Strategic Disclosure in Research Races (Preliminary Version)

Kalyan Chatterjee *　　Kaustav Das†　　Miaomiao Dong‡

November 30, 2022

Abstract

We study a research race between two players. Each works on an identical two-step project, in which step 1 needs to be accomplished before working on step 2. Each stage is completed with a discovery. Once a discovery is made, a player decides whether and when to disclose it. Disclosure of an intermediate discovery gives an immediate reward to the player, but it also allows the opponent to copy it and compete for a final reward from the final discovery. We show that the effect of increasing the final reward on the disclosure time of the intermediate and final findings is U-shaped: it speeds up disclosure only when the final reward is not high.

Keywords: R&D competition, Strategic Disclosure

JEL Codes: D83, L13, O31

This paper seeks to analyse the nature of competition in academic research races. There are three aspects of academic we wish to emphasise. First, the result of the research is a public good. Once published or made public, everyone who reads it knows about it and cannot be excluded from any potential benefits. Second, the research is cumulative, i.e. later researchers can “stand on the shoulders of giants”, as Newton reportedly said. Thus, there could well be benefits from reading a published piece of academic research. Third, the benefit is usually not directly in terms of money. The first to publish gets credit for the discovery, which might be small or large based on its perceived importance at the time of publication. However, publication does not give the initial discoverer any share in the credit accruing to someone else who uses the discovery to make another significant one.[1]

*Penn State University.

†University of Leicester.

‡Penn State University.

[1] This excludes events such as rediscovery-a paper might have languished without becoming known until someone else, perhaps of greater fame, rediscovers it and gives credit to the original discoverer. We do not address such episodes here.

There are two basic questions that arise about these research races. First, the incentive to exert effort. This clearly depends on the rewards that might result but the issue is complicated by the phenomenon of others using the discoverer's publication for subsequent discovery. This leads to the second, important question. What is the incentive to disclose a discovery once made? This depends on the reward to announcing the discovery traded off against the potential losses arising from other researchers having access to the information. These questions are interlinked. Would providing access to potential competitors reduce the incentive to pursue the research in the first place? Surely there would be an element of free riding here, if someone else's discovery is available for one to use. This paper seeks to address these questions, in particular to analyse the effect of unrestricted access to early research on the incentive to expend research effort.

We now provide a brief description of our framework. Our model has two ex-ante identical players who compete to accomplish a final result, denoted as S. The discovery process involves two steps: to discover S, one has to make a first discovery, denoted as F. Each discovery is modelled as a Poisson arrival. The players simultaneously choose effort between 0 and 1 in continuous time. The effort level (linearly) affects the arrival rate of a discovery. A player enters step 2 if and only if she has discovered F. The flow cost per unit of effort is $c > 0$ in step 1 and 0 in step 2 (for simplicity), the same for both players.

Once a player has made the first discovery, she decides whether and when to disclose it by publication. Once it is disclosed, the other player, if he is still in step 1, can immediately copy it and discover F. The prize allocation is winner-take-all. The first player who discloses the intermediate result receives a reward F and the first player who discovers S receives a reward S (abusing notation for simplicity). If no disclosure has taken place, the first discoverer of S gets $F + S$.

An example we have in mind is that of researchers competing to be the first to prove Fermat's last theorem.[2]

In 1984, a path to the proof involving two steps was outlined. These steps were (i) Proving that the Taniyama-Shimura conjecture implies Fermat's Last Theorem and (ii) Proving the Taniyama-Shimura conjecture. Ribet accomplished the first step in 1986 and made his finding public. Andrew Wiles accomplished the second step in 1995. If some researcher had discovered this connection and wanted to prove Fermat's theorem would it have been optimal for him to disclose. Note also that there was no way in which any of the researchers who had worked on the first step could have laid claim to the credit from proving Fermat's theorem-there was no way to "patent" academic results.[3]

[2]Others have mentioned this example , which is not surprising given its importance, but we will highlight the features that are relevant to our model.

[3]Andrew Wiles, who (with Richard Taylor), obtained the long and complex proof of Fermat's last theorem,

We now describe the main results, confining ourselves to symmetric equilibria. (The equilibrium concept, a standard one, is discussed in the next section.)

First, we characterise circumstances under which in equilibrium the discoverer of the first stage discloses his results. In equilibrium, disclosure of the initial discovery can be immediate, or "mixed" over time or none at all.

In addition to the relative magnitudes of F and S, it is interesting to determine whether the cost of effort plays any role in disclosure. This relates to the effect of disclosure strategies on free riding incentives mentioned earlier. We find that if the cost is in an intermediate range, as the final reward increases, equilibrium disclosure goes up when the final reward is low and goes down when it is high. Thus, heavily rewarding the final result can backfire if a policy maker wants to bring forward the discovery time of the final result.

We now explain the intuition. The cost of disclosing immediately rather than later is that it empowers the opponent to discover S quicker, if the opponent has not discovered F yet. Disclosing immediately, however, is beneficial because it avoids discounting and the risk of being preempted by the opponent. These cost and benefit are affected by effort cost in step 1, the prize ratio $\frac{S}{F}$, and an endogenous parameter μ the probability that the opponent has discovered F:

- At a given point in time, the higher the probability that the opponent has discovered F, the higher is a player's disclosure incentive. The reason is that the more likely j has discovered F, the lower the cost of disclosure (lower chance of empowering the opponent), and the higher the risk of i being preempted.

- Higher effort cost reduces the incentive to disclose. The higher the cost, the lower the effort the opponent expends. Thus, at a given point in time, the probability that the opponent has succeeded is lower. Consequently, a player has a lower incentive to disclose.

- The effect of the final reward on disclosure is not monotone. On the one hand, the higher the final reward, the more costly it is to disclose, since by disclosure, a player essentially shares the final reward with the opponent. On the other hand, the higher the final reward, the higher the effort the opponent exerts, meaning that at a given point in time, the chance the opponent has discovered is higher. This, as we have discussed, increases a player's disclosure incentive. Thus, increasing the final reward has two opposing effects—an effort-enhancing effect and a disclosure-promoting effect.

using the 1984 framework and Ribet's theorem, actually kept the fact he was working on this problem secret in order not to attract competitors. (This aspect of the example is not present in our analysis, though in our model, choice of effort by individual researchers is unobservable.)

In general, when the effort cost and the final reward are intermediate, a player's effort is low, and consequently, the effort-enhancing effect dominates the disclosure-promoting effect. Otherwise, increasing the final reward promotes earlier disclosure.

Why does increasing the final reward have a U-shaped effect on the discovery time of the final result? As discussed above, increasing the final reward has two effects—more effort and (possibly) delayed disclosure. At low levels of S, the first effect dominates, and as a result, increasing S shortens the discovery time. At a high level of S, effort is already high; the second effect dominates and increasing S therefore lengthens the discovery time.

Our last result concerns whether mandatory transparency brings forward the discovery time of S. Mandatory transparency refers to a situation where firms are obliged to ensure complete spillover of intermediate findings. Thus, it can be interpreted as a variant of Research Joint Ventures (RJV) where participants agree to full disclosure of any intermediate discoveries among themselves, which can be a part of the contract setting up the RJV.

We find that mandatory transparency and increasing the final reward are policy complements; that is, mandatory transparency is good only if the final reward is sufficiently high. The intuition is that while mandatory transparency kills the inefficiency of delayed disclosure, it reduces players' first-step effort as they now can free-ride and wait for knowledge spillover from the opponent. When the final reward is high, the first effect is negligible as players' efforts are already high. In this case, mandatory disclosure is good. When the final reward is low, the first effect dominates, and introducing mandatory disclosure thus lengthens the discovery time.

Discussion. Note that there might be another reason for delaying disclosure, not to keep competitors from learning intermediate results but to try to convince the other player that the research is unlikely to end in success. Typically this would be done with a model with an unknown state of nature, where someone who made the first discovery would learn that the state of nature was "Good."[4] In order to isolate the effect of disclosure of the content of a paper, we have avoided introducing this other incentive in our model.

We now discuss how our framework differs from that of, say, Bhattacharya et al. (1992) or other papers dealing with patenting and licensing. In these papers, there is a basic research phase and once it is concluded, the discovery is licensed to secondary developers for a fee. The fee also gives an incentive to exert effort to the potential participants in basic research. We might illustrate the difference between our setup and the licensing problem by considering a successful application for a patent filed by Brams and Taylor for a fair division algorithm based on their academic paper (Brams and Taylor (1995)). The US patent, which

[4]For example in an experimentation model a.la. Keller et al. (2005)

the authors assigned to New York University, relates to uses of their work in actual fair division problems, not in their result being used to prove other academic theorems. One can foresee the difficulty in enforcing a patent on academic use. What is the value of such a patent? If someone infringes it, what would the loss be to a patent holder? Would he or she have a claim to the other party's lifetime earnings? But these would presumably be dependent on the entire career of the infringer, so a court would have to divide up the earnings of the infringer based on the volume and quality of his other publications. This does not seem a possible undertaking and we have not heard of any such cases. A patent on, say, a method for synthesising leukotrienes, could lead to payments from manufacturers of asthma drugs but surely not from someone who used the same techniques for another academic paper.

1 Related literature

Early attempt to analyze the issue of optimal disclosure of information in research races can be traced back to the work by Scotchmer and Green (1990). They study how stringency of the novelty requirements for the award of a patent affects the amount of technical information that is disclosed among firms. Song and Zhao (2021) analyse a model where two players compete to achieve a final success that involves two stages. Once the first stage is solved, the expected time to achieve the second stage success is finite. However, a-priori, it is not known with certainty whether the first stage success will arrive or not. A player after completing the first stage can either disclose or withhold the information about his discovery. Post disclosure, the game is similar to the model in the current paper. If a player becomes too pessimistic about the viability of the first stage success, then it can irreversibly exit the race. They describe a symmetric equilibrium which is characterised by two time thresholds. Each player discloses if the first stage success is achieved before the first threshold, withholds if the success occurs between the two thresholds, and exits the race if no success occurs until the second threshold. In their model, non-disclosure in equilibrium solely arises from the fact that in equilibrium the opponent will exit after not having success until a time point. Their setting abstracts from interior effort in the first stage, and continuous disclosure of the intermediate success. Further, their equilibrium exists only if the prior belief that the first stage is viable is below a certain threshold. Our model differs from their paper in two significant ways. First, we abstract from learning but demonstrate there are interesting considerations for strategic disclosure of intermediate success even in the absence of any uncertainty about the viability of the first stage. As such, we show that there can be non-disclosure in equilibrium without learning. Secondly, equilibrium disclosure has a public good characteristic. As such, it is important to analyse the interaction between efforts put by players in the first stage and the rate of disclosure of the intermediate success. Our equilibrium analysis does this comprehensively, and this enables us to discuss the two comparative static results mentioned above under “main results”.

Hopenhayn and Squintani (2016) also consider a model with some similarity to ours,where firms compete in R&D. Each firm makes no progress until it achieves a breakthrough, the arrival of which follows a Poisson process with known intensity. Once a firm achieves a breakthrough, the “development phase” begins and the innovation value grows at a deterministic rate. A firm, after achieving a breakthrough, can decide to end the game by patenting the innovation. This gives a stream of profit to the firm, but it also discloses the technology to the opponents. However, the opponents receive a lower payoff, and this difference is determined by the patent strength. The authors examine the effect of patent strength on the time of disclosure and resulting incentives to conduct the first-stage research . Thus, unlike in our case, disclosure in their paper has nothing to do with whether the opponents can be advantaged in obtaining a more substantial future benefit. Further, in our paper, there is no development phase in the absence of disclosure—the future prize is fixed. Hence, in our model non-disclosure arises purely from the strategic aspect related to the second prize as well as the payoff from the first prize. Finally, they look at the effect of patent strength, which increases the payoff to the discloser; in our case, if the payoff from the final good is increased, both players benefit.

Kocourek (2018) analyses strategic disclosure in a R&D model in which two firms compete to success, which requires two consecutive breakthroughs. A firm achieving the first breakthrough can choose whether to reveal it or not. However, disclosure does not provide any interim payoff and also the rival firm cannot copy the first stage success. The incentive of non-disclosure arises from the fact that it will discourage the rival and lead to a decline in his first stage effort. Wang (2020) studies the design of research contests to achieve the socially optimal outcome within a model similar to ours. Interestingly, they find that to incentivize disclosure and effort, the designer has to offer a prize not just to the first discoverer, but also to the second discoverer.

2 The model

Two players are competing in a research race in continuous time, indexed by $t \geq 0$. Each is working on an identical, two-step project. A player completes step 1 (2, respectively) if and only if he or she makes the discovery F (S, respectively). A player enters step 2 if and only if he or she completes step 1.

At each $t \geq 0$, player i ($i = 1, 2$) chooses a level of effort $e_i \in [0, 1]$. If F has not been discovered by i, the arrival rate for F during $[t, t+dt)$ is γe_i, where $\gamma > 0$.[5] If i has discovered F but has not discovered S, the arrival rate of S during $[t, t + dt)$ is λe_i, where $\lambda > 0$. First discovery of S ends the game. The flow cost per unit of effort is the same for both players. It is $c > 0$ in step 1 and $c' \geq 0$ in step 2 such that $c > c'$. In our analysis, W.L.O.G we assume $c' = 0$.

If i has discovered F at some time $\tau > 0$, at any time $t \geq \tau$, she chooses whether to disclose it, provided neither player has disclosed already. Once F is disclosed, a player who has not completed step 1 can immediately copy it and discover F as well.

Before the game ends, the only public information is whether a disclosure of F has been made, and if so, when. That is, players' past efforts and whether the opponent has discovered F (before a disclosure is made) are private information.

Preferences and payoffs. The first player who discloses F receives $F > 0$ and the first player who discovers S receives $S > 0$. If no disclosure of F has taken place until the first discovery of S, then the discoverer gets $F + S$. Both players discount the future using a common continuous time discount rate $r > 0$. Thus, a prize of F (S) received by a player i at a random time τ is worth $e^{-r\tau}F$ ($e^{-r\tau}S$) to i. Effort costs for each player are similarly discounted.

Histories. Until the first discovery of S, there are only two types of public histories: ones such that F is disclosed and ones without a disclosure. For each player i, let $N_i(t)$ be given by

$$N_i(t) = \begin{cases} 0 & \text{if } i \text{ has not discovered } F \text{ by time } t; \\ 1 & \text{if } i \text{ has discovered } F \text{ and neither player has disclosed } F \text{ by time } t; \\ 2k & \text{if } F \text{ is disclosed by player } k \text{ by time } t. \end{cases}$$

A private history at time t for player i is $(e_{i,s}, N_{i,s})_{s<t}$, which encodes i's private effort history before t, whether i has discovered F, and whether F is disclosed before t.

Effort strategy. Given our assumption of $c' = 0$, once a player enters step 2, effort cost reduces to 0. This means after entering step 2, the player will put full effort until the first discovery of S. Therefore, we only need to specify players i's ($i = 1, 2$) effort strategy before he or she discovers F, that is, conditional on $N_{i,t} = 0$. Let $e_i : \mathbb{R}_+ \to [0, 1]$ be i's effort strategy, where $e_i(t)$ is i's effort at time t in step 1.[6] We require effort strategies to be

[5] The usual interpretation is that the probability of arrival in [t,t+dt] is $\gamma e_i dt$. This is independent of the past.

[6] Our definition does not specify i's strategy after her own deviation. Our results do not change if we assume that after an unobserved deviation, a player reverts to her equilibrium path. This discussion applies

right-continuous.

Disclosure strategy. If player i discovers F at time $\tau \geq 0$, he chooses a CDF over $[\tau, \infty]$, denoted $G_i(|\tau)$. The interpretation is that i plays a randomized stopping time strategy, such that the probability that he has disclosed F by time t is $G_i(t|\tau)$.

Equilibrium. Our equilibrium notion is perfect Bayesian equilibrium (PBE), in which strategies are mutual best responses given beliefs and beliefs satisfy Bayes' rule whenever possible. Note that the only kind of deviation that is observable to an opponent is that a player discloses F when he or she is not supposed to do so on the equilibrium path. Since we assume that disclosing F requires disclosing the method to discover F, after such a deviation, N jumps to 2 and both players enter step 2.

3 Preliminaries

3.1 The continuation equilibrium after disclosure

Suppose one of the players, after completing the first step, discloses the intermediate finding. Once F is disclosed, both players enter step 2 and put full effort until S is discovered. Let U denote a player's equilibrium payoff after F is disclosed. We have

$$U = \frac{\lambda S}{r + 2\lambda}, \tag{1}$$

which is the expected payoff per unit of time, discounted by the discounted rate r and the arrival rate of S, 2λ.

3.2 Beliefs

Players' disclosure incentives are affected by the cost of effort c, the price ratio $\frac{\lambda U}{F}$, and an endogenous variable— a player's belief about whether the opponent has discovered F. Let $\mu_j(t)$ denote the probability that player i assigns to the event that j ($j \neq i$) has discovered F by time t, conditional on no disclosure by t. We call $\mu_j(t)$ j's *reputation* at t. Given players' strategies, let

$$Y_j(t) := E_{(e_1,e_2,d_1,d_2)}(G_i(t|\tau))$$

be the probability that j has revealed by time t, where $E_{(e_1,e_2,d_1,d_2)}$ is the expectation taken over j's discovery time of F, induced by players' strategies. Thus, $dY_j(t)$ denotes the probability that j discloses F during the time interval $[t, t + dt)$. By Bayes rule, j's reputation

to our definition of the disclosure strategy below.

evolves according to[7]

$$d\mu_j(t) = (1-\mu_j(t))(e_j(t)\gamma - \mu_j(t)\lambda)dt - (1-\mu_j(t))dY_j(t)/(1-Y_j(t)). \tag{2}$$

Whenever $Y_j(t)$ is absolutely continuous in time, we define $y_j(t) = \frac{dY_j(t)/dt}{1-Y_j(t)}$ as j's *revealing rate* at time t. For example, if j reveals immediately after discovering F, we have $y_j(t) = e_j(t)\gamma$; if j does not disclose F during some time interval, we have $y_j(t) = 0$.

3.3 Existence and uniqueness

Proposition 1. *There is a symmetric PBE. All symmetric PBE have the same equilibrium outcome (i.e., distribution of equilibrium paths).*

All formal proofs are in the appendix. We now present a symmetric PBE. Depending on the parameters, it is one of the following three types of equilibria:

Definition 1. *A symmetric PBE is*

- *immediately revealing (IR) if each player discloses F immediately after discovering it.*
- *non-revealing (NR) if each player never discloses F (before the game ends).*
- *gradually revealing (GR) if there is a time T such that, before T, neither player discloses F, and after T, each player discloses at a positive rate.*

That is, a GR equilibrium has two phases: in phase 1 (before time T), neither player discloses; in phase 2 (after time T), each player discloses at Poisson rate.

Figures 1 and 2 illustrate the equilibrium, where the horizontal axis represents the cost of effort c and the vertical axis represents the ratio between the two prizes $\frac{\lambda U}{F}$. Figure 1 depicts the situation when the first step is more difficult to accomplish than the second step, i.e., $\lambda > \gamma$. Figure 2 depicts the opposite situation, $\lambda \leq \gamma$. We now read Figures 1 from left to right, fixing a prize ratio. We postpone the intuitions to the next section.

First, as cost increases, players reduce their first-step effort. Given a prize ratio $\frac{\lambda U}{F}$, for low effort costs, effort is 1. As the cost becomes higher, the effort starts falling and eventually goes to zero.

[7]To see this, suppose j's reputation at t is $\mu_j(t)$. By Bayes rule,

$$\mu_j(t+dt) = \frac{\mu_j(t) - \mu_j(t)\lambda dt - dY_j(t)/(1-Y_j(t)) + (1-\mu_j(t))e_j(t)\gamma dt}{1 - \mu_j(t)\lambda dt - dY_j(t)/(1-Y_j(t))}.$$

Subtracting $\mu_j(t)$ on both sides and defining $d\mu_j(t) = \mu_j(t+dt) - \mu_j(t)$, we obtain the above law of motion.

Second, as cost increases, players disclose at a later time. If the final prize S is sufficiently low (i.e., $\frac{\lambda U}{F} < r$), players disclose immediately. If S is moderately low (i.e., $\frac{\lambda U}{F} \in (r, r+\gamma)$, the equilibrium involves immediately revealing at low costs, and gradually revealing at higher costs. If S is moderately high (i.e., $\frac{\lambda U}{F}(r+\gamma, \frac{\lambda}{\lambda-\gamma}(r+\gamma))$), IR disappears, and the equilibrium involves GR at low costs and NR at high costs. Finally, if S is sufficiently high (i.e., $\frac{\lambda}{\lambda-\gamma}(r+\gamma)$), a player never discloses.

Figure 2 shares the same features with Figures 1, except that at low costs, disclosure always occurs, whatever the prize level.

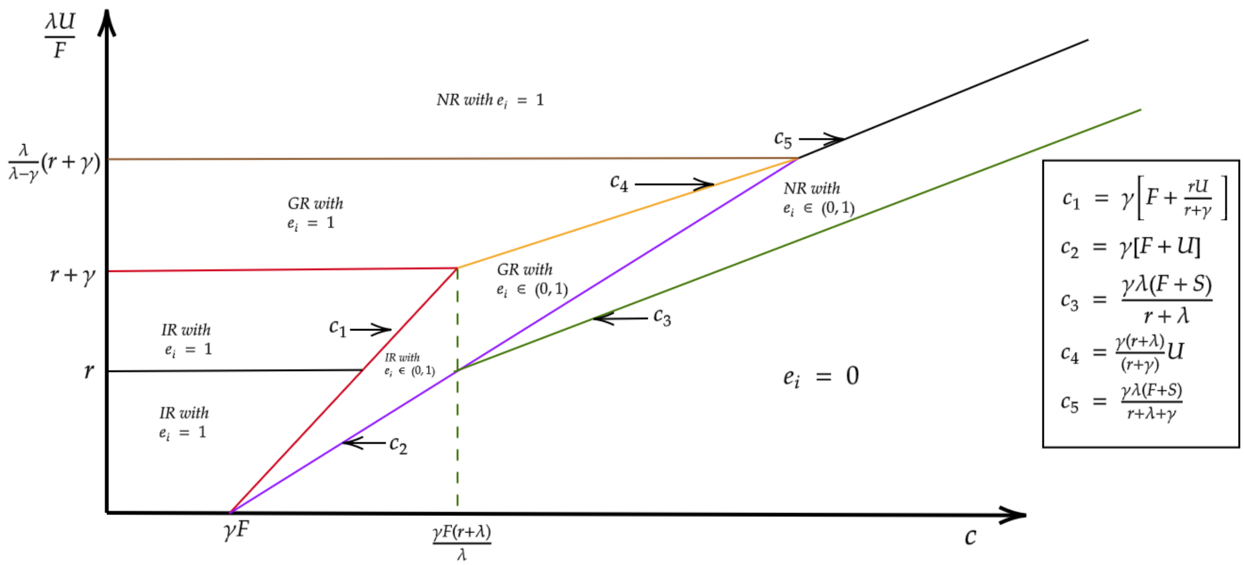

Figure 1: Symmetric PBE: $\lambda > \gamma$

4 Main results

Our main results shed light on the role of increasing the final reward on the expected date at which each result becomes public. The intermediate result F becomes public when it is disclosed, and S becomes public when it is discovered (i.e., when the game ends). We confine ourselves to symmetric PBE here.

4.1 Does increasing S make players disclose F later?

In our setup, disclosing the intermediate finding empowers the opponent to achieve the final result. That is, disclosure is effectively sharing the final reward with the opponent. One would naturally expect that a higher final reward S reduces the incentive of sharing and therefore slows down disclosure. Our first result shows that this is not necessarily true:

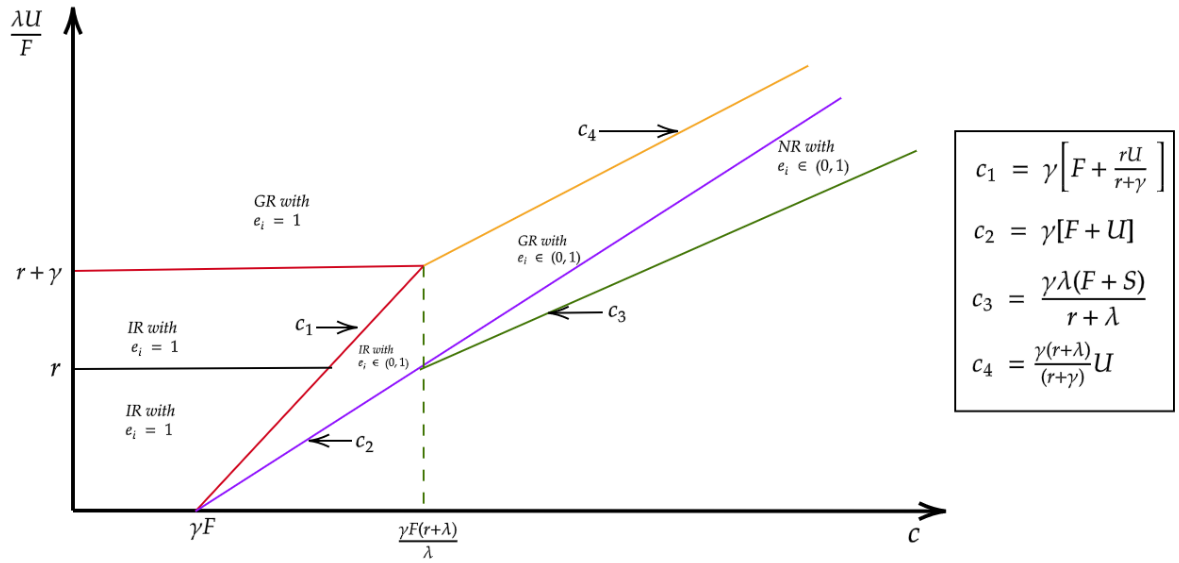

Figure 2: Symmetric PBE: $\lambda < \gamma$

increasing the reward S can increase players' incentives to share the reward S. To formally describe the result, we first define $\bar{c} := \frac{\gamma(r+\lambda)}{\lambda-\gamma}F$ if $\lambda > \gamma$ and $\bar{c} := \infty$ otherwise.[8] The following theorem illustrates the result.

Theorem 1. *In any symmetric PBE, if the effort cost c is intermediate, that is, if $c \in (\frac{\gamma(r+\lambda)}{\lambda}F, \bar{c})$, then increasing S encourages disclosure of F when S is low and discourages it otherwise.*
Specifically, if $c \in (\frac{\gamma(r+\lambda)}{\lambda}F, \bar{c})$, then there is a $\hat{S}$, such that as S increases, the equilibrium disclosure rate increases if $S < \hat{S}$, and decreases if $S > \hat{S}$.

We now explain the intuition of this result. Suppose that i has discovered F and is considering whether to disclose F now or at dt time later. If she discloses now, then, in the event the opponent has not discovered F yet, the opponent can use her finding to discover S. The probability that this event occurs during this dt duration of time is $(1-\mu_j(t))\lambda dt$; if it occurs, the game ends and player i thus suffers a loss of value U. That is, the expected cost from disclosing now is

$$(1-\mu_j(t))\lambda U dt.$$

The benefit from disclosing now is twofold: the prize F will neither be discounted nor be preempted by the opponent. That is, the expected gain from disclosing now is

$$(r + y_j(t) + \mu_j(t)\lambda)F dt,$$

[8] Referring to Figure 3: $\bar{c}$ is the intersection of the horizontal line $\lambda U/F = \lambda\frac{\gamma(r+\lambda)}{\lambda-\gamma}$ and the line representing c_2: $\lambda U/F = \lambda c/(\gamma F) - \lambda$.

where $y_j(t)$ is j's disclosure rate of F and $\mu_j(t)\lambda$ j's discovery rate of S.

Therefore, the prize ratio $\frac{\lambda U}{F}$ has a direct impact on players' disclosure incentive: the higher the final reward, the smaller the disclosure incentive. This is intuitive: since disclosure is essentially sharing the final reward with the opponent, the higher the reward, the lower incentive to share.

In addition to this direct impact, the prize ratio $\frac{\lambda U}{F}$ has an indirect impact on players' disclosure incentive, by affecting players' first stage equilibrium effort and thus changing players' reputations. Intuitively, as the final reward increases, the opponent increases effort (until it reaches 1), which raises his reputation. Therefore, the cost of disclosure is reduced while the gain is increasing (due to a higher preemption risk). That is, the higher the final reward, the higher the disclosure incentive, which is the opposite of the direct effect.

The two combined effects leads to the non-monotone effect of increasing the final reward. Figure 3 presents the qualitative features of the symmetric equilibrium. We focus on the case that $c \in (\frac{\gamma(r+\lambda)}{\lambda}F, \bar{c})$, the middle region shown in Figure 3. In this region, there is no symmetric equilibrium that is immediately revealing. We can divide the region into five subregions: A, B, C, D, E. In region A, rewards are so low that neither player works. In regions D and E, rewards are sufficiently high so that both exert full effort. In between, efforts are interior in equilibrium.

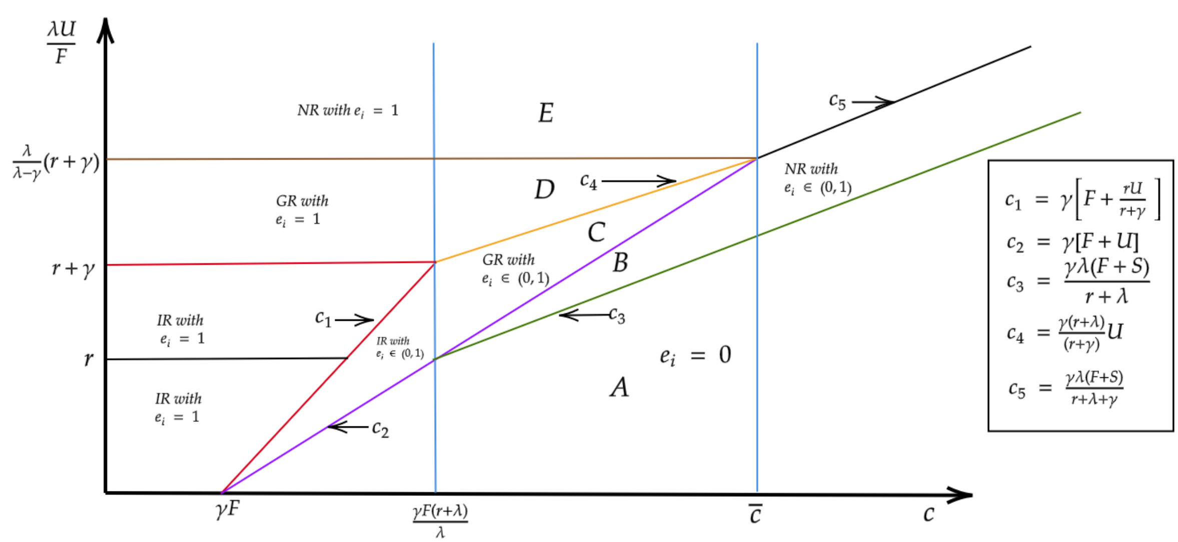

Figure 3: Symmetric PBE

We now explain disclosure behavior, starting with region B. In region B, players put low effort; as a result, the preemption risk is so low that neither player discloses F. As the reward moves up to region C, players compete with high effort. The risk of preemption thus rises, and players begin to disclose continuously.

Once we move to region D, the final reward is sufficiently high so that both players put full effort in the first stage. Therefore, the indirect effect of increasing S ceases to exist and only the direct effect is present. Disclosing F becomes increasingly costly so that players start revealing at a late date and at a lower rate. Finally, as the reward moves up to region E, the first reward F becomes too insignificant compared with the final reward; consequently, neither player discloses F.

4.2 Does increasing S make players discover S quicker?

Our second main result says that if a policy maker's objective is to achieve the final result as early as possible, increasing the final reward can backfire if it is already at a very high level. We illustrate this in the following theorem.

Theorem 2. *Increasing the final reward S has a U-shaped effect on the expected time of the final discovery.*

Specifically, for all c, there is a threshold $\tilde{S}(c)$ such that increasing S (weakly) shortens the time of the final discovery if $S < \tilde{S}(c)$, and (weakly) lengthens it otherwise. Moreover,

- *If cost is sufficiently low: $c \in (0, \gamma F)$, then $\tilde{S}(c) = 0$. The effect is J-shaped.*
- *If cost is high: $c \in (\bar{c}, \infty)$, then $\tilde{S}(c) = \infty$. The effect is L-shaped.*
- *In between: $c \in (\gamma F, \bar{c})$, then $\tilde{S}(c) > 0$. That is, the effect is U-shaped.*

Increasing the final reward affects the time to the final discovery via two channels: effort in the first step, and disclosure of F. As is shown in Theorem 1, the second channel is not monotone.

We first focus on intermediate cost: $c \in (\frac{\gamma(r+\lambda)}{\lambda}F, \bar{c})$. Recall that increasing S has two opposing effects on disclosure: the direct effect—increasing S discourages disclosure, and the indirect effect (via reputation changes)—increasing S encourages disclosure. We have heuristically argued that in regions B and C, the indirect effect dominates. Since increasing S boosts effort, both channels shorten the time of the final discovery. We have also argued that in regions D and E, the indirect effort disappears. Since effort is fixed at 1 in these regions, only the direct effect matters. Thus, increasing S slows down the final discovery.

Therefore, only the effort channel is present: increasing S thus shortens the time to final discovery. Otherwise, players' equilibrium effort is 1, and thus only the disclosure channel is in effect: increasing S lengthens the time to final discovery.

Next, we explain the high cost region: $c \in (\bar{c}, \infty)$. In this case, the symmetric equilibrium is non-revealing. When the final reward is low so that players do not put full effort, only the effort channel is present. When the final reward is high so that players put full effort, since the equilibrium disclosure rate is already 0, increasing S has no effect. Overall, increasing S reduces the time to final discovery.

Finally, we move to the sufficiently low-cost region: $c \in [0, \gamma F]$. Since players' efforts are 1, only the disclosure channel is in effect. As a result, increasing S increases the time to final discovery. The current theorem is depicted in figure 4.

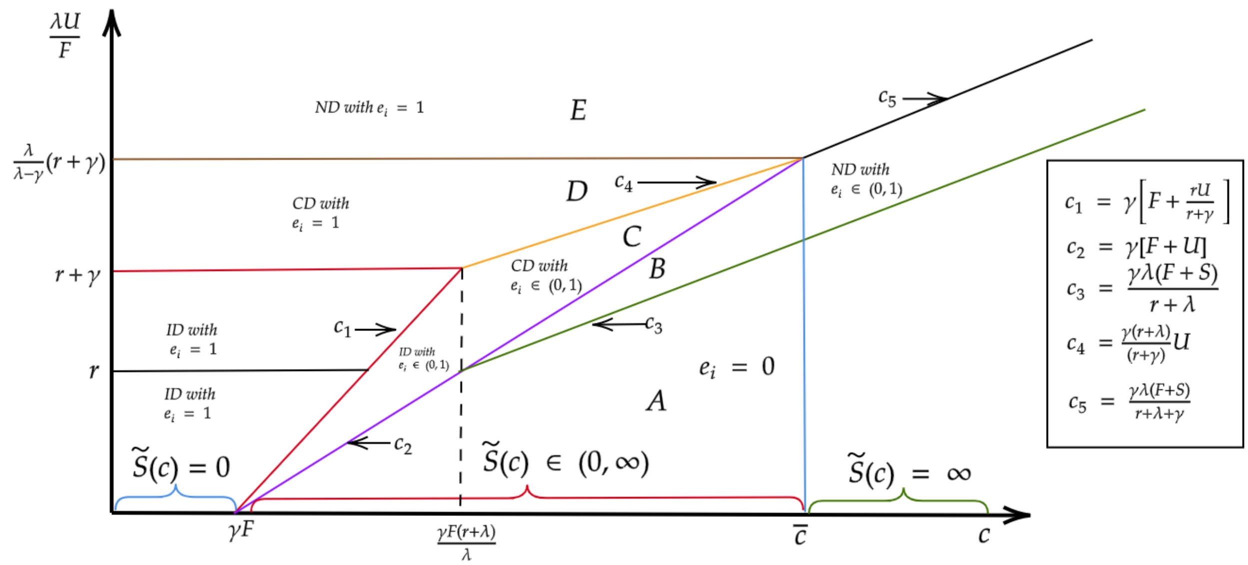

Figure 4: Symmetric PBE

4.3 Does mandatory transparency make players discover S quicker?

We have shown that players delay their disclosure when the final reward is high or at high effort cost. If exactly one player has discovered F, disclosure delay slows down the final discovery: if F were disclosed immediately, the other player can spend all his resources on discovering the final result, instead of wasting it on F. If a policymaker aims to speed up the final discovery, would she want to adopt a mandatory-transparency policy, which (legally) forces players to disclose F immediately? We find that she wants to do so if and only if the final reward is high enough:

Proposition 2. *Mandatory transparency and the final reward are policy complements. That is, mandatory transparency reduces the time to final discovery if and only if the final reward is high enough.*

The intuition is as follows. On the one hand, mandatory transparency eliminates the inefficiency caused by delayed disclosure. On the other hand, it aggravates the free-riding problem: expecting that the opponent will disclose F immediately, each player now has a higher incentive to wait and copy the other's result. When the final reward is low, the second effect is more severe; consequently, mandatory transparency slows down the discovery of S. When the final reward is high so that players are willing to put high effort, the first effect dominates. Thus, increasing S speeds up the discovery of S.

5 Discussion

5.1 Positive effort cost in step 2

We have assumed that effort is costless in step 2. Our main results extend situations in which effort is costly in step 2. The reason is that, since step 2 is the final step and the prize is winner-takes-all, players will put maximal effort as long as the final prize is higher than a threshold and zero effort if it is below the threshold. Thus, if the final reward exceeds the threshold, our main results continue to hold.

6 Conclusion

Our model has addressed the issue of strategic disclosure of intermediate success in research races. Our model has comprehensively captured the interplay between the effort during an intermediate stage of research and the intermediate finding's disclosure incentives.

First, we have demonstrated that heavily rewarding the final success can sometimes backfire because it can delay the final success. We have comprehensively characterised the parameter ranges for which increasing the reward from final success can bring forward the expected discovery time.

Secondly, we have shown that a policy of *Mandatory transparency* is not always beneficial. One of the reasons for the advocacy of mandatory transparency is that it reduces the intermediate step's inefficient duplicative research. However, we have shown that because of its effect on the first stage effort due to the players' free-riding, such a policy can bring forward the expected time of success only when the reward from the second prize is high enough.

References

Bhattacharya, Sudipto, Jacob Glazer, and David EM Sappington, "Licensing and the sharing of knowledge in research joint ventures," *Journal of Economic Theory*, 1992, *56* (1), 43–69.

Brams, Steven J and Alan D Taylor, "An envy-free cake division protocol," *The American Mathematical Monthly*, 1995, *102* (1), 9–18.

Hopenhayn, Hugo A and Francesco Squintani, "Patent rights and innovation disclosure," *The Review of Economic Studies*, 2016, *83* (1), 199–230.

Keller, Godfrey, Sven Rady, and Martin Cripps, "Strategic experimentation with exponential bandits," *Econometrica*, 2005, *73* (1), 39–68.

Kocourek, P, "Revealing Private Information in a Patent Race," *Mimeo NYU*, 2018.

Scotchmer, Suzanne and Jerry Green, "Novelty and disclosure in patent law," *The RAND Journal of Economics*, 1990, pp. 131–146.

Song, Yangbo and Mofei Zhao, "Dynamic RD competition under uncertainty and strategic disclosure," *Journal of Economic Behavior Organization*, 2021, *181*, 169–210.

Wang, Yuehui, "Knowledge Sharing in Research Contests," 2020.

Appendix

A Proof of Proposition 1

Let $c_1, c_2, ..., c_5$ be the same as in Figures 1. Define $\hat{\mu}$ by

$$\hat{\mu} := (\frac{\lambda U}{F} - r)/(\frac{\lambda U}{F} + \lambda). \tag{3}$$

We will show that $\hat{\mu}$ is the maximum of a player's reputation during continuous disclosure in any symmetric PBE. Note that $\hat{\mu} \geq 0$ if and only if $\frac{\lambda U}{F} \geq r$.

Let $\bar{\mu} := \min\{1, \gamma/\lambda\}$, which is the limit of a player's reputation if he exerts full effort and never discloses. This is an upper bound of reputation levels in our setup, since there is no feasible strategy by which a player's reputation can exceed it. We assume that both players' reputations are lower than $\bar{\mu}$: $\mu_i(t) \leq \bar{\mu}$, for all $t \geq 0$ and $i \in \{1, 2\}$. Recall that $\hat{\mu}$ is defined by (3). If $\lambda \leq \gamma$, then $1 = \hat{\mu} < \bar{\mu}$; if $\lambda > \gamma$, then $\hat{\mu} < \bar{\mu}$ if and only if $\frac{\lambda U}{F} < \frac{\lambda(r+\gamma)}{\lambda-\gamma}$.

Take a PBE. Let $W_i^0(t)$ and $W_i^1(t)$ be player i's equilibrium payoffs at time t if he has not discovered F and that if he has discovered F, respectively. Let $e_i(\cdot)$ be i equilibrium effort strategy. Let $R_i(t)$ denote the rate at which i discloses F; that is, $R_i(t)$ is equal to $e_i(t)\gamma$ if i discloses immediately, to $y_i(t)$ if he discloses continuously, and to 0 if he does not disclose.

A.1 Immediate disclosure PBE

Lemma 1. *If $\lambda U/F < r$, every PBE is immediately disclosing.*

Proof. Suppose by negation there is a PBE in which j does not immediately discloses F during some time interval $(t', t' + \epsilon)$. Let $t'' > t'$ be the earliest time at which j starts disclosing. Then, on $(t', t''))$, j's value function satisfies $W_j^1(t) \geq F + U$ and the HJB equation

$$(r + \lambda + \mu_i(t)\lambda + R_i(t))W_j^1(t) \quad = \quad \lambda(F + S) + R_i(t)U + \frac{dW_j^1(t)}{dt}. \tag{4}$$

Subtracting $(r + \lambda + \mu_i(t)\lambda + R_i(t))(F + U)$ from both sides and using $\lambda S = (r + 2\lambda)U$, we have

$$(r + \lambda + \mu_i(t)\lambda + R_i(t))(W_j^1(t) - F - U) \quad = \quad [(1 - \mu_i(t))\lambda U - (r + \mu_i(t)\lambda)F] - R_i(t)F + \frac{dW_j^1(t)}{dt}. \tag{5}$$

The term in the square brackets on the right-hand side is less than $\lambda U - rF$. Thus, if $\lambda U/F < r$, then it is always negative. Since the left-hand side is weakly positive, for (5) to hold, we must have $\frac{dW_j^1(t)}{dt} \geq -(\lambda U - rF)$. Using $\lambda U/F < r$ again, we have $\frac{dW_j^1(t)}{dt} > 0$. And since $W_j^1(t') \geq F + U$, we have $W_j^1(t) > F + U$ over $(t', t'']$. By definition of t'', j discloses at t'', meaning that if $t'' < \infty$ $W_j^1(t'') = F + U$. A contradiction. Therefore, we have $t'' = \infty$, i.e., j never discloses after t'. But then, the fact that $\frac{dW_j^1(t)}{dt} > 0$ implies that $W_j^1(t)$ grows unboundedly as $t \to \infty$, which is impossible in our setup.

Therefore, in every PBE, each player discloses immediately. ■

Lemma 2. *If $\lambda U/F > r$, there is no symmetric PBE such that an ID phase with full effort is followed by a ND phase.*

Proof. Suppose by negation that there is a symmetric PBE such that it has an ID phase with full effort during time interval (t_1, t_2) and a ND phase during time interval (t_2, t_3), where $t_1 < t_2 < t_3$.

During the ID phase (t_1, t_2), each player's reputation is 0; that is, $\mu_i(t) = 0$. At t_2, μ_i cannot jump upward since success arrives at Poisson rate. It cannot jump downward since $\mu_i(t_2-) = 0$. That is, we have $\mu_i(t_2-) = \mu_i(t_2+) = 0$.

Moreover, if j has discovered F, j's value function W_j^1 satisfies $W_j^1(t) = F + U$ on (t_1, t_2). At t_2, W_j^1 must be continuous because, first, if $W_j^1(t_2-) > W_j^1(t_2+)$, then $W_j^1(t) < F + U$ during some time interval $(t_2, t_2 + \epsilon)$, implying that j strictly prefers to disclose, which contracts the fact that he does not disclose during (t_2, t_3) in equilibrium. Second, if $W_j^1(t_2-) < W_j^1(t_2+)$, then j strictly prefers not to disclose during some time interval $(t_2 - \epsilon, t_2)$, which contracts the fact that he discloses immediately during (t_1, t_2) in equilibrium.

During the ND phase (t_2, t_3), W_j^1 satisfies

$$(r + \lambda + \mu_i(t)\lambda)(W_j^1(t) - F - U) \quad = \quad [(1 - \mu_i(t))\lambda U - (r + \mu_i(t)\lambda)F] + \frac{dW_j^1(t)}{dt}. \tag{6}$$

Applying $W_j^1(t_2) = F + U$ and $\mu_j(t_2) = 0$, we have,

$$0 \quad = \quad [\lambda U - rF] + \frac{dW_j^1(t_2+)}{dt}. \tag{7}$$

Or, $\frac{dW_j^1(t_2+)}{dt} = -[\lambda U - rF] < 0$. Since $W_j^1(t_2) = F + U$, this implies that $W_j^1(t) < F + U$ on some time

interval $(t_2, t_2+\epsilon)$. Thus, j strictly prefer to disclose during the time interval $(t_2, t_2+\epsilon)$. A contradiction.
■

Lemma 3. *If $\lambda U/F > r$, there is no symmetric PBE such that a ND phase is followed by an ID phase.*

Proof. Suppose by negation that there is a symmetric PBE such that it has a ND phase during time interval (t_1, t_2) and an ID phase with full effort during time interval (t_2, t_3), where $t_1 < t_2 < t_3$. Then t_2, we have $\mu_i(t_2) > 0$. Thus, each player discloses with strictly positive probability at t_2. As a result, each player receives $\frac{1}{2}F + U$ by disclosing at t_2. As a result i's value function $W_i^1(t) < F + U$ on some time interval $(t_2 - \epsilon, t_2)$. However, if player i discloses at $t_2 - \epsilon$ instead, he receives payoff $F + U$, which is strictly higher than his continuation value in equilibrium. That is, i has a profitable deviation at time $t_2 - \epsilon$. A contradiction.

■

Lemma 4. *If player j discloses continuously on some time interval (t', t''), then player i's disclosure rate on (t', t'') satisfies*[9]

$$R_i(t) = (1 - \mu_i(t))\lambda U/F - (r + \mu_i(t)\lambda). \tag{8}$$

Moreover, if during (t', t'') player j exerts interior effort, then on this interval,

$$\mu_i(t) = \mu^* \quad := \quad 1 - \frac{\gamma F(r+\lambda)}{\lambda c}, \tag{9}$$

$$e_i(t) = e^* \quad := \quad \frac{(r+\lambda)U}{c} - \frac{r}{\gamma}. \tag{10}$$

Proof. Whenever j discloses continuously, his value function satisfies (5), $W_j^1(t) = F + U$, and $\frac{dW_j^1(t)}{dt} = 0$. We thus have (8).

Moreover, W_j^0 satisfies

$$(r + \mu_i(t)\lambda)W_j^0(t) \quad = \quad \max_e e(\gamma(W_j^1(t) - W_j^0(t)) - c) + R_i(t)(U - W_j^0(t)) + \frac{dW_j^0(t)}{dt}. \tag{11}$$

Since j discloses continuously, if j exerts interior effort, we have $W_j^0(t) = W_j^1(t) - c/\gamma = F + U - c/\gamma$. Thus, $dW_j^0(t)/dt = 0$, and the first term on the right-hand side is 0. These further imply that

$$(r + \mu_i(t)\lambda)(F + U - c/\gamma) \quad = \quad R_i(t)(U - (F + U - c/\gamma)). \tag{12}$$

This equation and (8) imply that $\mu_i(t) = \mu^*$.

Thus, $\mu_i(t)$ is a constant, that is, $d\mu_i(t) = 0$. Using the law of motion of $\mu_i(t)$ during continuous disclosure, (2), we have that $e_i(t)\gamma = R_i(t) + \mu^*\lambda$. Combining (8), we have $e_i(t) = e^*$. ■

Corollary 1. *A symmetric equilibrium has a CD phase with interior effort only if one of the following holds:*

(a) $\frac{\lambda U}{F} \in (r, r+\gamma]$ *and* $c \in (\gamma F[\frac{r+\lambda}{\lambda}], c_2]$,

[9]This part also holds if j's payoff satisfies $W_j^1(t) = F + U$ on (t', t'') if j chooses not to disclose during this time interval.

(b) $\frac{\lambda U}{F} > r+\gamma$ *and* $\lambda \leq \gamma$ *and* $c \in (c_4, c_2]$,

(c) $\frac{\lambda U}{F} \in (r+\gamma, \frac{\lambda}{\lambda-\gamma}(r+\gamma))$ *and* $\lambda > \gamma$ *and* $c \in (c_4, c_2]$.

Proof. By Lemma 4, A symmetric equilibrium has a CD phase with interior effort only if during the CD phase, $R_i(t) > 0$, $\mu^* \in (0,1)$, and $e^* \in (0,1)$.

Using the definition of $\hat{\mu}$, if $\mu_i(t) = \mu^*$ in (8), then $R_i(t) > 0$ is equivalent to $\mu_i^* < \hat{\mu}$. Combining (9), we have

$$R_i(t) > 0 \Leftrightarrow \mu^* < \hat{\mu} \Leftrightarrow c < \gamma(U+F) \equiv c_2. \tag{13}$$

Using (9), we have

$$\mu^* > 0 \Leftrightarrow c > \gamma F[\frac{r+\lambda}{\lambda}]. \tag{14}$$

In (10), $e^* > 0$ iff

$$e^*\gamma = (1-\mu^*)\frac{\lambda U}{F} - r \Rightarrow c < \frac{\gamma}{r}(r+\lambda)U. \tag{15}$$

$e^* < 1$ iff

$$c > \frac{\gamma(r+\lambda)U}{(r+\gamma)} \equiv c_4. \tag{16}$$

From Lemma 1, every equilibrium is an ID equilibrium. Thus, for a CD phase to exist, we must have $\frac{\lambda U}{F} \geq r$. If $\frac{\lambda U}{F} \geq r$, then $\hat{\mu} = 0$ and thus $\mu^* = 0$ (by (13)) and $e^* = 0$ (by (15)). This is not a CD phase because neither player works and the game ends. Thus, we only need to consider $\frac{\lambda U}{F} > r$.

We first compare the upper bounds in (13) and (15). If $\frac{\lambda U}{F} > r$, then $\frac{\gamma}{r}(r+\lambda)U > c_2$. Thus, we can ignore (15) in this case.

We now compare the lower bounds in (14) and (16). We have,

$$\gamma F[\frac{r+\lambda}{\lambda}] > c_4 \Leftrightarrow \frac{\lambda U}{F} < r+\lambda.$$

That is, if $\frac{\lambda U}{F} < r+\gamma$, then (14) is tighter, and $\frac{\lambda U}{F} > r+\gamma$, then (16) is tighter. Moreover, if $\frac{\lambda U}{F} > r+\gamma$, the set $(c_4, c_2]$ is nonempty iff the following: $\lambda \leq \gamma$ or ($\lambda > \gamma$ and $\frac{\lambda U}{F} < \frac{\lambda}{\lambda-\gamma}(r+\gamma)$).

Therefore, for a CD phase with interior effort to exist, we must have one of the conditions in the corollary holds. ■

Lemma 5. *Assume that both players disclose immediately during some time interval* (t', t'') *in a PBE. Then, during* (t', t''),

- *If* j *exerts interior effort, then*

$$e_i(t) = e_{ID} := \frac{r}{\gamma}[\frac{\gamma(F+U)-c}{c-\gamma F}]. \tag{17}$$

- j *exerts interior effort only if* $\frac{\lambda U}{F} < r$ *and* $c \in (c_1, c_2]$ *or* $r < \frac{\lambda U}{F} < r+\gamma$ *and* $c \in (c_1, \frac{\gamma F(r+\lambda)}{\lambda}]$.

Proof. Part 1. Given that both players disclose immediately, if j exerts interior effort around some time t, then his value function W_j^0 satisfies (11) and (12), with $\mu_i(t) = 0$ and $R_i(t) = e_i(t)\gamma$. (12) then implies (17). Note that $e_{ID} \in (0,1)$ iff $c \in (c_1, c_2)$.

Part 2. We now derive a necessary condition under which immediate disclosures are mutual best replies. Suppose i is disclosing immediately. j is willing to disclose immediately only if the payoff from delaying disclosure by dt duration of time, denoted $\tilde{W}_j^1(t)$, is at most $F+U$. Note that $\tilde{W}_j^1(t)$ satisfies the HJB equation (4), if we let $R_i(t) = e_i(t)\gamma$ and $\mu_i(t) = 0$ (as i is disclosing immediately). We thus have

$$\tilde{W}_j^1(t) = \frac{\lambda(F+S) + e_i(t)\gamma U + d\tilde{W}_j^1(t)/dt}{r + \lambda + e_i(t)\gamma}.$$

Since at $t+dt$, j discloses immediately, we have $\tilde{W}_j^1(t+dt) = F+U$. Since $\tilde{W}_j^1(t) \leq F+U$, we must have $d\tilde{W}_j^1(t)/dt \geq 0$. Thus, $\tilde{W}_j^1(t) \geq \frac{\lambda(F+S)+e_i(t)\gamma U}{r+\lambda+e_i(t)\gamma}$. If i's effort is interior, we have $e_i(t) = e_{ID}$. For an ID equilibrium with interior effort to exist, the following conditions must hold: $e_{ID} \in (0,1)$ and $\frac{\lambda(F+S)+e_{ID}\gamma U}{r+\lambda+e_{ID}\gamma} \leq F+U$. These conditions are equivalent to $\frac{\lambda U}{F} < r$ and $c \in (c_1, c_2]$ or $r < \frac{\lambda U}{F} < r+\gamma$ and $c \in (c_1, \frac{\gamma F(r+\lambda)}{\lambda}]$.

■

In the sequel, we focus on simple equilibria: both players exert positive effort at any t.

Lemma 6. *Consider a PBE in which both players disclose immediately and exert strictly positive effort at any $t \geq 0$. If both players exert interior effort over some time $(t', t'+\epsilon)$, then $e_1(t) = e_2(t) = e_{ID}$, for all $t \geq 0$.*

Proof. From Lemma 5, over time $(t', t'+\epsilon)$, both players put effort e_{ID}, which implies that $c \in (c_1, c_2)$. Let $t'' > t'$ be the earliest time at which at least one player switches to full effort. Lemma 5 implies that it is impossible to have one player exerts effort 1 and the other interior effort. Since we focus on equilibria in which both players always put strictly positive effort, it must be that at t'', both players switch to full effort. Suppose their effort is 1 over $(t'', t''+\eta)$ for some $\eta > 0$. Since $W_j^1 = F+U$, for j to put full effort, we must have $W_j^0(t) \leq F+U-c/\gamma$ on $(t'', t''+\eta)$, and $dW_j^0(t''+)/dt \leq 0$.

During $(t', t''+\eta)$ j's value function satisfies (11), with $\mu_i(t) = 0$ and $R_i(t) = e_i(t)\gamma$. Subtracting $r(F+U-c/\gamma)$ from both sides, we have

$$\begin{aligned} r(W_j^0(t) - (F+U-c/\gamma)) &= \max_e e(\gamma(W_j^1(t) - W_j^0(t)) - c) + \frac{dW_j^0(t)}{dt} \\ &\quad + e_i(t)\gamma(U - W_j^0(t)) - r(F+U-c/\gamma). \end{aligned} \tag{18}$$

Over $(t', t'']$, the left-hand side and the first line on the right-hand side are zero. Thus, the second line on the right-hand side is also zero, implying that $U - W_j^0(t) > 0$. Since over $(t'', t''+\eta)$, both players put higher effort and $W_j^0(t)$ is lower than over $(t', t'']$, we have, for all $t \in (t'', t''+\eta)$

$$e_i(t)\gamma(U - W_j^0(t)) - r(F+U-c/\gamma) \geq e_i(t''-)\gamma(U - W_j^0(t'')) - r(F+U-c/\gamma) = 0. \tag{19}$$

We now show that $\frac{dW_j^0(t)}{dt} < 0$ for all $t > t''$. Suppose not. Let $\hat{t}$ be the earliest time after t'' such that $\frac{dW_j^0(t)}{dt} \geq 0$. Over time $(t'', \hat{t})$, $\frac{dW_j^0(t)}{dt} < 0$, thus j must strictly prefer to put full effort. Using Lemma again, we must have $e_i(t) = 1$ on $(t'', \hat{t})$. At $\hat{t}$, the right-hand side of (18) is strictly positive since j is willing to put full effort, $\frac{dW_j^0(\hat{t}+)}{dt} \geq 0$, and (19) holds. But since left-hand side of (18) is zero at t'' and is strictly decreasing on $(t'', \hat{t})$, it is strictly negative at $\hat{t}$. A contradiction. We now show that before t', both players must put interior effort as well; then we are done. Suppose $t' > 0$ and there is some $\epsilon' > 0$ such that on $(t'-\epsilon', t')$, both players put full effort. The first term on the right-hand side over $(t'-\epsilon', t')$ is weakly higher than it is over $(t', t'+\epsilon)$, since players put full effort in the former interval and interior in the latter. Moreover, we have shown that $U > W_j^0(t)$ over a neighborhood of t'. Since e_i jumps downward at t', For (18) to hold, we must have $\frac{dW_j^0(t)}{dt} \geq 0$ jumps upward at t'. Since $\frac{dW_j^0(t)}{dt} = 0$ over $(t', t'+\epsilon)$, we have $\frac{dW_j^0(t)}{dt} < 0$ over a left neighborhood of t'. But then, over this neighborhood, we must have $W_j^0(t) > W_j^0(t') = F + U - c/\gamma$, meaning that j is unwilling to put any effort. This contradicts our assumption that he puts effort 1.

■

We thus have

Corollary 2. *Consider a PBE in which both players disclose immediately and exert strictly positive effort at any $t \geq 0$.*

- *If $c \leq c_1$, $e_1(t) = e_2(t) = 1$ for all t.*
- *If $\frac{\lambda U}{F} < r$ and $c \in (c_1, c_2]$ or $r < \frac{\lambda U}{F} < r + \gamma$ and $c \in (c_1, \frac{\gamma F(r+\lambda)}{\lambda}]$, then $e_1(t) = e_2(t) = e_{ID}$ for all t.*

Proof. From Lemma 6, in an equilibrium described in this corollary, either $e_1(t) = e_2(t) = e_{ID}$, for all $t \geq 0$, or $e_1(t) = e_2(t) = 1$, for all $t \geq 0$.

From Lemma 5, if $c \leq c_1$, it is impossible to have players exerting interior effort in immediate-disclosure equilibrium. Thus, we have $e_1(t) = e_2(t) = 1$, for all $t \geq 0$.

If $c \in (c_1, c_2)$, suppose by negation that there is an immediate-disclosure equilibrium in which both players put full efforts. Both players value functions satisfy (18), with $W_j^0(t) \leq W_j^1(t) - c/\gamma = F + U - c/\gamma$. Moreover, e_{ID} satisfies

$$e_{ID}\gamma(U - (F + U - c/\gamma)) - r(F + U - c/\gamma) = 0.$$

Since $e_{ID} < 1$ and $W_j^0(t) \leq F + U - c/\gamma$, there is some $\eta > 0$ such that

$$\gamma(U - W_j^0(t)) - r(F + U - c/\gamma) < \eta.$$

Applying this inequality to (18), as the left-hand side is weakly negative, we must have $dW_j^0(t)/dt < -\eta$ for all t. But then, $W_j^0(t)$ will become negative at some point, which is contradicts the assumption that j is willing to put full effort (because j can guarantee at least 0 payoff by putting no effort). ■

Lemma 7. *There is a symmetric PBE with immediate disclosure only if $\lambda U/F \leq r + \gamma$. Moreover, in any symmetric PBE with immediate disclosure,*

- *both players put full effort iff $c \leq c_1$.*
- *both players put effort e_{ID}, defined by (17), iff the following holds: $\frac{\lambda U}{F} < r$ and $c \in (c_1, c_2]$ or $r < \frac{\lambda U}{F} < r + \gamma$ and $c \in (c_1, \frac{\gamma F(r+\lambda)}{\lambda}]$*

Proof. We first show that immediate disclosure are mutual best replies iff $\lambda U/F \leq r + \gamma$. Suppose player i is disclosing immediately. We know that j's payoff from immediate disclosure is $W^d = F + U$. If j never discloses, then his payoff after he has succeeded in step 1, $W_j^1(t)$, satisfies

$$W_j^1(t) = \lambda\, dt[F + S] + \gamma\, dtU + (1 - r\, dt)(1 - \lambda\, dt - \gamma\, dt)(W_j^1(t + dt)),$$

where it can be shown that $W_j^1(t) = W_j^1(t + dt)$. We thus have,

$$rW_j^1(t) = \lambda((F + S) - W_j^1(t)) + \gamma(U - W_j^1(t)).$$

Or,

$$W_j^1(t) = \frac{\lambda(F + S) + \gamma U}{r + \lambda + \gamma}.$$

Disclosing immediately is a best response only if $W^d \geq W_j^1$, or equivalently, $\frac{\lambda U}{F} \leq r + \gamma$.

We now prove the second part. Take a symmetric ID equilibrium. i's value function W_i^0 satisfies the following HJB.

$$rW_i^0 = \max_{e_i} e_i[-c + \gamma(W_i^1 - W_i^0)] + e_i\gamma\left(U - W_i^0\right). \tag{20}$$

During immediate disclosure, we have $w_i^1 = F + U$. Applying this equation and $e_i = 1$ in (20), we have,

$$W_i^0 = \frac{\gamma(F + 2U) - c}{r + 2\gamma}.$$

Given that the other player is choosing $e_j = 1$, it is a best response for player i to choose $e_i = 1$ iff

$$-c + \gamma(W_i^1 - W_i^0) \geq 0.$$

$$\Leftrightarrow -c + \gamma\left(F + U - \frac{\gamma}{r + 2\gamma}(F + 2U) + \frac{c}{r + 2\gamma}\right) \geq 0.$$
$$\Leftrightarrow c \leq c_1.$$

From the proof of lemma 5, we can conclude about the interior effort part. ■

Remark: Assume $\lambda U/F \leq r + \gamma$. Corollary 2 and Lemma 7 imply that if $c \leq c_1$, in the Pareto-dominating symmetric equilibrium, both players put full effort. If $\frac{\lambda U}{F} < r$ and $c \in (c_1, c_2]$ or $r < \frac{\lambda U}{F} < r + \gamma$ and $c \in (c_1, \frac{\gamma F(r+\lambda)}{\lambda}]$, in any symmetric PBE with immediate disclosure, players put effort e_{ID}.

A.2 Some best replies

Lemma 8. *Before player j discovers F, if j's payoff reaches 0 at some t', his payoff will remains at 0 before he succeeds; that is, if $W_j^0(t') = 0$ for some t', then $W_j^0(t)$ for all $t \geq t'$.*

Assume that $\lambda U/F > r$ so that $\hat{\mu} > 0$ (defined by (3)). After player j discovers F, if j's payoff reaches $F + U$ at some t' and i's reputation $\mu_i(t') = \hat{\mu}$, then $W_j^1(t) = F + U$ and $\mu_i(t') = \hat{\mu}$ for all $t \geq t'$.

Proof. Part 1. j's value function satisfies (11). The first term on the right-hand side of (11) is weakly positive. Therefore, for all t such that $W_j^0(t) \leq U$, we have

$$(r + \mu_i(t)\lambda)W_j^0(t) \quad \geq \quad \frac{dW_j^0(t)}{dt}.$$

Since $W_j^0(t') = 0$, if $t'' > t'$ is such that $W_j^0(t) < U$ for all $t \in (t', t'')$, the above inequality is satisfied on this interval. Using Grönwall's inequality, we have, for all $t \in (t', t'')$, $W_j^0(t) \leq W_j^0(t') \exp \int_t^{t'} (r + \mu_i(s)\lambda)ds = 0$. Since W_j^0 is continuous, we have $W_j^0(t) = 0$ for all $t \geq t'$.

We now show Part 2. Let t' be the earliest time at which at least one player's reputation reaches $\hat{\mu}$. Let i be (one of) the player whose reputation is $\hat{\mu}$ at t'. We have $\mu_j(t') \leq \hat{\mu}$. j's value function satisfies (5). The term in the square brackets on the right-hand side of (5) is decreasing in $\mu_i(t)$ and is positive if $\mu_i(t) < \hat{\mu}$ and negative if $\mu_i(t) > \hat{\mu}$. There are two cases: (a) there is some $\epsilon > 0$ such that $\mu_i(t) \geq \hat{\mu}$ for all $t \in (t', t' + \epsilon)$; (b) there is some $\epsilon > 0$ such that $\mu_i(t) < \hat{\mu}$ for all $t \in (t', t' + \epsilon)$.

In Case (a), for all $t \in [t', t'+\epsilon)$,

$$(r+\lambda+\mu_i(t)\lambda+R_i(t))(W_j^1(t)-F-U) \quad \geq \quad \frac{dW_j^1(t)}{dt}. \tag{21}$$

Since $W_j^1(t')-F-U=0$, Grönwall's inequality implies that $W_j^1(t)-F-U=0$ for such t's. (5) implies that $\mu_i(t)=\hat{\mu}$ and $R_i(t)=0$. Since W_j^1 is continuous, we must have $W_j^1(t)=F+U$ for all $t \geq t'$.

We now show that Case (b) never holds, and our proof of Part 2 is then completed. First, at t', we have $\frac{dW_j^1(t+)}{dt} \geq 0$, because otherwise, we would have $W_j^1(t) < F+U$ for some $t > t'$ and j would prefer to disclose with probability 1. This means that the first three terms on the right-hand side of (5) is weakly negative, and thus (21) holds. Grönwall's inequality implies that $W_j^1(t)-F-U=0$ for such t's. Together with $\mu_i(t) < \hat{\mu}$, (5) implies that $R_i(t) > 0$ and i is thus disclosing continuously. For μ_i to decrease, i's effort must be interior. Using again $W_i^1(t)=F+U$, these imply that $W_i^0(t)=F+U-c/\gamma$ and $dW_i^0(t)/dt=dW_i^1(t)/dt=0$ and .

We have two subcases: (b.1) $\mu_j(t'') \in (0,\hat{\mu}]$ or (b.2) $\mu_j(t'')=0$. In Case (b.1), j is either not disclosing or disclosing continuously. To satisfy $W_i^1(t)=F+U=W_i^0(t)+c/\gamma$, we must have $\mu_j(t)=\mu^*$ and $e_j(t)=e^*$, by Lemma 4 (after swapping i and j). This then means that j's effort is interior. Since we have shown that $W_j^1(t)=F+U$, we have $W_j^0(t)=F+U-c/\gamma$. These are possible only if $\mu_i(t)=\mu^*$ and $e_i(t)=e^*$. Since $\mu^* < \hat{\mu}$, this contradicts $\mu_i(t'')=\hat{\mu}$.

In Case (b.2), $\mu_j(t)=0$ means that j is disclosing immediately. To keep $W_j^1(t)=F+U$, i's disclosing rate at $\hat{\mu}$ must be 0. That is, we have $R_i(t)$ close to 0 on a small right neighborhood of t'. For j to be willing to disclose immediately, we must have $\lambda U/F \leq r+R_i(t)$. Taking limit as $t \to t'$, we must have $\lambda U/F \leq r$. This contradicts our assumption $\lambda U/F > r$. A contradiction. ■

A corollary from this lemma is the following:

Corollary 3. *If there is a time interval (t_1,t_2) during which i's disclosure rate $R_i(t) > 0$, then*

1. *at all $t \in (t_1,t_2)$, $W_j^0(t) > 0$.*
2. *If, moreover, j discloses at a positive rate and $e_j(t) \in (0,1)$ on $(t_2-\epsilon,t_2)$ for some $\epsilon > 0$, then $W_j^0(t_2) > 0$.*

Proof. To see Part 1, suppose by negation that there is some $t \in (t_1,t_2)$ such that $W_j^0(t)=0$, then $dW_j^0(t+)/dt \geq 0$; otherwise, W_j^0 will be negative, which is impossible in equilibrium. If i discloses continuously, that is, if $R_i(t) > 0$, then (11) will be violated, since the left-hand side is 0 and the right-hand side is strictly positive: the first term on the right-hand side is nonnegative, the second term is strictly positive, and the third term satisfies $dW_j^0(t+)/dt \geq 0$.

To see Part 2, since j discloses continuously during $(t_2-\epsilon,t_2)$, $W_j^1(\cdot)$ is a constant on this interval; and since j's effort is interior, we have, $W_j^0(\cdot)$ is a constant as well, and hence $dW_j^0(t)/dt=0$. From Part 1, $W_j^0(t) > 0$ on $(t_2-\epsilon,t_2)$, and since it is a constant and is continuous, we have $W_j^0(t_2) > 0$. ■

Lemma 9. *If there is a time interval during which both players disclose continuously and j's reputation is not decreasing, then $\hat{\mu} \in (0,\gamma/\lambda)$, that is, either $\gamma \geq \lambda$ and $\frac{\lambda U}{F} > r$, or $\gamma < \lambda$ and $\frac{\lambda U}{F} \in (0,\frac{\lambda(r+\gamma)}{\lambda-\gamma})$.*

Proof. Recall that when both players disclose continuously, j's disclosing rate is given by (8)and the rate is strictly positive only if $\mu_j(t) < \hat{\mu}$; thus, we must have $\hat{\mu} > 0$. Also, j's reputation is not decreasing during a time interval if and only if $0 \leq e_j(t)\gamma-\mu_j(t)\lambda-R_j(t)$. Using (8)and $\mu_j(t) < \hat{\mu}$, we have,

$$0 \leq e_j(t)\gamma-\mu_j(t)\lambda-R_j(t)=e_j(t)\gamma-(1-\mu_j(t))\frac{\lambda U}{F}-r < \gamma-(1-\hat{\mu})\frac{\lambda U}{F}-r=\gamma-\hat{\mu}\lambda,$$

where the last equality is due to the definition of $\hat{\mu}$. ■

A.3 Continuous-disclosure PBE

Assume that $F+U-c/\gamma>0$. If $F+U-c/\gamma\geq 0$, then there is no continuous-disclosure PBE in which both players puts positive effort, since $W_i^1(t)-W_i^0(t)=F+U-W_i^0(t)<F+U-c/\gamma\leq 0$, where the strictly inequality follows from Part 1 of Corollary 3.

Lemma 10. *Take a symmetric PBE. If there is a time interval (t',t'') during which both players discloses continuously, then they discloses continuously for all $t>t'$, and their reputation is a constant on $[t',\infty)$. Moreover, during continuous disclosure,*

- *players exert full effort only if $c\leq c_4$;*
- *players exert interior effort e^* only if $c\in(c_4,c_3)$.*

If players exert full effort during continuous disclosure, then their reputations stay at μ^{**} defined by

$$\mu^{**}:=1-\frac{r+\gamma}{\lambda U/F}. \tag{22}$$

Proof. Step 1. We show that if there is a time interval (t',t'') during which both players disclose continuously, and if, starting from t', $\tilde{t}$ is the earliest time at which at least one player stops disclosing, then for all $t\geq\tilde{t}$, the reputation of the player who stops discloses at $\tilde{t}$ will always be $\hat{\mu}$, where $\hat{\mu}$ is defined by (3).

Starting from time t', let j denote the first player who stops disclosing, and $\tilde{t}$ be the earliest time at which j stops disclosing. Take a small $\epsilon>0$. During the time interval $(\tilde{t}-\epsilon,\tilde{t})$, since both are disclosing continuously, i's payoff satisfies $W_i^1(t)=F+U$, meaning that $\frac{dW_i^1(t)}{dt}=0$. Moreover, it satisfies

$$(r+\lambda+\mu_j(t)\lambda)W_i^1(t)=\lambda(F+S)+R_j(t)(-F)+\frac{dW_i^1(t)}{dt}. \tag{23}$$

We now show that $\mu_j(\tilde{t})=\hat{\mu}$. First, we must have $\mu_j(t)\leq\hat{\mu}$, for all $t\in(t',\tilde{t})$; otherwise, j's revelation rate (given by (8)) would be negative and such a strategy is infeasible. Suppose by negation that $\mu_j(\tilde{t})<\hat{\mu}$, then j's revelation rate right before t is bounded away from 0: $\lim_{t\uparrow\tilde{t}}R_j(t)>0$; (23) then implies that

$$(r+\lambda+\mu_j(\tilde{t})\lambda)W_i^1(\tilde{t}-)<\lambda(F+S). \tag{24}$$

Since j stops disclosure at $\tilde{t}$ (by assumption), at time $t\in(\tilde{t},\tilde{t}+\epsilon')$ for some small $\epsilon'>0$, if i hides F from $\tilde{t}$ until time t, then i's payoff satisfies

$$(r+\lambda+\mu_j(t)\lambda)W_i^1(t)=\lambda(F+S)+\frac{dW_i^1(t)}{dt}. \tag{25}$$

Note that at $\tilde{t}$, we must have $\frac{dW_i^1(\tilde{t}+)}{dt}\geq 0$, because $W_i^1(\tilde{t})=F+U$ and W_i^1 cannot be lower than $F+U$ (as i always can disclose F and gets $F+U$). Thus, the above equation implies that

$$(r+\lambda+\mu_j(\tilde{t})\lambda)W_i^1(\tilde{t}+)\geq\lambda(F+S).$$

This inequality and (24) imply that W_i^1 is discontinuous at $\tilde{t}$, which is impossible, since neither player's revelation strategy can have an atom when both players' reputation levels are strictly positive.

Therefore, we must have $\mu_j(\tilde{t}) = \hat{\mu}$.

Using Lemma 8, we have $W_i^1(t) = F + U$ and $\mu_j(t) = \hat{\mu}$ for all $t \geq \tilde{t}$, that is, j does not disclose after $\tilde{t}$. Since $\mu_j(t) \leq \hat{\mu}$ right before $\tilde{t}$, j's reputation is not decreasing for some interval before $\tilde{t}$. By Lemma 9, we have $\gamma > \hat{\mu}\lambda$. From this inequality and the fact that j does not disclose after $\tilde{t}$ and his reputation stays at $\hat{\mu}$, we have, j's effort is interior (using the law of motion of $\mu_j(t)$, (2)).

Step 2. We show that it is impossible for continuous revelation to lasts until both players' reputation reach $\hat{\mu}$. Step 1 implies that, if there is some $\tilde{t}$ such that $\mu_i(t) = \mu_j(t) = \hat{\mu}$, then neither player will disclose after $\tilde{t}$; moreover, both exert interior effort. Moreover, for both players, we have $W_j^1(t) = F + U$ and $W_j^0(t) = F + U - c/\gamma > 0$ (as we assume that $F + U - c/\gamma > 0$), for all $t \geq \tilde{t}$. Then, since i does not disclose and j exerts interior effort and $dW_j^0(t)/dt = 0$, the HJB equation (11) implies that $W_j^0(t) = 0$. A contradiction.

Remark: Steps 1 and 2 imply that in any symmetric PBE, if there is some time interval during which both disclose continuously, then disclosure must lasts forever with no gaps. We thus have only two possibilities: continuous disclosure continues forever and players' reputation converges to 0, and continuous disclosure continues forever and players' reputation does not converge to 0. In Step 3, we show that players' reputation must be weakly increasing and converges to a positive constant.

Step 3. We show that there is no time interval during which both players disclose continuously, and players' reputation is strictly decreasing.

By Lemma 4, during continuous disclosure, if at least one player puts interior effort, then both players' reputations are a constant. Thus, we only need to prove that there is no time interval during which both players disclose continuously and exert full effort, and at least one player's reputation is strictly decreasing. Suppose by negation that there is such an interval (t', t''). Note that since both players exert effort 1 and disclose continuously during (t', t''), applying (8) to the law of motion of reputation (2), we have

$$d\mu_j(t)/dt = (1 - \mu_j(t))(\gamma + r - (1 - \mu_j(t))\frac{\lambda U}{F}). \tag{26}$$

If j's reputation is decreasing, we have

$$\gamma + r - (1 - \mu_j(t))\frac{\lambda U}{F} < 0. \tag{27}$$

We have shown that once continuous disclosure starts, then disclosure must last forever. Thus, as long as $\mu_j(t) > 0$, we have continuous disclosure, and (27) implies that

$$d\mu_j(t)/dt = (1 - \mu_j(t))(e_j(t)\gamma + r - (1 - \mu_j(t))\frac{\lambda U}{F}) < 0.$$

That is, starting from time t', $\mu_j(t)$ must strictly decrease until it hits 0.

Once $\mu_j(t)$ hits 0, if j discloses immediately, then $\mu_j(t)$ stays at 0; if he discloses continuously, then (8) implies that $\mu_j(t)$ stays at 0. In other words, after $\mu_j(t)$ hits 0, j must disclose immediately. Note that inequality (27) implies that $\lambda U/F > r + \gamma$. Since we have shown that under this condition, if one player discloses immediately, then the other player does not find it optimal to disclose immediately. This means that once players' reputation hits 0, at least one player will stop disclosing. From Step 1, this cannot happen in symmetric PBE.

Step 4. In a symmetric PBE, if both player disclose continuously during (t', t''), then their reputation will be a constant after t'.

If both players put interior effort during (t', t''), then Lemma 4 shows that their reputation is fixed at μ^*. Thus we are done with this step.

For the rest of this step, suppose that both put full effort during (t', t''). Then, j's reputation evolves according to (26). Since Step 3 shows that μ_j is weakly increasing, we have $\mu_j(t) \geq \mu^{**}$, where μ^{**} is such that the right-hand side of (26) is 0. That is, $d\mu_j(t)/dt = 0$ if $\mu_j(t) = \mu^{**}$ in (26). Combining the definition of μ^*, (9), we have, $\mu^{**} > \mu^*$ iff $c < c_4$.

Suppose first that $c \leq c_4$ so that $\mu^{**} \geq \mu^*$. We show that $\mu_j(t) = \mu^{**}$ on (t', ∞). Suppose by negation that $\mu_j(t) > \mu^{**}$ at some $\tilde{t} \geq t'$, then μ_j strictly increases after time $\tilde{t}$. That is, $\mu_j(t) > \mu^{**}$ for all $t > \tilde{t}$. Moreover, since $\mu_j(t) > \mu^{**} \geq \mu^*$, it is impossible for the equilibrium to transit to continuous disclosure with interior effort. And since it is impossible to have nondisclosure after continuous disclosure, we have, for all $t > t'$, both players disclose continuously and exert full effort. But then, $\mu_j(t)$ will exceed $\hat{\mu}$, implying that players' disclosure rate will become negative, which is infeasible.

Suppose next that $c > c_4$ so that $\mu^{**} < \mu^*$. We now show that it is impossible to have continuous disclosure with full effort in this case. Recall the HJB equation (11) as

$$(r + \mu_i(t)\lambda)W_j^0(t) \quad = \quad \max_e e(\gamma(W_j^1(t) - W_j^0(t)) - c) + R_i(t)(U - W_j^0(t)) + \frac{dW_j^0(t)}{dt}. \tag{28}$$

By Lemma 4, suppose there is an interval of time during which continuous disclosure with interior effort occurs; then, players' reputation is μ^*. Let R^* denote the disclosure rate during this interval, defined by (8) with $\mu_i(t) = \mu^*$. Let a player's continuation value before he or she discovers F be W_*^0. We thus have

$$(r + \mu^*\lambda)W_*^0 \quad = \quad R^*(U - W_*^0). \tag{29}$$

We now show that it is impossible to have any interval with continuous disclosure and full effort. Suppose by negation that there is such an interval (t', t''). We have two cases $\mu_i(t) = \mu^{**}$ for all $t > t'$, or $\mu_i(t) > \mu^{**}$ for some $t > t'$.

In the first case, if there is some $\tilde{t}$ such that $\frac{dW_j^0(t)}{dt} = 0$; otherwise, W_j^0 grows unboundedly, which is impossible. At $\tilde{t}$, we have

$$\begin{aligned}(r + \mu^{**}\lambda)W_j^0(t) \quad &= \quad \max_e e(\gamma(W_j^1(t) - W_j^0(t)) - c) + R_i(t)(U - W_j^0(t)) \\ &\geq \quad R_i(t)(U - W_j^0(t)).\end{aligned} \tag{30}$$

But this inequality contradicts (29) because, first, the left-hand side of (30) is strictly lower than that of (29), because $\mu^{**} < \mu^*$ and $W_j^0(t) \leq W_*^0$ (since j is willing to put full effort); second, the right-hand side of (30) is strictly higher than that of (29), because $R_i(t) > R^*$ (as $R_i(t)$ defined by (8) is strictly decreasing $\mu_i(t)$).

In the second case, either $\mu_i(t)$ converges to μ^* from below, after which, both put interior effort (that is, equilibrium transits to continuous disclosure with interior effort); or $\mu_i(t)$ eventually exceeds $\hat{\mu}$, which has been shown to be impossible in previous steps. Suppose at $\tilde{t}$, $\mu_i(t)$ reaches μ^*. We have $dW_j^0(\tilde{t}-)/dt \geq 0$. Then, on a small left neighborhood of $\tilde{t}$, the first and third terms on the right-hand side of (28) is positive, the second term is strictly higher than the right-hand side of (29), due to the same reason as in the first case. Similarly, the left-hand side of (28) is strictly lower than the left-hand side of (29). Thus these two equations cannot hold, meaning that it is impossible for the equilibrium to transit to continuous disclosure with interior effort. This case is impossible.

To summarize, take a symmetric PBE in which both player disclose continuously during some time interval (t', t''). if $c \leq c_4$, then for all $t > t'$, both both player disclose continuously and exert full effort. Step 4 also shows that a symmetric equilibrium has a CD phase with full effort only if $c \leq c_4$. If $c > c_4$, then for all $t > t'$, both both player disclose continuously and exert interior effort. Note that $R^* \leq 0$ if $c \geq c_3$, and $\mu^* \leq 0$ if $c \leq \gamma(r+\lambda)F/\lambda$; that is, if $c \geq c_3$ or $c \leq \gamma(r+\lambda)F/\lambda$, there is no symmetric PBE in which players disclose continuously and exert interior effort. We are done with this lemma.

■

Corollary 4. *A symmetric equilibrium has a CD phase with full effort only if one of the following holds:*

(a) $\frac{\lambda U}{F} > r + \gamma$ *and* $\lambda \leq \gamma$ *and* $c \in [0, c_4]$,

(b) $\frac{\lambda U}{F} \in (r+\gamma, \frac{\lambda}{\lambda-\gamma}(r+\gamma))$ *and* $\lambda > \gamma$ *and* $c \in [0, c_4]$.

Proof. From Step 4 of the proof of Lemma 10, a first necessary condition for a CD phase with full effort to exist is that $c \leq c_4$.

From Step 4 of the proof of Lemma 10, during a CD phase with full effort, each player's reputation stays at μ^{**}, which is defined by (22). Another necessary condition for a CD phase with full effort to exist is $\mu^{**} > 0$. Using (22), we have

$$\mu^{**} > 0 \Longleftrightarrow \frac{\lambda U}{F} > r + \gamma.$$

Thus, the second necessary condition is $\frac{\lambda U}{F} > r + \gamma$.

Finally, we use positive disclosure rate to drive another necessary condition. By Lemma 4, each player's expected revelation rate $R_i(t)$ satisfies (8). Plugging in the expression of μ^{**}, we have

$$R_i(t) = \gamma - \mu^{**}\lambda.$$

Continuous disclosure requires that $R_i(t) > 0$, or equivalently, $\mu^{**} < \gamma/\lambda$. If $\lambda \leq \gamma$, then $\mu^{**} < \gamma/\lambda$ always holds. If $\lambda \leq \gamma$, then $\mu^{**} < \gamma/\lambda$ iff $\frac{\lambda U}{F} > \frac{\lambda}{\lambda-\gamma}(r+\gamma)$.

Combining the three necessary conditions, we are done with the proof. ■

A.4 Nondisclosure PBE

Lemma 11. *In any symmetric PBE, in any time interval during which neither player discloses, each player's reputation is weakly decreasing.*

Proof. We assume that both players' reputations are lower than $\bar{\mu}$, where $\bar{\mu} := \min\{1, \gamma/\lambda\}$, since there is no feasible strategy by which a player's reputation can exceed this level.

Suppose by negation that there is a time interval (t', t'') during which neither player discloses, and that i's reputation is strictly decreasing. Since $\mu_i(t) < \bar{\mu}$, i must exert interior effort during (t', t''). Thus, i's value functions must satisfy have $W_i^1 - W_i^0 = c/\gamma$. Since both players do not disclose, i's value functions also satisfy

$$(r + \mu_j(t)\lambda)W_i^1(t) = \lambda(F + S - W_i^1(t)) + \frac{dW_i^1(t)}{dt}, \tag{31}$$

$$(r + \mu_j(t)\lambda)W_i^0(t) = \frac{dW_i^0(t)}{dt}. \tag{32}$$

Applying $W_i^1(t) - W_i^0(t) = c/\gamma$, we have

$$W_i^1(t) = F + S - (\frac{r}{\lambda} + \mu_j(t))\frac{c}{\gamma}, \quad (33)$$

$$W_i^0(t) = F + S - (\frac{r}{\lambda} + \mu_j(t) + 1)\frac{c}{\gamma}. \quad (34)$$

From (32), if $\frac{dW_i^0(t)}{dt} < 0$, then $W_i^0(t) < 0$, which is impossible in equilibrium. Therefore, $\frac{dW_i^0(t)}{dt} \geq 0$. Using the explicit form of W_i^0 we just derived, we have $\frac{d\mu_j(t)}{dt} \leq 0$ for all t in (t', t''). Since $\mu_j(t) < \bar{\mu}$, j must exert interior effort as well. Therefore, j's value functions satisfy (31) to (34) as well, if we swap j with j in the equations. If $\frac{d\mu_j(t)}{dt} = 0$, then $\frac{dW_i^0(t)}{dt} = 0$ and $W_i^0(t) = 0$.

First, suppose $\frac{d\mu_j(t)}{dt} < 0$ for some $\hat{t}$ in (t', t''), then $W_i^0(\hat{t}) > 0$ and $\frac{dW_i^0(t)}{dt} > 0$ on $(\hat{t}, t'')$. These further imply that $\frac{d\mu_j(t)}{dt} < 0$ on $(\hat{t}, t'')$. Since both W_i^1 and W_j^1 are strictly increasing on $(\hat{t}, t'')$, neither player will disclose. Let $\tilde{t}$ be the earliest time at which there is some $k \in \{1, 2\}$ such that $\frac{dW_k^1(t)}{dt} = 0$. Since W_k^1 is increasing on $(t', \tilde{t})$, we have $\frac{d^2W_k^1(t)}{dt^2} \leq 0$ at $\tilde{t}$. Taking derivative w.r.t t on both sides of (31) and applying $\frac{dW_k^1(t)}{dt} = 0$ at $\tilde{t}$, we have $\frac{d^2W_k^1(t)}{dt^2} = \lambda W_k^1(t)\frac{d\mu_m(t)}{dt}$, for $m \neq k$. Thus, at at $\tilde{t}$, we have $\frac{d\mu_m(t)}{dt} \leq 0$, implying that m exerts interior effort. Going through the same arguments as above, we have, both m's and k's value functions satisfy (31) to (34) around $\tilde{t}$. Then, the fact that $\frac{dW_k^1(\tilde{t})}{dt} = 0$ implies that $\frac{dW_k^0(\tilde{t})}{dt} = 0$. Using (32), we have $W_k^0(\tilde{t}) = 0$. But since $W_k^0(t') \geq 0$ and W_k^1 is strictly increasing on $(t', \tilde{t})$, we have $W_k^1(\tilde{t}) - W_k^0(\tilde{t}) > W_k^1(t') - W_k^0(t') = c/\gamma$. That is, k would strictly prefer to exert full effort. A contradiction.

Second, suppose $\frac{d\mu_j(t)}{dt} = 0$ for all t in (t', t''). We have $W_i^0(t) = 0$ on this interval. Let $\tilde{t}_j$ be the earliest time at which $\frac{dW_j^1(t)}{dt} = 0$. If either $\frac{dW_i^1(t)}{dt} = 0$ on $(t', \tilde{t}_j)$ or there is some $\check{t}$ in $(t', \tilde{t}_j)$ such that $\frac{dW_i^1(t)}{dt} \geq 0$ on $(t', \check{t})$ and $\frac{dW_i^1(\check{t})}{dt} > 0$, then the above proof still applies. We now show that there is no t_1 such that $\frac{dW_i^1(t)}{dt} \leq 0$ on (t', t_1) and $\frac{dW_i^1(t_1)}{dt} < 0$. Suppose by negation that there are t_1, t_2 in $(t', \tilde{t}_j)$ such that $\frac{dW_i^1(t)}{dt} = 0$ on (t', t_1) and $\frac{dW_i^1(t_1)}{dt} < 0$ on (t_1, t_2). Thus, for all t in (t_1, t_2), we have $W_i^1 - W_i^0 < c/\gamma$, meaning that i's effort is 0, and hence $d\mu_i(t) = -\mu_i(t)(1 - \mu_i(t))\lambda dt$.

Then, since j is not disclosing, using (31) and $\frac{dW_i^1(t_1)}{dt} = 0$, we have $\frac{d^2W_i^1(t_1+)}{dt^2} = \lambda W_i^1(t)\frac{d\mu_j(t_+)}{dt}$. And since $\frac{d^2W_i^1(t_1+)}{dt^2} < 0$, we have $\frac{d\mu_j(t_1+)}{dt} < 0$. This implies that j's effort is interior in a right neighborhood of t_1 and hence his value function satisfy (31) to (33). From (33) and $d\mu_i(t) = -\mu_i(t)(1 - \mu_i(t))\lambda dt$, we have $\frac{dW_i^1(t)}{dt} = \mu_i(t)(1 - \mu_i(t))\lambda c/\gamma$. But this cannot satisfy both (33) and (31). A contradiction. ■

Lemma 12. *In any symmetric PBE, if there is a time interval (t', t'') during which both players do not disclose and put interior effort, they continue this forever.*

Proof. Nondisclosure with interior effort on (t', t'') imply that $W_i^0(t) = 0$, and thus $W_i^1(t) = c/\gamma$ on (t', t'').

Lemma 8 implies that $W_i^0(t) = 0$ for all $t \geq t'$. Corollary 3 implies that players' disclosure rate must be 0, that is, neither player disclose, for all $t \geq t'$.

To conclude our proof, we show that there is no $\tilde{t} > t'$ such that during $(t', \tilde{t})$ player exert interior effort and during $(\tilde{t}, \tilde{t} + \epsilon)$ player exert full effort for some $\epsilon > 0$. Suppose by negation that there is such a $\tilde{t}$. Since effort on $(\tilde{t}, \tilde{t} + \epsilon)$ is 1 while it is interior on $(t', \tilde{t})$, players' reputation must strictly increase on $(\tilde{t}, \tilde{t} + \epsilon)$ if ϵ is small enough. Player i's reputation satisfies (31); since on $(t', \tilde{t})$ $dW_i^1(t)/dt = 0$, W_i^1 is weakly lower on $(t', \tilde{t})$ than on $(\tilde{t}, \tilde{t} + \epsilon)$ (otherwise, i will not put effort), and $\mu_j(t)$ is strictly lower on $(t', \tilde{t})$ than on $(\tilde{t}, \tilde{t} + \epsilon)$, we must have $dW_i^1(t)/dt > 0$ on $(\tilde{t}, \tilde{t} + \epsilon)$. We now show that players must put full effort for all $t > \tilde{t}$, which then implies that $dW_i^1(t)/dt > 0$ for all $t > \tilde{t}$ and thus W_i^1 grows unboundedly, which is absurd. Suppose by negation that there is some $\hat{t}$ at which players stop exert full effort (for some time interval). Lemma 11 implies that $\mu_j(\hat{t}) > \mu_j(\tilde{t})$. Moreover, at $\hat{t}$ $dW_i^1(t-)/dt \leq 0$, and $W_i^1(\hat{t}) = W_i^1(\tilde{t}) = c/\gamma$. But then, (31) cannot be satisfied at both $\tilde{t}$ and $\hat{t}$. A contradiction.

■

Corollary 5. *A symmetric equilibrium has a ND phase with interior effort only if one of the following holds:*

(a) $\frac{\lambda U}{F} > r$ *and* $\lambda \leq \gamma$ *and* $c \in (c_2, c_3]$,

(b) $\frac{\lambda U}{F} \in (r, \frac{\lambda}{\lambda-\gamma}(r+\gamma)$ *and* $\lambda > \gamma$ *and* $c \in (c_2, c_3]$,

(c) $\frac{\lambda U}{F} \in (\frac{\lambda}{\lambda-\gamma}(r+\gamma), \infty)$ *and* $\lambda > \gamma$ *and* $c \in (c_5, c_3]$.

Proof. Consider a symmetric PBE that has a ND phase with interior effort during the time interval (t', t''). By Lemma 12, this phase lasts for all $t < t'$.

We first show that $W_i^0(t) = 0$ and $W_i^1(t) = \frac{c}{\gamma}$ for all $t > t'$. During the ND phase with interior effort, i's first stage payoff $W_i^0(t)$ satisfies the following HJB equation.

$$rW_i^0(t) = \max_{e_i} e_i[-c + \gamma(W_i^1(t) - W_i^0(t))] + \lambda[0 - W_i^0(t)] + \frac{dW_i^0(t)}{dt}. \tag{35}$$

Since it is optimal for i to choose interior effort, we have

$$-c + \gamma(W_i^1(t) - W_i^0(t)) = 0. \tag{36}$$

Thus, (35) reduces to

$$(r+\lambda)W_i^0(t) = \frac{dW_i^0(t)}{dt}. \tag{37}$$

For an ND phase with interior effort to lasts forever, we must have $W_i^0(t)$ for all $t > t'$. Indeed, if there were some $\hat{t} > t'$ at which $W_i^0(\hat{t}) > 0$, then the ODE (37) implies that $W_i^0(t)$ will grow to infinity as $t \to \infty$. This is impossible in equilibrium.

Therefore, $W_i^0(t) = \frac{dW_i^0(t)}{dt} = 0$ for all $t > t'$. Combining (36) again, we have $W_i^1(t) = \frac{c}{\gamma}$.

Next, we derive a player's reputation during this ND phase with interior effort. After discovering F, player i's payoff $W_i^1(t)$ should satisfy the following HJB.

$$rW_i^1(t) = \lambda(F + S - W_i^1(t)) + \mu_j(t)\lambda(0 - W_i^1(t)),$$

where we have omitted $\frac{dW_i^1(t)}{dt}$ since $\frac{dW_i^1(t)}{dt} = 0$. We thus have

$$\Rightarrow W_i^1(t) = \frac{\lambda(F+S)}{r + \lambda + \mu_j(t)\lambda}$$

Since $W_i^1(t) = \frac{c}{\gamma}$, j's reputation satisfies

$$\mu_j(t) = \frac{\gamma(F+S)}{c} - \frac{r}{\lambda} - 1 \equiv \mu_{ND}, \forall t > t'. \tag{38}$$

We have

$$\mu_{ND} \geq 0 \Longleftrightarrow c \leq \frac{\lambda\gamma(F+S)}{r+\lambda} \equiv c_3. \tag{39}$$

Then, we derive a player's effort during this ND phase with interior effort. From (38), we have $d\mu_j(t) = 0$. Using the law of motion of a player's reputation (2), this implies that

$$e_j(t)\gamma = \mu_{ND}\lambda.$$

Using (38), we have

$$e_j(t) = \frac{\lambda(F+S)}{c} - \frac{r+\lambda}{\gamma} \equiv e_{ND}, \forall t > t'. \tag{40}$$

We now derive necessary conditions for effort to be interior. We have,

$$e_i \geq 0 \iff c \leq c_3. \tag{41}$$

$$e_i \leq 1 \iff c \geq \frac{\lambda\gamma(F+S)}{r+\lambda+\gamma} \equiv c_5. \tag{42}$$

Finally, we need to ensure that given the other player is not disclosing, it is a best response of i to not disclose. This requires $F + U \leq W_i^1(t)$. Since $W_i^1(t) = \frac{c}{\gamma}$, nondisclosure is a mutual best reply if

$$c \geq \gamma(F+U) = c_2. \tag{43}$$

We now compare the two lower bounds in (42) and (43). We have $c_2 > c_5$ iff $(\lambda - \gamma)\frac{\lambda U}{F} < \lambda(r+\gamma)$. Thus, (43) is tighter iff $\lambda \leq \gamma$ or $\lambda > \gamma$ and $\frac{\lambda U}{F} < \frac{\lambda}{\lambda-\gamma}(r+\gamma)$. In such a case, a necessary condition for a symmetric equilibrium to have a ND phase with interior effort is $c \in (c_2, c_3)$. (42) is tighter iff $\lambda > \gamma$ and $\frac{\lambda U}{F} < \frac{\lambda}{\lambda-\gamma}(r+\gamma)$. In such a case, a necessary condition for a symmetric equilibrium to have a ND phase with interior effort is $c \in (c_5, c_3)$.

Using the same reasoning as in the proof of Corollary 1, another necessary condition for a symmetric equilibrium to have a ND phase with interior effort is $\frac{\lambda U}{F} > r$. We are done with the proof. ■

Lemma 13. *Take a symmetric equilibrium. If there is a t' such that after time t', the equilibrium is nondisclosure with full effort, then we must have $\frac{\lambda U}{F} \geq \frac{\lambda}{\lambda-\gamma}(r+\gamma)$, $\gamma < \lambda$, and $c \leq c_5$.*

Proof. After time t', since players do not disclose and exert full effort, by (2), j's reputation $\mu_j(\cdot)$ satisfies the law of motion

$$d\mu_j(t) = (1-\mu_j(t))(\gamma - \mu_j(t)\lambda)dt. \tag{44}$$

And since $\mu_j(t) < \gamma/\lambda$, $\mu_j(\cdot)$ is weakly increasing and converges to $\bar{\mu} \equiv \min\{1, \gamma/\lambda\}$ as $t \to \infty$.

Player j's value function satisfies $W_j^1(t) \geq F + U$ and the HJB equation

$$(r+\lambda+\mu_i(t)\lambda)W_j^1(t) = \lambda(F+S) + \frac{dW_j^1(t)}{dt}. \tag{45}$$

Taking derivative on both sides, we have

$$(r+\lambda+\mu_i(t)\lambda)\frac{dW_j^1(t)}{dt} + \frac{d\mu_j(t)}{dt}W_j^1(t) = \frac{d^2W_j^1(t)}{dt^2}. \tag{46}$$

Taking limit as $t \to \infty$ and using $\mu_j(\cdot)$ converges to $\bar{\mu}$, we have,

$$(r+\lambda+\bar{\mu}\lambda)\lim_{t\to\infty}\frac{dW_j^1(t)}{dt} = \lim_{t\to\infty}\frac{d^2W_j^1(t)}{dt^2}. \tag{47}$$

We must have $\lim_{t\to\infty} \frac{dW_j^1(t)}{dt} \to 0$ because otherwise, (47) implies that $\lim_{t\to\infty} \frac{dW_j^1(t)}{dt}$ will grow unbounded, and by (46), $W_j^1(t)$ will also grow unbounded, which is impossible in equilibrium. Taking limit on both sides of (46) and plugging in $\lim_{t\to\infty} \mu_j(t) \to \bar{\mu}$ and $\lim_{t\to\infty} \frac{dW_j^1(t)}{dt} \to 0$, we have

$$\lim_{t\to\infty} W_j^1(t) = \frac{\lambda(F+S)}{r+\lambda+\bar{\mu}\lambda}. \tag{48}$$

Suppose $\gamma \geq \lambda$. This means $\bar{\mu} = 1$. Thus, j's equilibrium payoff (from not revealing) is $\frac{\lambda(F+S)}{r+2\lambda}$, which is strictly lower than the payoff from disclosing, $F+U$, if we use $U = \frac{\lambda S}{r+2\lambda}$. Therefore, it is impossible to have a ND phase with full effort from t' onward in this case.

Suppose $\gamma < \lambda$. This means that $\bar{\mu} = \frac{\gamma}{\lambda}$. In this case, j's equilibrium payoff (from not revealing) is $\frac{\lambda(F+S)}{r+\lambda+\gamma}$. Not revealing constitutes a best response only if $\frac{\lambda(F+S)}{r+\lambda+\gamma} \geq F+U$, or equivalently, $\frac{\lambda U}{F} > \frac{\lambda}{\lambda-\gamma}(r+\gamma)$.

Finally, player j is willing to put full effort in stage 1 only if $\gamma\left(\frac{\lambda(F+S)}{r+\lambda+\gamma}\right) \geq c$, or equivalently, $c \leq c_5$.

In sum, if a symmetric PBE has a ND phase with full effort that lasts forever, we must have $\frac{\lambda U}{F} \geq \frac{\lambda}{\lambda-\gamma}(r+\gamma)$, $\gamma < \lambda$, and $c \leq c_5$. ■

The above lemmas imply that we have only five types of symmetric PBE:

- players immediately disclosure at all t; such an equilibrium exists only if
- CD equilibria: there is some time T such that before T, both players put full effort and do not disclose, and after T, both players disclose continuously with a constant effort and reputation level; such an equilibrium exists only if
- ND equilibria with interior effort: there is some time T such that before T, both players put full effort and do not disclose, and after T, both players put interior effort and do not disclose; such an equilibrium exists only if
- ND equilibria with full effort: both players put full effort and do not disclose at all t; such an equilibrium exists only if

B Proof of Theorem 1

Proof. First consider $c \in (\frac{\gamma(r+\lambda)}{\lambda}F, \bar{c})$.

There exists $\underline{S}c > 0$ such that below this threshold either the equilibrium rate of disclosure is zero or both the first stage equilibrium effort and the rate of disclosure are zero. As S increases above this threshold, we are initially in the CD phase with interior effort in stage 1 and gradually enter CD with full effort in stage 1 and to ND (if $\lambda > \gamma$ with full effort in stage 1. An increase in S increases the equilibrium disclosure rate y in the GR phase without changing the length of the ND phase. Hence equilibrium disclosure rate is increasing in S.

For each c, there exists a $\tilde{S}(c)$ such that for any S above this threshold, the equilibrium is CD with full effort is stage 1. We know that an increase in S decreases the equilibrium disclosure rate y in the CD phase and lengthens the ND phase. Hence, for $S \geq \tilde{S}(c)$, an increase in S reduces equilibrium disclosure. When $\lambda > \gamma$, as S increases, eventually we enter the region where equilibrium is non-disclosure with full effort in stage 1. In this case, S has no effect on the expected time to discovery. ■

C Proof of Theorem 2

Proof. First, suppose for a given c, for a range of values of U, the equilibrium is immediate disclosure with interior effort in stage 1. Recall from the equilibrium description that the effort in the first stage is given by

$$e_{ID} = \frac{r}{\gamma}[\frac{\gamma(F+U)-c}{c-\gamma F}]$$

Since e_{ID} is strictly increasing in U and the disclosure remains immediate, the expected time to final success is decreasing in U, i.e as S goes up, the final discovery is expected to take sooner.

Next, suppose for a given c, for a range of values of U the equilibrium is continuous disclosure with interior efforts. In a CD equilibrium with interior effort in stage 1, we know that the steady state belief μ^* satisfies $\mu^* = 1 - \frac{\gamma(r+\lambda)F}{\lambda c}$. This belief in unaffected by an increase in U. This means the time to enter the CD phase from the non-disclosure phase is also unaffected. Recall that the effort in first stage satisfies $e\gamma = \frac{\gamma(r+\lambda)U}{c} - r$. Since this is strictly increasing U, we can conclude that an increase in U increases the level of effort in the first stage.

We know that $\mu^* y = \frac{\lambda U}{F}(1-\mu^*) - r - \mu^*\lambda$. Since μ^* is unaffected by an increase in U, we can say that equilibrium y goes up. Thus, an increase in U increases the first stage effort and the equilibrium disclosure rate in the CD phase and does not alter the length of the non-disclosure phase. Hence we can conclude that for this rage of U, $\frac{dE(t_2)}{dU} < 0$. On the other hand if for a given c and a range of values of U, if the equilibrium is continuous disclosure with full effort in stage 1, the effect of increasing U on the expected time to discovery would be opposite. In the CD phase, the long run belief μ^{**} is strictly increasing in U. This means, following an increase in U, the length of the non-disclosure phase goes up. Since we have $\mu y = \gamma - \mu^{**}\lambda$, following an increase in U, equilibrium disclosure rate in the CD phase goes down. Both these effects imply that the expected time to success is longer and hence, $\frac{dE(t_2)}{dU} > 0$.

Finally, if for a given c and for a range of values of U the equilibrium is non-disclosure with interior effort, then since we have $e_{ND} = \frac{\lambda(F+S)}{c} - \frac{(r+\lambda)}{\gamma}$, we know that an increase in U leads to an increase in first stage effort. Increase in U has no effect if the equilibrium is non-disclosure with full effort in stage 1.

These results will allow us to prove the theorem. First, consider $c \in (0, \gamma F)$.In this case for $S = 0$, the equilibrium is ID with full effort is stage 1 and as S increases, we gradually enter the CD and ND (if $\gamma < \lambda$) with full effort is stage 1. This means for any $S \geq 0$, increase in S can weakly delay the expected time to final success. Hence, $\tilde{S}(c) = 0$.

Next, consider $c > \bar{c}$. This exists only when $\lambda > \gamma$ In this case before a threshold S is reached, the equilibrium effort in stage 1 and disclosure rate are both 0. Once the value of S crosses a threshold, we start having ND equilibrium with interior efforts. An increase in S in this range leads to an increase in first stage equilibrium effort and hence brings down the expected time to discovery. As S keeps increasing, we enter the ND phase with full effort. Thus for any $S < \infty$, an increase in S brings the expected time to final discovery closer.

Finally, we consider $c \in (\gamma F, \bar{c})$.In this range for each value of c, there exists a $\underline{S}(c)$ such that for all $S < \underline{S}(c)$, equilibrium effort is stage 1 and the rate of disclosure is zero.As S increases from $\underline{S}(c)$, we either gradually move from an ID phase with interior effort to ID phase with full effort to CD and ND phase with full effort, or from a ND phase with interior effort to CD phase with interior effort to CD and ND phase with full effort. In either case, until we enter the CD phase with full effort, increase in S brings forward the expected time to discovery. Once the CD phase with full effort becomes relevant, an increase in S delays the expected time to discovery. This shows that $\tilde{S}(c) > 0$. ■

D Proof of Proposition 2

The main proposition on Mandatory transparency is as follows

Proposition 3. *Mandatory Transparency and the final reward are policy complements. That is, given a $c > 0$, there exists a unique $U\hat{(c)}$ such that when $U > (<) U\hat{(c)}$, mandatory transparency reduces (increases) the expected time to final success.*

Proof. We will refer to mandatory transparency, non-disclosure and continuous disclosure as MT, CD and ND respectively.

First, suppose we are in the region when the cost $c < \frac{\gamma F(r+\lambda)}{\lambda}$. In this case, implementation of MT has no effect on the time to final discovery as long as $\frac{\lambda U}{F} \leq r + \gamma$. For $\frac{\lambda U}{F} > r + \gamma$, in the absence of MT, equilibrium effort in the first stage is equal to 1, and we have either ND or CD. Since the implementation of MT implies unit effort in stage 1, the expected time to final success goes down as a result of MT. In this case, our $U\hat{(c)} = \frac{F}{\lambda}(r + \gamma)$.

The non-trivial region is where given U and c, we have $c > \frac{\gamma F(r+\lambda)}{\lambda}$, and $U \in [\underline{U}(c), \bar{U}(c)]$ such that $\bar{U}(c)$ and $\underline{U}(c)$ satisfy $c = \gamma[F + \frac{r\bar{U}(c)}{r+\gamma}]$, and $c = \gamma[F + \underline{U}(c)]$ respectively. When $U > U\bar{(c)}$, in the absence of MT, the first stage equilibrium effort is equal to 1, but there is CD or ND. Hence, implementation of MT unambiguously reduces the expected time to final success. On the other hand, when $U < U\hat{(c)}$, in the absence of MT we have ND in equilibrium with interior efforts in stage 1. Implementation of MT implies zero first stage effort level. Thus, MT unambiguously increase the expected time to final success.

We will now prove the following three lemmas.

Lemma 14. *Suppose parameters are such that without any intervention, in equilibrium, there is ND with unit effort in stage 1. Let $E(T_2|ND)$ be the expected time required to get the second success. Let $E(T_2|MT)$ be the expected time to get the second success if mandatory transparency is introduced. Then, conditional on remaining in the same parameter region,*

$$\frac{\partial[E(T_2|MT) - E(T_2|ND)]}{\partial U} < 0$$

Proof. Introduction of mandatory transparency means, the equilibrium effort level of both players in stage 1 is equal to that in the ID equilibrium with interior efforts. This level e is given by $e\gamma = r[\frac{\gamma(F+U)-c}{c-\gamma F}]$. From the expression of e, we can conclude that $\frac{\partial e}{\partial U} > 0$. This implies $\frac{\partial E(T_2|MT)}{\partial U} < 0$.

In a ND equilibrium with unit effort in the first stage we have $\frac{\partial E(T_2|ND)}{\partial U} = 0$. This concludes the proof of the lemma. ∎

Lemma 15. *Suppose parameters are such that in the absence of any intervention, the equilibrium is CD with full effort in stage 1. Let $E(T_2|CD_f)$ be the expected time to achieve the second success in the absence of any intervention. Let $E(T_2|MT)$ be the expected time to achieve the second success under mandatory transparency. Then, conditional on remaining in the same parameter region,*

$$\frac{d[E(T_2|MT) - E(T_2|GR_f)]}{dU} < 0$$

Proof. We know that in CD equilibrium, before entering the CD phase, there is a ND phase. Let T_f be the time required to enter the CD phase.

In the CD phase, the steady state belief μ^{**} is given by $\mu^{**} = 1 - \frac{(r+\gamma)F}{\lambda U}$. The equilibrium revelation rate y satisfies $\mu^{**}y = \gamma - \mu^{**}\lambda$. Thus, μ^{**} is strictly increasing in U, and y is strictly decreasing in U. First stage effort is not altered by a change in U as in both the ND and CD phase the first stage effort is equal to 1.

An increase in U leads to an increase in μ^{**} . Thus, the length of the NR phase goes up, and hence for an increased length of time there will be no revelation of intermediate success. This effect leads to an increase in expected time to success. Further, an increase in U leads to a decrease in y. This also leads to an increase in the expected time to success. Thus, we can conclude that $\frac{dE(T_2|CD_f)}{dU} > 0$. On the other hand, an increase in U leads to an increase in the first stage effort under mandatory transparency, implying $\frac{dE(T_2|MT)}{dU} < 0$. Thus we have

$$\frac{d[E(T_2|MT) - E(T_2|CD_f)]}{dU} < 0$$

This concludes the proof of the lemma. ■

Lemma 16. *Suppose parameters are such that in the absence of any intervention, the equilibrium is CD with interior effort in stage 1. Let $E(T_2|CD_i)$ be the expected time to achieve the second success in the absence of any intervention. Let $E(T_2|MT)$ be the expected time to achieve the second success. In the current parameter region, if there exists a $U = \hat{U}$ such that $E(T_2|MT) \geq E(T_2|CD_i)$. Then we have*

$$\frac{d[E(T_2|MT) - E(T_2|CD_i)]}{dU} < 0$$

for all $U \leq \hat{U}$.

Proof. Under Mandatory transparency, we have

$$E(T_2|MT) = \frac{1}{2e\gamma} + \frac{1}{2\lambda}$$

where

$$e\gamma = r[\frac{\gamma(F+U) - c}{c - \gamma F}] = \frac{r\gamma U}{c - \gamma F} - r$$

Suppose we start from the CD phase. In that case,

$$E_1(T_2|CD_i) = \frac{1}{2e^*\gamma} + \frac{1}{2\lambda} + \frac{1}{2(e^*\gamma + y + \lambda)}$$

where

$$e^*\gamma = \gamma\frac{(r+\lambda)U}{c} - r$$

Let μ^* be the stationary belief in the CD phase. We know that μ^* is given by $\frac{\lambda c - \gamma(r+\lambda)}{\lambda c}$.

From the expressions of the effort levels during MT and CD phase we have

$$e^*\gamma - e\gamma = \gamma U[\frac{\lambda c - \gamma F(r+\lambda)}{c(c-\gamma F)} = \lambda\mu^*[\frac{\gamma U}{c - \gamma F}]$$

$$\Rightarrow e^*\gamma - e\gamma - \lambda\mu^* = \lambda\mu^*[\frac{\gamma U}{c-\gamma F} - 1] = \lambda\mu^*\frac{e\gamma}{r}$$

$$\Rightarrow e^*\gamma = e\gamma[\frac{\lambda\mu^*}{r} + 1] + \lambda\mu^*$$

$$\Rightarrow \frac{e^*}{e} = [\frac{\lambda\mu^*}{r} + 1] + \frac{\lambda\mu^*}{e\gamma} \tag{49}$$

Since, e is strictly increasing in U, from (49) we can conclude that $\frac{e^*}{e}$ is strictly decreasing in U. This implies, $\log \frac{e^*}{e}$ is decreasing in U. This gives us

$$\frac{d \log \frac{e^*}{e}}{dU} < 0$$

$$\Rightarrow \frac{e'(U)}{e} > \frac{{e^*}'(U)}{e^*} \tag{50}$$

Next, we will determine $E(T_2|CD_i)$, the expected time required without the introduction of mandatory transparency.

Let T be the time required to enter the CD phase. Let t_1 be the time of first success in stage 1, and let x be the additional time required for the first success in stage 2 after the first success in stage 1 has arrived. Thus, if T_2 is the time required to achieve the first success in stage 2, then we have $t_2 = t_1 + x \Rightarrow E(T_2) = E(t_1) + E(x)$.

We will first compute the density of x under the following different contingencies:

- $t_1 > T$: Let $f(x)$ be the density of x in this case. $f(x)$ is computed on the basis that either the first player who obtained the first stage success will get the second stage success first, or player 2 will achieve success in both stages, and will be the first player to achieve the second stage success.

$$f(x) = \frac{\exp^{-x(e\gamma+y+2\lambda)}(2\exp^{x(e\gamma+y)}\lambda(e\gamma+y) - \exp^{x\lambda}(e\gamma+y+\lambda)}{e\gamma+y-\lambda}.$$

  In this case, we have

$$E(x|t_1 > T) = \int_{x=0}^{\infty} x f(x)\, dx = \frac{1}{2\lambda} + \frac{1}{2(e^*\gamma+y+\lambda)}.$$

- $t_1 < T$: In this case there are two possibilities. Either $x < (T - t_1)$ which means the first success in second stage has occurred before entering the CD phase. Let $g(x)$ be the density of this event.

$$g(x) = \frac{\exp^{-x(\gamma+2\lambda)}(2\exp^{x\lambda}\lambda\gamma - \exp^{x\lambda}\lambda(\gamma+\lambda)}{\gamma-\lambda}.$$

  The other possibility is that the first success in stage 2 has occurred either before entering the CD phase or after entering the CD phase. Let $h(x)$ be the density of this event. This is obtained by taking the derivative of the CDF $P(x < t : t > (T - t_1))$.

$$h(x, T, t_1) = \frac{\exp^{-T\gamma - x\lambda - (t_1+x)(e\gamma+y+\lambda)}\lambda[a+b+c}{(\gamma-\lambda)(-e\gamma-y+\lambda)}$$

  where

$$a = -2\exp^{T\gamma + x(e\gamma+y)+t_1(e\gamma+y+\lambda)}\gamma(e\gamma+y-\lambda)$$

$$b = 2\exp^{x(e\gamma+y)+t_1(\gamma+e\gamma+y)+T\lambda}(-\gamma+e\gamma+y)\lambda$$

$$c = \exp^{T(e\gamma+y)+x\lambda+t_1(\gamma+\lambda)}(\gamma-\lambda)(e\gamma+y+\lambda)$$

Hence, we have

$$E(x|t_1 < T) = \int_{x=0}^{x=T-t_1} xg(x)\,dx + \int_{x=T-t_1}^{x=\infty} xh(x)\,dx = F(T,t_1) + \frac{e^{-(T-t_1)(\gamma+\lambda)}}{2(e^*\gamma + y + \lambda)}$$

One can show that for any $t_1 < T$, $F(T,t_1) > \frac{1}{2\lambda}$.
Direct computation shows that

$$E(t_1) = (1 - e^{-2T\gamma})\frac{1}{2\gamma} + e^{-2T\gamma}\frac{1}{2e^*\gamma}$$

Also, from our above calculations, we have

$$E(x) = e^{-2T\gamma}[\frac{1}{2\lambda} + \frac{1}{2(e^*\gamma + y + \lambda)}]$$

$$+2\gamma \int_{t_1=0}^{t_1=T} e^{-2t_1\gamma}\{F(T,t_1) + \frac{e^{-(T-t_1)(\gamma+\lambda)}}{2(e^*\gamma + y + \lambda)}\}\,dt_1$$

Thus, we have

$$E(t_2) = E(t_1) + E(x)$$

$$= (1 - e^{-2T\gamma})\frac{1}{2\gamma} + e^{-2T\gamma}\frac{1}{2\lambda} + \int_{t_1=0}^{t_1=T} e^{-2t_1\gamma} 2\gamma F(T,t_1)\,dt_1$$

$$+e^{-2T\gamma}(\frac{1}{2e^*\gamma}) + (\frac{1}{2(e^*\gamma + y + \lambda)})[e^{-2T\gamma} + \int_0^T e^{-2t_1\gamma} 2\gamma e^{-(T-t_1)(\gamma+\lambda)}\,dt_1]$$

$$= e^{-2T\gamma}(\frac{1}{2e^*\gamma}) + (\frac{1}{2(e^*\gamma + y + \lambda)})[e^{-2T\gamma} + \int_0^T e^{-2t_1\gamma} 2\gamma e^{-(T-t_1)(\gamma+\lambda)}\,dt_1] + \bar{F}(T)$$

$$= \frac{B_1}{2e^*\gamma} + \frac{B_2}{2(e^*\gamma + y + \lambda)} + \bar{F}(T)$$

where $B_1 = e^{-2T\gamma}$ and $B_2 = e^{-2T\gamma} + \int_0^T e^{-2t_1\gamma} 2\gamma e^{-(T-t_1)(\gamma+\lambda)}\,dt_1$. From the dynamics of the belief, we know that at the steady state belief μ^*, we have

$$\lambda + y = \frac{e^*\gamma}{\mu^*}$$

$$\Rightarrow e^*\gamma + y + \lambda = \frac{e^*\gamma(1+\mu^*)}{\mu^*}$$

Thus we have

$$E(T_2|GR_i) = \frac{B_1(1+\mu^*) + B_2\mu^*}{2e^*\gamma(1+\mu^*)} = \frac{\bar{B}}{2e^*\gamma} + \bar{F}(T)$$

This gives us

$$\frac{d[E(T_2|MT) - E(T_2|CD_i)]}{dU} = -\frac{1}{2\gamma e^2}e'(U) + \frac{\bar{B}}{2\gamma e^{*2}}e^{*'}(U) = -\frac{1}{2\gamma e}\frac{e'(U)}{e} + \frac{\bar{B}}{2\gamma e^*}\frac{e^{*'}(U)}{e^*}$$

$$< [-\frac{1}{2\gamma e} + \frac{\bar{B}}{2\gamma e^*}]\frac{e'(U)}{e}$$

as $\frac{e^{'}(U)}{e} > \frac{e^{*'}(U)}{e^*} > 0$. Suppose there exists a $\hat{U}$ such that $E(T_2|MT) \geq E(T_2|GR_i)$. This means at $\hat{U}$, we have

$$\frac{1}{2e\gamma} + \frac{1}{2\lambda} \geq \frac{\bar{B}}{2e^*\gamma} + \bar{F}(T)$$

Direct computation shows that $\bar{F}(T) > \frac{1}{2\lambda}$. Hence, for the above inequality to hold, at $U = \hat{U}$, we must have $\frac{1}{2e\gamma} > \frac{\bar{B}}{2e^*\gamma}$. Thus we can conclude that at $U = \hat{U}$ we have

$$\frac{d[E(T_2|MT) - E(T_2|GR_i)]}{dU} < [-\frac{1}{2\gamma e} + \frac{\bar{B}}{2\gamma e^*}]\frac{e^{'}(U)}{e} < 0$$

At $U = \hat{U}$, we have $\frac{1}{2e\gamma} > \frac{\bar{B}}{2e^*\gamma} \Rightarrow \frac{e^*}{e} > \bar{B}$. Since $\bar{B}$ is independent of U, and $\frac{e^*}{e}$ is strictly decreasing in U, for all $U < \hat{U}$, we will continue to have $\frac{e^*}{e} > \bar{B} \Rightarrow \frac{1}{2e\gamma} > \frac{\bar{B}}{2e^*\gamma}$. Thus, we can conclude that for all $U \leq \hat{U}$, we will continue to have

$$\frac{d[E(T_2|MT) - E(T_2|CD_i)]}{dU} < [-\frac{1}{2\gamma e} + \frac{\bar{B}}{2\gamma e^*}]\frac{e^{'}(U)}{e} < 0$$

This concludes the proof of the lemma.

■

Consider $U = \bar{U}(c)$. At this point, introduction of MT does not affect the first stage effort which is equal to 1. However, it makes the intermediate revelation rate higher. Hence introduction of MT reduces the expected time to the second success implying MT is beneficial.

On the other hand consider $U = \underline{U}(c)$. At the right neighborhood of this point, the ratio of the first stage effort without MT and with MT is given by $\frac{e^*}{e} \to \infty$. This implies introduction of MT reduces the expected time to final success. Hence there must exist $\hat{U}_1, \hat{U}_2 \in (\underline{U}(c), \bar{U}(c))$ such that $\hat{U}_1 \geq \hat{U}_2$, for all $U < \hat{U}_2$ MT increases the expected time to final success, and for all $U > \hat{U}_1$, MT reduces the expected time to final success. Also at both $\hat{U}_1$ and $\hat{U}_2$, $E(T_2|MT) = (T_2|NMT)$, where NMT means no-mandatory transparency.

We shall now prove that $\hat{U}_1 = \hat{U}_2$.

Suppose this $\hat{U}$ is not unique. Let $\hat{U} = \sup\{U \in (\underline{U}(c), \bar{U}(c)) : E(T_2|MT) = E(T_2|NMT)\}$. For all $U > \hat{U}$, $E(t_2|MT) < E(t_2|NMT)$.

We need to consider two cases. First, suppose $\hat{U}$ is in the region such that without MT the equilibrium is either NR or GR with full effort in stage 1. From lemma (14) and (15) we can conclude that for all $U \in [\tilde{U}, \hat{U})$, $E(t_2|MT) > E(t_2|NMT)$ where $\tilde{U}$ is such that $c = \frac{\gamma(r+\lambda)}{(r+\gamma)}\tilde{U}$.

Since at $U = \tilde{U}$, the equilibrium without MT continuously transforms from a GR with interior effort to a GR with full effort, at a left neighborhood of $\tilde{U}$, we will continue to have $E(t_2|MT) > E(t_2|NMT)$. Using lemma (16) we can conclude that for all $U < \tilde{U}$, $E(t_2|MT) > E(t_2|NMT)$. Hence, for all $U < \hat{U}$ we have $E(t_2|MT) > E(t_2|NMT)$. Thus $\hat{U}$ is unique.

Next, consider $\hat{U} < \tilde{U}$. Again using lemma (16) we have $E(t_2|MT) > E(t_2|NMT)$ for all $U < \hat{U}$. Hence, $\hat{U}$ is unique.

This concludes the proof of the proposition ■