha\tilde{\theta}(\mu)}$ with respect to μ, one then have that

$$\frac{d-\mu e^{-\alpha\tilde{\theta}(\mu)}}{d\mu}=e^{-\alpha\tilde{\theta}(\mu)}[\alpha\mu\tilde{\theta}'(\mu)-1]$$

Take the second derivative of $-\mu e^{-\alpha\tilde{\theta}(\mu)}$ with respect to μ, one thne have that

$$\frac{d^2-\mu e^{-\alpha\tilde{\theta}(\mu)}}{(d\mu)^2}=e^{-\alpha\tilde{\theta}(\mu)}\alpha[2\tilde{\theta}'(\mu)+\mu\tilde{\theta}''(\mu)-\alpha(\tilde{\theta}'(\mu))^2]<0.$$

The inequality is due to Equation (34).

Therefore, we argue that if $u(a_2;\omega_2) - u(a_1;\omega_1) \geq 0$, then the function $\tilde{V}(\mu)$ is strictly concave in μ. All our results hold. On the other hand, if $u(a_2;\omega_2) - u(a_1;\omega_1) < 0$, then the function $\tilde{V}(\mu)$ may be convex in μ. We then need persuasion to concavify $\tilde{V}(\mu)$.

C.12. **Derivation of research progress process $\hat{\mathcal{P}}$.** The derivation of the research progress process $\tilde{\mathcal{P}}$ is similar to that of the process $\mathcal{P}$ except that the intensity coefficient λ in the later process being replaced with the coefficient $\lambda(1-2\psi)$ in the former process.

The rest part of this subsection is devoted into the derivation of the belief martingale $\boldsymbol{\tau}^F$[16], induced by expert's information acquisition. Specifically, at any time t, the support of τ_t^F is given by supp$(\tau_t^F) = \{0, p_t, 1\}$. Moreover, at any time t, Bayesian persuasion constraint then implies that

$$\tau_t^F(0) + \tau_t^F(p_t) + \tau_t^F(1) = 1 \tag{35}$$

$$\tau_t^F(0)0 + \tau_t^F(p_t)p_t + \tau_t^F(1)1 = p_0 \tag{36}$$

At time $t + dt$, the probability $\tau_{t+dt}(1)$ on belief 1 is given by

$$\tau_{t+dt}^F(1) = \tau_t^F(1) + \tau_t^F(p_t)p_t\lambda\psi dt \implies d\tau_t^F(1) = \tau_t^F(p_t)p_t\lambda\psi dt$$

Introducing Equation (36) into the equation above, one then have that

$$d\tau_t^F(1) = [p_0 - \tau_t^F(1)]\lambda\psi dt \Rightarrow \frac{d\tau_t^F(1)}{p_0 - \tau_t^F(1)} = \lambda\psi dt \Rightarrow d\left(-\ln(p_0 - \tau_t^F(1))\right) = d\lambda\psi t$$

$$\Rightarrow \ln(p_0 - \tau_0^F(1)) - \ln(p_0 - \tau_t^F(1)) = \lambda\psi t \Rightarrow \ln\frac{p_0 - \tau_0^F(0)}{p_0 - \tau_t^F(1)} = \lambda\psi t$$

$$\Rightarrow \tau_t^F(1) = p_0 - (p_0 - \tau_0^F(t))e^{-\lambda\psi t} = p_0 - p_0 e^{-\lambda\psi t}$$

Besides, based on Equation (36), the probability $\tau_t^F(p_t)$ on belief p_t is given by

$$\tau_t^F(p_t) = \frac{p_0 - \tau_t^F(1)}{p_t} = \frac{p_0 - \left(p_0 - p_0 e^{-\lambda\psi t}\right)}{p_t} = \frac{p_0}{p_t}e^{-\lambda\psi t}.$$

Finally, based on Equation (35) implies that the probability $\tau_t(0)$ on belief 0 is given by

$$\tau_t^F(0) = 1 - \tau_t^F(1) - \tau_t^F(p_t) = 1 - p_0\left(1 - e^{-\lambda\psi t}\right) - \frac{p_0}{p_t}e^{-\lambda\psi t} = (1-p_0) - \frac{p_0}{p_t}e^{-\lambda\psi t}(1-p_t) = (1-p_t)\left[\frac{1-p_0}{1-p_t} - \frac{p_0}{p_t}e^{-\lambda\psi t}\right]$$

[16]The up-script F stands for full information disclosure.

C.13. **Derivation of margin belief process $\tilde{\mathcal{Q}}$.** Let me denote $\zeta_t \in \Delta(\Delta\{\omega_1, \omega_2\})$ as a belief martingale with support $\{0, q_{t+dt}, 1\}$. To meet the condition of perfectly disclosing breakthrough signal b_2 immediately whenever receiving it, we therefore set $\zeta_t(1)$ as the probability that the researcher privately observes this signal. Mathematically, the following equation holds,

$$\zeta_t(1) = \frac{\tilde{q}_t}{\tilde{p}_t}\tilde{p}_t \lambda\psi dt = \tilde{q}_t \lambda\psi dt \tag{37}$$

Besides, increment $d\tilde{q}_t = \tilde{q}_{t+dq} - \tilde{q}_t$ is set such that the policymaker is indifferent between making decision immediately and funding research for another dt period, being better informed under belief martingale ζ_t and putting off decision making dt period later. Mathematically, $\tilde{q}_{t+dt}$ (or $d\tilde{q}_t$) solves the following equation

$$m(\tilde{q}_t) = -cdt + e^{-\rho dt}\left[\zeta_t(0)m(0) + \zeta_t(\tilde{q}_{t+dt})m(\tilde{q}_{t+dt}) + \zeta_t(1)m(1)\right] \tag{38}$$

Denote $\underline{q} = \frac{\lambda\psi - c}{\lambda\psi+\rho}$, one could then solve for $d\tilde{q}_t$ as

$$d\tilde{q}_t = (\rho + \lambda\psi)\tilde{q}_t\left(\tilde{q}_t - \underline{q}\right)dt \tag{39}$$

Bayesian plausibility implies that

$$\begin{cases} \zeta_t(0) + \zeta_t(\tilde{q}_{t+dt}) + \zeta_t(1) = 1 \\ \zeta_t(0)0 + \zeta_t(\tilde{q}_{t+dt})\tilde{q}_{t+dt} + \zeta_1(1)1 = \tilde{q}_t \end{cases} \tag{40}$$

Combining Equation (37), (38) and (40) then implies Equation (39).

Furthermore, At first, we have the following equation

$$\frac{1}{\underline{q}}\left(\frac{d\tilde{q}_t}{\tilde{q}_t - \underline{q}} - \frac{d\tilde{q}_t}{\tilde{q}_t}\right) = \frac{d\tilde{q}_t}{(\tilde{q}_t - \underline{q})\tilde{q}_t} = (\lambda\psi + \rho)dt$$

As for the LHS of the equation above, one then have that

$$\frac{1}{\underline{q}}\left(\frac{d\tilde{q}_t}{\tilde{q}_t - \underline{q}} - \frac{d\tilde{q}_t}{\tilde{q}_t}\right) = \frac{1}{\underline{q}}[d\ln(\tilde{q}_t - \underline{q}) - d\ln\tilde{q}_t] = \frac{1}{\underline{q}}d\ln\left(\frac{\tilde{q}_t - \underline{q}}{\tilde{q}_t}\right).$$

For any $q_0 \in \left[\frac{1}{2}, 1\right] \cap \left(\underline{q}, 1\right]$, one, through combine these two equations above, then have that

$$\frac{1}{\underline{q}}\left[\ln\left(\frac{\tilde{q}_t - \underline{q}}{\tilde{q}_t}\right) - \ln\left(\frac{q_0 - \underline{q}}{q_0}\right)\right] = (\lambda\psi + \rho)t \Rightarrow \ln\left(\frac{\tilde{q}_t - \underline{q}}{\tilde{q}_t}\frac{q_0}{q_0 - \underline{q}}\right) = (\lambda\psi + \rho)t$$
$$\Rightarrow \frac{\tilde{q}_t - \underline{q}}{\tilde{q}_t}\frac{q_0}{q_0 - \underline{q}} = e^{\underline{q}(\lambda\psi+\rho)t} \Rightarrow \frac{\tilde{q}_t - \underline{q}}{\tilde{q}_t} = \frac{q_0 - \underline{q}}{q_0}e^{\underline{q}(\lambda\psi+\rho)t} \Rightarrow 1 - \frac{\underline{q}}{\tilde{q}_t} = \frac{q_0 - \underline{q}}{q_0}e^{\underline{q}(\lambda\psi+\rho)t}$$
$$\Rightarrow \frac{\underline{q}}{\tilde{q}_t} = 1 - \frac{q_0 - \underline{q}}{q_0}e^{\underline{q}(\lambda\psi+\rho)t} \Rightarrow \tilde{q}_t = \frac{\underline{q}}{1 - \frac{q_0 - \underline{q}}{q_0}e^{\underline{q}(\lambda\psi+\rho)t}} = \frac{\underline{q}}{1 - \frac{q_0 - \underline{q}}{q_0}e^{(\lambda\psi - c)t}}$$

C.14. **Properties of function $\tilde{\eta}(\cdot)$ and $\tilde{\theta}(\cdot)$.** Let me define function $\tilde{\chi}(t;\tilde{t})$ as the difference between the margin belief process $\widetilde{\mathcal{Q}}(\tilde{p}_{\tilde{t}},\tilde{t})$ and the research progress process $\widetilde{P}$ at time t. Or mathematically, it is defined as

$$\tilde{\chi}(t;\tilde{t}) := \tilde{q}_t(\tilde{p}_{\tilde{t}},\tilde{t}) - \tilde{p}_t.$$

Denote $\underline{s} := \max\left\{T(\frac{1}{2};\widetilde{P}),0\right\}$. In the first step, we will show that $\tilde{\chi}(t;\tilde{t})$ is convex in t within interval $t \in \left[\underline{s}, T\left(1;\widetilde{\mathcal{Q}}(\tilde{p}_{\tilde{t}},\tilde{t})\right)\right]$, for any $\tilde{t} \in \mathbf{T}$. Specifically, we have that

$$\begin{aligned}\frac{\partial^2\tilde{\chi}(t;\tilde{t})}{\partial t^2} &= \frac{\partial\frac{\partial\tilde{q}_t(\tilde{p}_{\tilde{t}},\tilde{t})}{\partial t}}{\partial t} - \frac{\partial\frac{\partial\tilde{p}_t}{\partial t}}{\partial t} = \frac{\partial\left[(\rho+\lambda\psi)\tilde{q}_t(\tilde{p}_{\tilde{t}},\tilde{t})\left(\tilde{q}_t(\tilde{p}_{\tilde{t}},\tilde{t})-\underline{t}\right)\right]}{\partial t} - \frac{\partial\left[\lambda(1-2\psi)\tilde{p}_t\left(1-\tilde{p}_t\right)\right]}{\partial t} \\ &= (\rho+\lambda\psi)\left(2\tilde{q}_t(\tilde{p}_{\tilde{t}},\tilde{t})-\underline{q}\right)\frac{\partial\tilde{q}_t(\tilde{p}_{\tilde{t}},\tilde{t})}{\partial t} + \lambda(1-2\psi)\left(2\tilde{p}_t-1\right)\frac{\partial\tilde{p}_t}{\partial t} > 0\end{aligned}$$

Moreover, one also have that

$$\begin{aligned}\frac{\partial\chi(\tilde{t},\tilde{t})}{\partial t} &= (\rho+\lambda\psi)\tilde{q}_{\tilde{t}}(\tilde{p}_{\tilde{t}},\tilde{t})\left(\tilde{q}_{\tilde{t}}(\tilde{p}_{\tilde{t}},\tilde{t})-\underline{t}\right) - \lambda(1-2\psi)\tilde{p}_{\tilde{t}}(1-\tilde{p}_{\tilde{t}}) = (\rho+\lambda\psi)\tilde{p}_{\tilde{t}}(\tilde{p}_{\tilde{t}}-\underline{t}) - \lambda(1-2\psi)\tilde{p}_{\tilde{t}}(1-\tilde{p}_{\tilde{t}}) \\ &= \tilde{p}_{\tilde{t}}\left[(\rho+\lambda\psi)(\tilde{p}_{\tilde{t}}-\underline{t}) - \lambda(1-2\psi)(1-\tilde{p}_{\tilde{t}})\right] = [\lambda(1-\psi)+\rho]\tilde{p}_{\tilde{t}}\left(\tilde{p}_{\tilde{t}}-\tilde{q}^P\right) \le 0,\end{aligned}$$

where equality holds if and only if $\tilde{p}_{\tilde{t}} = \tilde{q}^P$. Therefore, for any $\tilde{t} \in \mathbf{T}$, $\tilde{\chi}(t;\tilde{t}) = 0$ has exactly two solutions if $\tilde{t} < T(\tilde{q}^P;\widetilde{\mathcal{P}})$ and has a single solution if $\tilde{t} = T(\tilde{q}^P;\widetilde{\mathcal{P}})$. Therefore, both the function $\tilde{\eta}(\cdot)$ and $\tilde{\theta}(\cdot)$.

For any $p, p' \in \left[\tilde{q}^P,\tilde{q}^R\right]$ such that $p < p'$, one then have that $\tilde{q}_{T(p;\widetilde{\mathcal{P}})}(p',T(p';\widetilde{\mathcal{P}})) < \tilde{p}_{T(p;\widetilde{\mathcal{P}})} = p = \tilde{q}_{T(p;\widetilde{\mathcal{P}})}(p,T(p;\widetilde{\mathcal{P}}))$, which then implies that

$$\tilde{\eta}(p') = \tilde{q}_{T(\tilde{q}^P;\widetilde{\mathcal{P}})}(p',T(p';\widetilde{\mathcal{P}})) < \tilde{q}_{T(\tilde{q}^P;\widetilde{\mathcal{P}})}(p,T(p;\widetilde{\mathcal{P}})) = \tilde{\eta}(p).$$

Moreover, $\tilde{\theta}(\mu)$ is decreasing in μ.

C.15. **The belief martingale induce by the optimal disclosure policy.** In this subsection, we characterize the belief martingale γ^*, induced by the optimal disclosure policy proposed in Proposition 4.

For any $t \in [0,\tilde{t}^*)$, then $\gamma_t^*(\cdot)$ is given as

$$\gamma_t^*(p) = \begin{cases} p_0\left[\frac{1-p_0}{p_0} - \frac{1-\tilde{p}_t}{\tilde{p}_t}e^{-\lambda\psi t}\right] & \text{if } p=0 \\ \frac{p_0}{p}e^{-\lambda\psi t} & \text{if } p=\tilde{p}_t \\ p_0 - p_0 e^{-\lambda\psi t} & \text{if } p=1 \\ 0 & \text{otherwise}\end{cases}$$

For any $t \in \left[\tilde{t}^*, \tilde{\theta}(\bar{\mu}(\tilde{t}^*))\right)$, then $\gamma_t^*(\cdot)$ is given as

$$\gamma_t^*(p) = \begin{cases} p_0 \left[\frac{1-p_0}{p_0} - \frac{1-\tilde{q}_t(\tilde{p}_{\tilde{t}^*}, \tilde{t}^*)}{\tilde{q}_t(\tilde{p}_{\tilde{t}^*}, \tilde{t}^*)} e^{-\lambda\psi t}\right] & \text{if } p = 0 \\ \frac{p_0}{p} e^{-\lambda\psi t} & \text{if } p = \tilde{q}_t(\tilde{p}_{\tilde{t}^*}, \tilde{t}^*) \\ p_0 - p_0 e^{-\lambda\psi t} & \text{if } p = 1 \\ 0 & \text{otherwise} \end{cases}$$

Denote belief $\beta_t = \min\left\{\tilde{q}_t\left(\underline{\mu}(\tilde{t}^*), T(\tilde{q}^P; \widetilde{\mathcal{P}})\right), \tilde{p}_{\tilde{\theta}(\underline{\mu}(\tilde{t}^*))}\right\}$ and probability $\upsilon = \frac{\tilde{q}_{\tilde{\theta}(\underline{\mu}(\tilde{t}^*))}\left(\bar{\mu}(\tilde{t}^*), T(\tilde{q}^P; \widetilde{\mathcal{P}})\right) - \tilde{q}_{\tilde{\theta}(\underline{\mu}(\tilde{t}^*))}(\tilde{p}_{\tilde{t}^*}, \tilde{t}^*)}{\tilde{q}_{\tilde{\theta}(\underline{\mu}(\tilde{t}^*))}\left(\bar{\mu}(\tilde{t}^*), T(\tilde{q}^P; \widetilde{\mathcal{P}})\right) - \tilde{p}_{\tilde{\theta}(\underline{\mu}(\tilde{t}^*))}}$. For any $t \geq \tilde{\theta}(\bar{\mu}(\tilde{t}^*))$, $\gamma_t^*(\cdot)$ is given by

$$\gamma_t^*(p) = \begin{cases} \frac{p_0}{\tilde{q}_{\tilde{\theta}(\bar{\mu}(\tilde{t}^*))}(\tilde{p}_{\tilde{t}^*}, \tilde{t}^*)} \upsilon e^{-\lambda\psi\tilde{\theta}(\bar{\mu}(\tilde{t}^*))} & \text{if } p = \tilde{p}_{\tilde{\theta}(\bar{\mu}(\tilde{t}^*))} \\ \frac{p_0}{\tilde{q}_{\tilde{\theta}(\bar{\mu}(\tilde{t}^*))}(\tilde{p}_{\tilde{t}^*}, \tilde{t}^*)} (1-\upsilon) \frac{\tilde{q}_{\tilde{\theta}(\bar{\mu}(\tilde{t}^*))}\left(\underline{\mu}(\tilde{t}^*), T(\tilde{q}^P; \tilde{\mathcal{P}})\right)}{\beta_t} e^{-\lambda\psi \min\left\{t, \tilde{\theta}(\underline{\mu}(\tilde{t}^*))\right\}} & \text{if } p = \beta_t \\ p_0\left(1 - e^{-\lambda\psi\tilde{\theta}(\bar{\mu}(\tilde{t}^*))}\right) + & \\ \frac{p_0}{\tilde{q}_{\tilde{\theta}(\bar{\mu}(\tilde{t}^*))}(\tilde{p}_{\tilde{t}^*}, \tilde{t}^*)} (1-\upsilon) \tilde{q}_{\tilde{\theta}(\bar{\mu}(\tilde{t}^*))}\left(\underline{\mu}(\tilde{t}^*), T(\tilde{q}^P; \tilde{\mathcal{P}})\right)\left(e^{-\lambda\psi\tilde{\theta}(\bar{\mu}(\tilde{t}^*))} - e^{-\lambda\psi \min\left\{t, \tilde{\theta}(\underline{\mu}(\tilde{t}^*))\right\}}\right) & \text{if } p = 1 \\ 1 - \gamma_t^*(\tilde{p}_{\tilde{\theta}(\bar{\mu}(\tilde{t}^*))}) - \gamma_t^*(\beta_t) - \gamma_t^*(1) & \text{if } p = 0 \\ 0 & \text{otherwise} \end{cases}$$

Proposition 4 can also be rewritten as the following proposition.

Proposition 5. *The optimal disclosure policy* $(\tilde{\pi}^*, \tilde{M}^*)$ *induces a belief martingale* $\boldsymbol{\gamma}^*$.

C.16. **Proof of Proposition 4.** Our proof only focuses on the case where $p_0 \geq \frac{1}{2}$ and a similar proof can be extended to cover the case where $p_0 \in [\underline{p}_0, \frac{1}{2})$. One can easily show that both responsiveness with restricted interval $[\tilde{q}^P, \tilde{q}^R]$ and directness still hold in this scenario of general information structure. For any direct and responsive disclosure policy (π, M), let me denote the induced belief martingale as $\boldsymbol{\tau}^\pi$ and the stopping belief distribution as τ^π. Moreover, let me denote the weight accumulation speed $\tilde{\lambda}^\pi$ as

$$\tilde{\lambda}^\pi(t) = \tau_t^\pi(1) + \tau_t^\pi(0).$$

For any stopping belief distribution $\tau \in \Delta\left(0, 1 \cup \left[\tilde{q}^P, \tilde{q}^R\right]\right)$, we then define a belief martingale $\tilde{\boldsymbol{\tau}}^*(\cdot; \tau)$ with the fastest weight accumulation speed $\tilde{\lambda}^*(\cdot; \tau)$, which induces the stopping belief distribution τ. For any belief $p \in \left[\tilde{q}^P, \tilde{q}^R\right]$, let me denote $\tilde{\gamma}(p)$ as follows,

$$\tilde{\gamma}(p) \frac{\tilde{\eta}(p)}{\tilde{q}^P} \frac{\tilde{q}^P}{p} e^{-\lambda\psi\left[T(\tilde{p}; \widetilde{\mathcal{P}}) - T(\tilde{q}^P; \widetilde{\mathcal{P}})\right]} = \tau(p) \implies \tilde{\gamma}(p) = \tau(p) \frac{p}{\tilde{\eta}(p)} e^{\lambda\psi\left[T(\tilde{p}; \widetilde{\mathcal{P}}) - T(\tilde{q}^P; \widetilde{\mathcal{P}})\right]}$$

Basically, $\tilde{\gamma}(p)$ is defined as the probability on belief $\tilde{\eta}(p)$ such that the induced belief martingale reaches belief p with probability $\tau(p)$. Let me normalize $\gamma(p)$ by defining $\tilde{\gamma}'(p) = \frac{\tilde{\gamma}(p)}{\int_{p\in[\tilde{q}^P,\tilde{q}^R]}\tilde{\gamma}(p)dp}$. Then denote $\eta'(\tau)$ as

$$\tilde{\eta}'(\tau) = \int_{p\in[\tilde{q}^P,\tilde{q}^R]} \tilde{\eta}(p)\tilde{\gamma}'(p)dp.$$

Next, for any τ such that there exists some $\tilde{t} \in \mathbf{T}$ such that $\tilde{q}_{T(\tilde{q}^P;\widetilde{\mathcal{P}})}(\tilde{p}_{\tilde{t}}, \tilde{t}) \geq \tilde{\eta}'(\tau)$, we define $\tilde{\varphi}(\tau)$ implicitly as the solution to the following equation

$$\tilde{q}_{T(\tilde{q}^P;\widetilde{\mathcal{P}})}(\tilde{p}_{\tilde{\varphi}(\tau)}, \tilde{\varphi}(\tau)) = \tilde{\eta}'(\tau).$$

Basically, $\tilde{\varphi}(\tau)$ is defined as the tranisition time with which the reporting process reaches the belief $\tilde{\eta}'(\tau)$ at time $T(\tilde{q}^P;\widetilde{\mathcal{P}})$.

Finally, let me define the belief martingale $\boldsymbol{\tau}^*(\cdot;\tau)$. For any $t \in [0, \tilde{\varphi}(\tau))$, the belief martingale $\tilde{\boldsymbol{\tau}}^*(\cdot;\tau)$ is consistent with $\tilde{\boldsymbol{\tau}}^F$. For any $t \in \left[\tilde{\varphi}(\tau), T(\tilde{q}^P;\widetilde{\mathcal{P}})\right)$, the belief martingale $\tilde{\boldsymbol{\tau}}_t^*(\cdot;\tau)$ is defined as

$$\tilde{\boldsymbol{\tau}}_t^*(q;\tau) = \begin{cases} (1-p_0) - e^{-\lambda t\frac{1-\tilde{q}_t\left(\tilde{p}_{\tilde{\varphi}(\tau)},\tilde{\varphi}(\tau)\right)}{\tilde{q}_t\left(\tilde{p}_{\tilde{\varphi}(\tau)},\tilde{\varphi}(\tau)\right)}} & \text{if } q = 0 \\ \frac{p_0}{\tilde{q}_t\left(\tilde{p}_{\tilde{\varphi}(\tau)},\tilde{\varphi}(\tau)\right)} e^{-\lambda\psi t} & \text{if } q = \tilde{q}_t\left(\tilde{p}_{\tilde{\varphi}(\tau)}, \tilde{\varphi}(\tau)\right) \\ p_0\left(1 - e^{-\lambda\psi t}\right) & \text{if } q = 1 \\ 0 & \text{otherwise} \end{cases}$$

For any $t \geq T(\tilde{q}^P;\widetilde{\mathcal{P}})$, the belief martingale $\tilde{\boldsymbol{\tau}}_t^*(\cdot;\tau)$ is defined as follows

$$\tilde{\boldsymbol{\tau}}_t^*(q;\tau) = \begin{cases} \tau(q) & \text{if } q \in [\tilde{q}^P, \tilde{p}_t] \\ \frac{\tilde{\gamma}(p)\tilde{\eta}(p)}{\tilde{q}_t\left(\tilde{\eta}(p),T(\tilde{q}^P,\widetilde{\mathcal{P}})\right)} e^{-\lambda\psi\left(t-T(\tilde{q}^P,\widetilde{\mathcal{P}})\right)} & \text{if } q = \tilde{q}_t\left(\tilde{\eta}(p), T(\tilde{q}^P,\widetilde{\mathcal{P}})\right), \\ & \forall p \in \left(\tilde{p}_t, \tilde{q}^R\right] \\ p_0\left(1 - e^{-\lambda\psi T(\tilde{q}^P,\widetilde{\mathcal{P}})}\right) + \int_{p\in[\tilde{q}^P,\tilde{q}^R]} \frac{\tilde{\gamma}(p)\tilde{\eta}(p)}{\min\left\{p,\tilde{q}_t\left(\tilde{\eta}(p),T(\tilde{q}^P,\widetilde{\mathcal{P}})\right)\right\}} e^{-\lambda\psi\left[\min\left\{t,T(p;\widetilde{\mathcal{P}})\right\}-T(\tilde{q}^P,\widetilde{\mathcal{P}})\right]} & \text{if } p = 1 \\ 1 - \int_{q\in[\tilde{q}^P,\tilde{p}_t]} \tilde{\boldsymbol{\tau}}_t^*(q;\tau)dq - \int_{p\in\left(\tilde{p}_t,\tilde{q}^R\right]} \tilde{\boldsymbol{\tau}}_t^*\left(\tilde{q}_t\left(\tilde{\eta}(p), T(\tilde{q}^P,\widetilde{\mathcal{P}})\right);\tau\right)dp - \tilde{\boldsymbol{\tau}}_t^*(1;\tau) & \text{if } p = 0 \\ 0 & \text{otherwise} \end{cases}$$

Then let me define the corresponding weight accumulation speed $\tilde{\lambda}^*(t;\tau)$ as

$$\tilde{\lambda}^*(t;\tau) = \tilde{\boldsymbol{\tau}}_t^*(0;\tau) + \tilde{\boldsymbol{\tau}}_t^*(1;\tau).$$

Our next lemma shows that the weight accumulation speed $\tilde{\lambda}^*(t;\tau)$ is the fastest among all direct and responsive disclosure policies (π, M) which share the same stopping belief distribution, i.e., $\tau^\pi = \tau$.

Lemma 15. *For any direct and responsive (with restricted range) disclosure policy* (π, M)*, which induces weight accumulation speed* $\tilde{\lambda}^\pi$ *and stopping belief distribution* τ^π *such that*

$$\tilde{q}_{T(\tilde{q}^P;\widetilde{\mathcal{P}})}(\tilde{p}_{\tilde{t}}, \tilde{t}) \geq \tilde{\eta}'(\tau^\pi), \ \exists \tilde{t} \in \mathbf{T}$$

we then have that

$$\tilde{\lambda}^*(t; \tau^\pi) \geq \tilde{\lambda}^\pi(t), \forall t \geq 0.$$

Proof. Given any direct and responsive disclosure policy (π, M), the agent's belief $\alpha(t;\pi)$, contingent on event where he still funds the expert at time t, is given as

$$\alpha(t;\pi) = \frac{p_0 - \tilde{\boldsymbol{\tau}}_t^\pi(1) - \int_{p\in[\tilde{q}^P,\tilde{p}_t]} \tau^\pi(p)pdp}{1 - \tilde{\lambda}^\pi(t) - \int_{p\in[\tilde{q}^P,\tilde{p}_t]} \tau^\pi(p)dp}.$$

Moreover, under this disclosure policy, the expert and agent's payoff are given by

$$\begin{cases} \tilde{V}(\tilde{\lambda}^\pi, \tau^\pi) = \rho \int_{t=0}^{\infty} e^{-\rho t}\tilde{\lambda}^\pi(t)dt + \int_{p\in[\tilde{q}^P,\tilde{q}^R]} e^{-\rho T(p;\widetilde{\mathcal{P}})} m(p)dp \\ \tilde{U}(\tilde{\lambda}^\pi, \tau^\pi) = \tilde{V}(\tilde{\lambda}^\pi, \tau^\pi) - c\int_{t=0}^{T(\tilde{q}^R;\widetilde{\mathcal{P}})} \left[1 - \tilde{\lambda}^\pi(t) - \int_{p\in[\tilde{q}^P,\tilde{p}_t]} \tau^\pi(p)dp\right] \end{cases}$$

At time t, the expert and agent's continuation payoff, contingent on the event that the agent still funds the research, are given by

$$\begin{cases} \tilde{V}_t(\tilde{\lambda}^\pi, \tau^\pi) = \frac{\rho\int_t^\infty e^{-\rho(s-t)}[\tilde{\lambda}^\pi(s) - \tilde{\lambda}^\pi(t)]ds + \int_{p\in(\tilde{p}_t,\tilde{q}^P]} e^{-\rho[T(p;\widetilde{\mathcal{P}})-t]} m(p)\tau^\pi(p)dp}{1 - \tilde{\lambda}^\pi(t) - \int_{p\in[\tilde{q}^P,\tilde{p}_t]}\tau^\pi(p)dp} \\ \tilde{U}_t(\lambda^\pi, \tau^\pi) = \frac{\rho\int_t^\infty e^{-\rho(s-t)}[\tilde{\lambda}^\pi(s) - \tilde{\lambda}^\pi(t)]ds + \int_{p\in(\tilde{p}_t,\tilde{q}^P]} e^{-\rho[T(p;\widetilde{\mathcal{P}})-t]} m(p)\tau^\pi(p)dp - c\int_t^{T(\tilde{q}^R,\widetilde{\mathcal{P}})} e^{-\rho(s-t)}\left[1 - \tilde{\lambda}^\pi(s) - \int_{p\in[\tilde{q}^P,\tilde{p}_s]}\tau^\pi(p)dp\right]ds}{1 - \lambda^\pi(t) - \int_{p\in[\tilde{q}^P,\tilde{p}_t]}\tau^\pi(p)dp} \end{cases}$$

Please note that the combination of weight accumulation speed $\tilde{\lambda}^\pi$ and stopping belief distribution τ^π is sufficient statistic of any direct and responsive disclosure policy (π, M) in determining the payoffs.

For any pair of $(\tilde{\lambda}, \tau)$ which can be induced by some direct and responsive disclosure policy (π, M), let me define ϑ' as follows

$$\vartheta' := \inf\left\{t : \tilde{\lambda}(t') \leq \tilde{\lambda}^*(t', \tau), \ \forall t' \geq t\right\}.$$

Since $\tilde{\lambda}^\pi(T(\tilde{q}^R;\widetilde{\mathcal{P}})) = \tau(0) + \tau(1)$, one then have that $\vartheta' \leq T(\tilde{q}^R;\widetilde{\mathcal{P}})$. Moreover if $\vartheta' \leq \tilde{\varphi}(\tau)$, then $\vartheta' = 0$. This is due to the fact that $\tilde{\lambda}^*(\cdot;\tau)$ is consistent with that under full information disclosure, which reaches the upper-bound of weight accumulation speed. Our remaining task is to show that $\vartheta' = 0$.

Suppose the argument is not true, then $\vartheta' > \tilde{\varphi}(\tau)$. In the first step, we are going to show that

$$\tilde{\lambda}^*(\vartheta';\tau) \geq \tilde{\lambda}(\vartheta').$$

Suppose on the contrary, it is not true. Then at first, we have that $\tilde{\lambda}^*(\vartheta';\tau) < \tau(0)+\tau(1)$. Then, let us, at first, consider the case in which $\vartheta' \geq T(\tilde{q}^P;\widetilde{\mathcal{P}})$. The continuation payoff of the agent at time ϑ' is given by

$$\begin{aligned}
\tilde{U}_{\vartheta'}(\tilde{\lambda},\tau) &= \frac{\begin{bmatrix} \rho\int_{\vartheta'}^{\infty} e^{-\rho(s-\vartheta')}[\tilde{\lambda}(s)-\tilde{\lambda}(\vartheta')]ds + \int_{p\in(\tilde{p}_t,\tilde{q}^P]} e^{-\rho\left[T(p;\widetilde{\mathcal{P}})-\vartheta'\right]}m(p)\tau(p)dp \\ -c\int_{\vartheta'}^{T(\tilde{q}^R,\widetilde{\mathcal{P}})} e^{-\rho(s-\vartheta')}\left[1-\tilde{\lambda}(s)-\int_{p\in[\tilde{q}^P,\tilde{p}_s]}\tau(p)dp\right]ds \end{bmatrix}}{1-\tilde{\lambda}(\vartheta')-\int_{p\in[\tilde{q}^P,\tilde{p}_{\vartheta'}]}\tau(p)dp} \\
&\leq \frac{\begin{bmatrix} \rho\int_{\vartheta'}^{\infty} e^{-\rho(s-\vartheta')}[\tilde{\lambda}^*(t;\tau)-\tilde{\lambda}(\vartheta')]ds + \int_{p\in(\tilde{p}_t,\tilde{q}^P]} e^{-\rho\left[T(p;\widetilde{\mathcal{P}})-\vartheta'\right]}m(p)\tau(p)dp \\ -c\int_{\vartheta'}^{T(\tilde{q}^R,\widetilde{\mathcal{P}})} e^{-\rho(s-\vartheta')}\left[1-\tilde{\lambda}^*(\vartheta';\tau)-\int_{p\in[\tilde{q}^P,\tilde{p}_s]}\tau(p)dp\right]ds \end{bmatrix}}{1-\tilde{\lambda}(\vartheta')-\int_{p\in[\tilde{q}^P,\tilde{p}_{\vartheta'}]}\tau(p)dp} \\
&< \frac{\begin{bmatrix} \rho\int_{\vartheta'}^{\infty} e^{-\rho(s-\vartheta')}[\tilde{\lambda}^*(t;\tau)-\tilde{\lambda}^*(\vartheta';\tau)]ds + \int_{p\in(\tilde{p}_t,\tilde{q}^P]} e^{-\rho\left[T(p;\widetilde{\mathcal{P}})-\vartheta'\right]}m(p)\tau(p)dp \\ -c\int_{\vartheta'}^{T(\tilde{q}^R,\widetilde{\mathcal{P}})} e^{-\rho(s-\vartheta')}\left[1-\tilde{\lambda}^*(\vartheta';\tau)-\int_{p\in[\tilde{q}^P,\tilde{p}_s]}\tau(p)dp\right]ds \end{bmatrix}}{1-\tilde{\lambda}(\vartheta')-\int_{p\in[\tilde{q}^P,\tilde{p}_{\vartheta'}]}\tau(p)dp} \\
&= \frac{\int_{p\in(\tilde{p}_{\vartheta'},\tilde{q}^R]}\frac{\tilde{\gamma}(p)\tilde{\eta}(p)}{\tilde{q}_{\vartheta'}\left(\tilde{\eta}(p),T(\tilde{q}^P;\widetilde{\mathcal{P}})\right)}e^{-\lambda\psi[\vartheta'-T(\tilde{q}^P;\widetilde{\mathcal{P}})]}m\left(\tilde{q}_{\vartheta'}\left(\tilde{\eta}(p),T(\tilde{q}^P;\widetilde{\mathcal{P}})\right)\right)dp}{1-\tilde{\lambda}(\vartheta')-\int_{p\in[\tilde{q}^P,\tilde{p}_{\vartheta'}]}\tau(p)dp} \\
&= \frac{\int_{p\in(\tilde{p}_{\vartheta'},\tilde{q}^R]}\frac{\tilde{\gamma}(p)\tilde{\eta}(p)}{\tilde{q}_{\vartheta'}\left(\tilde{\eta}(p),T(\tilde{q}^P;\widetilde{\mathcal{P}})\right)}e^{-\lambda\psi[\vartheta'-T(\tilde{q}^P;\widetilde{\mathcal{P}})]}\tilde{q}_{\vartheta'}\left(\tilde{\eta}(p),T(\tilde{q}^P;\widetilde{\mathcal{P}})\right)dp}{1-\tilde{\lambda}(\vartheta')-\int_{p\in[\tilde{q}^P,\tilde{p}_{\vartheta'}]}\tau(p)dp} \\
&= \frac{p_0-\tilde{\boldsymbol{\tau}}^*_{\vartheta'}(1;\tau)-\int_{p\in[\tilde{q}^P,\tilde{p}_{\vartheta'}]}\tau(p)pdp}{1-\tilde{\lambda}(\vartheta')-\int_{p\in[\tilde{q}^P,\tilde{p}_{\vartheta'}]}\tau(p)dp} \leq \frac{p_0-\tilde{\boldsymbol{\tau}}^{\pi}_{\vartheta'}(1)-\int_{p\in[\tilde{q}^P,\tilde{p}_{\vartheta'}]}\tau(p)pdp}{1-\tilde{\lambda}(\vartheta')-\int_{p\in[\tilde{q}^P,\tilde{p}_{\vartheta'}]}\tau(p)dp} = m(\alpha(t;\pi))
\end{aligned}$$

The first inequality follows from the definition of ϑ', which implies that $\tilde{\lambda}(t) \leq \lambda^*(t;\tau)$ for any $t > \vartheta'$. The second inequality (the only strict inequality) follows from the assumption that $\tilde{\lambda}(\vartheta') > \tilde{\lambda}^*(\vartheta';\tau)$. The second equality follows from the following three facts. At first, the numerator of the term at the third line of the equation above is exactly the unconditional expected payoff to the agent if he keeps funding research until ϑ'. Secondly, at time t, the belief martingale $\tilde{\boldsymbol{\tau}}^*(\cdot;\tau)$ assigns probability $\frac{\tilde{\gamma}(p)\tilde{\eta}(p)}{\tilde{q}_{\vartheta'}\left(\tilde{\eta}(p),T(\tilde{q}^P;\widetilde{\mathcal{P}})\right)}e^{-\lambda\psi[\vartheta'-T(\tilde{q}^P;\widetilde{\mathcal{P}})]}$

on belief $\tilde{q}_{\vartheta'}\left(\tilde{\eta}(p), T(\tilde{q}^P;\widetilde{\mathcal{P}})\right)$. Finally, by the definition of margin belief process, the agent's continuation payoff upon receiving message with belief $\tilde{q}_{\vartheta'}\left(\tilde{\eta}(p), T(\tilde{q}^P;\widetilde{\mathcal{P}})\right)$ is given by $m\left(\tilde{q}_{\vartheta'}\left(\tilde{\eta}(p), T(\tilde{q}^P;\widetilde{\mathcal{P}})\right)\right)$. The forth equality follows from Bayesian plausibility constraint. The last inequality follows from the fact that $\tilde{\boldsymbol{\tau}}^*_{\vartheta'}(1;\tau)$ is always consistent with the one under full information disclosure, which reaches the upper bound of probability on belief 1 given stopping belief distribution τ. The last equality follows from the definition of $\alpha(t;\pi)$. Since the agent's continuation payoff at time ϑ', if he chooses to fund research, is strictly lower than the one if he chooses to cut research funding and make decision immediately, it is not incentive compatible for the dynamic disclosure policy (π, M) to implement the stopping belief distribution τ, which then leads to contradiction.

Secondly, let us consider the case where $\vartheta' < T(\tilde{q}^P;\widetilde{\mathcal{P}})$. The proof of the contradiction is similar and therefore ignored. All in all, it can not be true that $\tilde{\lambda}(\vartheta') > \tilde{\lambda}^*(\vartheta';\tau)$ otherwise the stopping belief distribution τ can not be implemented.

Finally, let us consider the case in which both $\tilde{\lambda}(\vartheta') \leq \tilde{\lambda}^*(\vartheta';\tau)$ and $\vartheta' > 0$. The proof of contraction is exactly the same as the one in the proof of Lemma 13 and we therefore ignore it. Accordingly, we conclude that $\vartheta' = 0$. □

We then show that for any direct and responsive dynamic disclosure policy (π, M), there always exists some $\tilde{t} \in \mathbf{T}$ such that $\tilde{q}_{T(\tilde{q}^P;\widetilde{\mathcal{P}})}(\tilde{p}_{\tilde{t}}, \tilde{t}) = \tilde{\eta}'(\tau^\pi)$. The proof is the similar to that of Corollary 4 and therefore ignored.

Finally, given any transition time $\tilde{t} \in \mathbf{T}$, the optimal stopping belief distribution τ is determined through concavifying the function $\tilde{V}(\cdot)$ at belief $\tilde{q}_{T(\tilde{q}^P;\widetilde{\mathcal{P}})}(\tilde{p}_{\tilde{t}}, \tilde{t})$ within interval $[\tilde{q}^P, \tilde{\eta}(\tilde{q}^R)]$.

WEI ZHAO, HEC PARIS AND GREGHEC-CNRS, 1 RUE DE LA LIBÉRATION, 78351 JOUY-EN-JOSAS, FRANCE

Email address: wei.zhao1(at)hec.fr