Intelligence Breakthroughs

Peter Bils* Gleason Judd† Bradley C. Smith‡

April 5, 2024

Abstract

We study a model of intelligence breakthroughs in a dynamic setting. An intelligence breakthrough occurs when an actor secretly obtains the capability to observe the behavior of an opponent, and this capability can be used to gain an advantage in conflict. We model this interaction as a game of asymmetric information in which a "Spy," is either able to observe the actions of a "Target" or not. The model captures key tradeoffs of wartime intelligence and espionage. If the target knows the spy is not observing, they prefer to communicate honestly. But if the Target thinks the Spy may be observing, then they are inclined to hedge their communications due to concerns about the spy leveraging its knowledge. In equilibrium, a listening Spy will always "bide its time" and induce the Target to communicate more openly, before eventually exploiting that honesty.

*Vanderbilt University, (peter.h.bils@vanderbilt.edu).
†Princeton University, (gleason.judd@princeton.edu).
‡Vanderbilt University, (bradley.carl.smith@gmail.com)

"The hardest part wasn't returning his serve — it was not letting him know that I knew this. I had to resist the temptation of reading his serve for the majority of the match and choose the moment when I was gonna use that information on a given point to execute a shot that would allow me to break the match open." — Andre Agassi

Secret intelligence is widely coveted. Whether in military conflict, cybersecurity, regulation, or even tennis, people want to have *information about others without their knowledge* in order to get "one step ahead." And for those who succeed, such intelligence provides critical advantages for anticipating how others will behave and react. Accordingly, espionage and other forms of intelligence gathering have attracted substantial attention and resources throughout human history.

Despite the clear appeal of *having* secret intelligence, however, there are murkier incentives for *using* it. On one hand, there is obviously upside in exploiting it in order to realize its potential. On the other hand, acting on it can erode its value going forward, since others may grow suspicious and more cautious. Thus, the possession of secret intelligence raises a dilemma: exploit it today but potentially diminish its usefulness tomorrow, versus conceal it today but preserve its usefulness tomorrow. Essentially, the more pressing decision for intelligence holders is not *whether* to act on intelligence, but *when* to act on it. We aim to shed new light on this fundamental tension.

We study the use of secret intelligence. Broadly, we probe how actors want to exploit *intelligence breakthroughs*—i.e., secretly obtaining the capability to observe others' private behavior. Specifically, our main question asks: when to exploit an intelligence breakthrough?

To do so, we analyze a game-theoretic model of a Target and Spy interacting over time. In each period, Target privately observes information about its current conditions and then takes an action. Meanwhile, the Spy may be secretly observing the Target's actions and, regardless, chooses how boldly to act towards Target. The Spy wants to act boldly in the same direction of Target's information. In contrast, the Target wants to match its own conditions while misdirecting the Spy to act in the wrong direction.

The model captures key tradeoffs of wartime intelligence and espionage. First, the Target wants to match its conditions but is worried about whether the Spy can observe those actions, since it could use such informative behavior to infer the underlying conditions. Second, the Spy wants to use its intelligence to act boldly in the correct direction but that can make the Target more suspicious and then act less sincerely in the future.

These tradeoffs combine to shape the dynamic incentives of both players. The Spy has an incentive to *feign ignorance* in order to appear as though it has not made an intelligence breakthrough. The longer Spy waits, Target grows more confident that there has not been a breakthrough and therefore they are more inclined to act boldly. In turn, a listening spy has increasingly strong incentives to exploit its information.

In equilibrium, behavior is deliberately unpredictable. Spy will 'bide its time' before striking, so repeated rounds of no action do not mean that Spy is not listening. Instead, in each period, Spy either (i) acts boldly and is 'found out' or (ii) does not act and the Target gets less suspicious about Spy's secret intel. In turn, Target gets more likely to send truthful message. But consequently, acting boldly grows more appealing for a listening spy.

Contributions to Related Literature

We contribute to theoretical understanding of secret intelligence and espionage. In various settings, there is an important feedback between incentives to gather secret intel and incentives to mitigate the fallout from being spied on. Essentially, some actors may want to get information about what others know or what they are planning to do but, in turn, that possibility can induce potential targets to obfuscate their plans, behavior, or information (Solan and Yariv, 2004). Parsing this feedback is important for understanding incentives for intelligence gathering. In static settings, it shapes whether actors want to spy (Solan and Yariv, 2004). We complement those insights by studying how these incentives unfold over time to shape *when* actors want to act on their intelligence sources. Moreover, our results shed light on whether actors want to set up a durable source of secret intelligence.

We probe the role of reputation in managing and exploiting a source of intelligence over time. In many situations, reputation is an important consideration because of how it can impact others' behavior (Kreps and Wilson, 1982; Milgrom and Roberts, 1982) Typically, those who are concerned about their reputation want to leverage it for some instrumental purpose (Morris, 2001). For instance, firms may want to build reputations for quality in order to attract customers or exploit them later on (Bar-Isaac, 2003; Bar-Isaac and Tadelis, 2008). Similarly, advisors may want to build trust with their clients by acting responsibly, but their incentives to exploit that credibility can impair their efforts to do so (Sobel, 1985; Benabou and Laroque, 1992). In equilibrium, early trust-building typically precedes exploitation, which destroys trust thereafter.

We uncover some familiar dynamics of reputation and behavior, but in a new setting. Rather than quality or credibility, we study a known adversary who may have a source of secret intel and wants to conceal it. Here, the Spy sits on information for a while in order to appear ignorant, which induces the Target to act on its information more openly and eventually increases Spy's temptation to exploit its intelligence enough that she acts on it, which thereafter destroys the illusion of ignorance. Thus, we focus on the strategic aspects of *bad reputations* (Ely and Välimäki, 2003) since Spy wants to avoid a certain reputation (for spying), rather than seek a certain reputation (e.g., for quality or reliability). [1] Furthermore, reflecting our application, we analyze the evolution of reputation between two long-lived players with two-sided asymmetric information. Specifically, each player has a distinct information advantage over the other, since Target privately observes the state of the world but does not know whether they are being spied on.

We also contribute to theoretical understanding of monitoring.[2] Broadly, we trace the dynamics and consequences of *uncertain monitoring* in a prolonged and adversarial relationship. Our setting features a strategic tension that is inherent in monitoring: it can affect how others behave if they think they are being watched, and that prospect can impact incentives to monitor and use any information obtain by doing so. This tension shapes the ways that actors would optimally monitor if they could commit (Tan, 2023) and how they do monitor if they cannot (e.g.,

[1] There are several other differences. For instance, in Sobel (1985) an advisor provides binary information before the decision maker takes a continuous action. In contrast, here the Target takes binary actions, the Spy takes continuous actions, their 'correct' actions can vary across time and they have opposing preferences. And a technical difference is that we have infinite horizon and no committed type.

[2] Strausz (2006) distinguishes *monitoring* (observing agent's actions in real time) from *auditing* (observing agent's actions/performance afterwards).

Graetz et al., 1986; Strausz, 1997). We study the latter and shed new light on how dynamic reputation considerations impacts monitors over time. Others have studied incentives to build reputations for monitoring capacity Halac and Prat (2016); Dilme and Garrett (2019) or quality (Marinovic and Szydlowski, 2022, 2023), oriented towards understanding regulatory applications.[3] In contrast, and motivated by our interest in secret intelligence and espionage, our setting features uncertainty about whether there is monitoring at all and thus emphasizes reputation for monitoring.

Model

A Target T and Spy S interact over an infinite number of periods, $t = 1, 2, ...$ S can be one of two types, $\tau_S \in \{L, N\}$, indicating whether S is listening ($\tau_S = L$) or not listening ($\tau_S = N$). Prior to the start of the game, S's type is drawn with prior probability $P[\tau_S = L] = p_0$.

At the beginning of each period t the state of the world, $\theta_t \in \{0, 1\}$ is drawn from a distribution with $P[\theta_t = 1] = 1/2$. T observes the state θ_t but S does not. States are drawn independently across time. After observing θ_t, T sends a message $m_t \in \{0, 1\}$. If $\tau_S = L$ then S observes m_t, otherwise, if $\theta_S = N$, then S does not observe m_t. Next, S chooses an action $a_t \in [0, 1]$, players observe their payoffs and the game continues to next period.

Per-period payoffs reflect two factors. First, they capture S's desire for a_t to match θ_t during each period t and T's desire for a_t to *not* match θ_t in each period. Second, they reflect T's desire, all else equal, for m_t to match the state. Formally, payoffs in period t for S and T are, respectively

$$u_S(a_t|\theta_t) = -(a_t - \theta_t)^2$$

and

$$u_T(a_t|\theta_t) = -(a_t - (1 - \theta_t))^2 - \mathbb{I}[m_t \neq \theta_t]c,$$

[3]Other work with dynamic reputation considerations for monitors: Diamond (1991) and Rajan (1992) for banking, Dye (1993) for auditors, Mathis et al. (2009) for ratings agencies, and Stolper (2009) for regulators.

where $c > 0$.

Dynamic payoffs are given by the (discounted) sum of per-period payoffs. Players discount payoffs at a rate $\delta \in [0, 1)$.

Strategies and equilibrium concept

Our analysis focuses on a selection of Perfect Bayesian Equilibrium (PBE). A PBE is an assessment (σ, μ) such that (i) the strategy profile σ is sequentially rational given beliefs μ and (ii) μ is derived from σ via Bayes's rule whenever possible. We select equilibria that are *stationary* and *symmetric*. We study equilibria that are stationary in the sense that (i) T's updated belief in period t only depends on the prior belief in period $t-1$ and the action choice of S in period $t-1$; (ii) T's choice of m_t depends only on the current state θ_t and (iii) Further, we study equilibria that are symmetric in the sense that the probability that $m_t = \theta_t$ in any period is independent of the value of θ_t. Henceforth, we refer to assessments with these features as "equilibria."

We now define stationary strategies and beliefs. Formally, a strategy for T is a mapping $\sigma_T : \{0, 1\} \to [0, 1]$ from the current period state into the set of probability distributions on messages.[4] A stationary strategy for type L of S is a mapping $\sigma_S^L : \{0, 1\} \to \Delta[0, 1]$ from the current period message m_t into the set of probability distributions on actions. As type N of S does not observe the message in each period, a strategy for this type is simply a choice of action.

We also specify beliefs for both T and S. A belief system for T is a mapping $\mu_T : [0, 1] \times [0, 1] \to [0, 1]$ from the set of actions and probability distributions on S's type in period $t-1$ into the set of probability distributions on S's type in period t.[5] A belief system for type L of S describes S's belief about the current period state and is a mapping $\mu_S : \{0, 1\} \to [0, 1]$ from the set of current period messages to the set of probability distributions on the current period state. Finally, it is unnecessary to define a system of beliefs for type N of S as they never observe T's messages.

[4]Note as the support of m_t is $\{0, 1\}$, a mixed strategy is given by a choice of Bernoulli distribution.

[5]As above, recall that S's type is distributed Bernoulli.

Analysis

Our analysis focuses on demonstrating the existence of a unique equilibrium and characterizing equilibrium behavior.

First, we characterize (i) the L type of S's beliefs about θ_t, given T's messaging strategy and (ii) each type of S's choice of a_t. With this in hand, we offer a sequence of Lemmas that describe features of equilibrium play. The first pins down continuation values for the case when T is certain that S is the L type. We then show that if T believes that S is sufficiently likely to be type S, they choose $m_t \neq \theta_t$ with positive probability, and that, conversely, if T believes that S is sufficiently likely to be type L they play $m_t = \theta_t$ with probability 1.

Then, we characterize equilibrium behavior as a function of p_t. With a series of Lemmas, we show that we can partition equilibrium behavior depending on whether $p_t > c$ or not. In particular, if $p_t > c$ then equilibrium behavior is probabilistic; T mixes between sending a "truthful" signal, $m_t = \theta_t$ and a "dishonest" signal, $m_t \neq \theta_t$, and the listening type of S mixes between acting on information and mimicking the not listening type.

We begin by characterizing the beliefs of S about the current state θ_t given the current period message m_t. To set notation, let p_t be T's belief in period t that S is type N and let $\mu(p_t)$ be the probability that T chooses $m_t = \theta_t$ given p_t.

Using Bayes's rule we have that S's beliefs after observing each message are

$$P(\theta_t = 1 | m_t = 1) = \frac{P(m_t = 1|\theta_t = 1)P(\theta_t = 1)}{P(m_t = 1)} = \frac{\mu(p_t)\frac{1}{2}}{\mu(p_t)\frac{1}{2} + (1 - \mu(p_t))\frac{1}{2}} = \mu(p_t)$$

and

$$P(\theta_t = 1 | m_t = 0) = \frac{P(m_t = 0|\theta_t = 1)P(\theta_t = 1)}{P(m_t = 0)} = \frac{(1 - \mu(p_t))\frac{1}{2}}{(1 - \mu(p_t))\frac{1}{2} + \mu(p_t)\frac{1}{2}} = 1 - \mu(p_t).$$

Given our focus on symmetric equilibria, we have that $P(\theta_t = 0|m_t = 0) = \mu(p_t)$ and $P(\theta_t = 0|m_t = 1) = 1 - \mu(p_t)$.

We now consider S's behavior. First, recall that the N type does not observe m_t. From this it follows immediately that, in equilibrium, the N type chooses $a_t = 1/2$ in every period. Because of this, throughout we assume that following any off-equilibrium-path action $a_t \neq 1/2$, T updates its belief in the following period to place probability 1 on the L type. Given that type N always chooses the same action, namely $a_t = 1/2$ we now consider the conditions under which type L chooses an action that is distinct from the N type's action. Moving forward, we say that S *acts on information* if they choose $a_t \neq 1/2$. Further, as type N always chooses the same action and is effectively a "behavioral type," we will henceforth refer to the L type simply as "S."

We now provide some characterization of S's decision to act on information. First, recall that after observing an action $a_t \neq 1/2$, T updates its belief such that $p_{t+1} = 1$. Given that all such actions lead to the same belief for T, if S chooses to act on information in period t, S will choose to maximize their static expected utility, given their belief $\mu(p_t)$. Formally, if S chooses to act on information, their choices of a_t after observing $m_t = 1$ $m_t = 0$ are

$$\max_{a_t} \quad -\mu(p_t)(a_t - 1)^2 - (1 - \mu(p_t))(a_t - 0)^2 = \mu(p_t),$$

and

$$\max_{a_t} \quad -\mu(p_t)(a_t - 0)^2 - (1 - \mu(p_t))(a_t - 1)^2 = 1 - \mu(p_t),$$

respectively.

Let $\alpha(p_t)$ be the probability that the informed S acts on information, and suppose that with probability $1-\alpha(p_t)$ S chooses $x_t = 1/2$ and mimics the uninformed type.

First, we characterize equilibrium behavior when T is certain that S is the listening type. Lemma 1 shows that, in this case, equilibrium continuation values are unique. This observation facilitates the rest of the analysis because it also pins down payoffs in all periods after L type of S chooses to act. To set notation, let $V_i(p_t)$ be player i's continuation payoff at the start of a period t in which T's belief is p_t.

Lemma 1. *If $p_t = 1$ then continuation values are given by*

$$V_S(1) = -\frac{1}{1-\delta}\frac{1-c^2}{4}$$
$$V_T(1) = -\frac{1}{1-\delta}\frac{(1+c)^2}{4}$$

Proof. First, note that if $p_t = 1$, then $p'_t = 1$ for all $t' > t$. Therefore, S will always choose x to maximize its static payoff, as it cannot influence T's beliefs. From above, we know that ,

$$V_S(1) = -\mu(1-\mu) + \delta V_S(1).$$

T must be indifferent between $m = \theta$ and $m \neq \theta$, which holds iff $\mu = \frac{1+c}{2}$ (see next proof and plug in $\alpha = 1$ and $p = 1$). Thus,

$$V_S(1) = \frac{1}{1-\delta}\Big(- \frac{1+c}{2}(1 - \frac{1+c}{2})\Big) = -\frac{1}{1-\delta}\Big(\frac{1-c^2}{4}\Big)$$

Additionally, since in each state T must be indifferent we have

$$\begin{aligned} V_T(1) &= -\mu^2 + \delta V_T(1) \\ \Leftrightarrow V_T(1) &= -\frac{\mu^2}{1-\delta} \\ \Rightarrow V_T(1) &= -\frac{1}{1-\delta}\frac{(1+c)^2}{4} \end{aligned}$$

□

As proof of Lemma 1 reveals, continuation values are uniquely pinned down by the fact that, once T becomes convinced that S is listening, behavior becomes quite simple. Knowing that T is aware it is listening, S maximizes its static utility. As a consequence, T uses a mixed strategy to obscure the true state. Solving the resulting indifference conditions yields the continuation values.

The next step in our analysis focuses on low values of p_t. In particular, we show that if T is sufficiently sure that S is not listening, there is a unique equilibrium.

Lemma 2. *If $p_t \leq c$ then, there is a unique equilibrium in which $\mu(p_t) = 1$ and $\alpha(p_t) = 1$.*

Proof. Without loss of generality given our focus on symmetric equilibria, suppose

$\theta_t = 1$. For T we have

$$U_T(m=1|\theta=1) = p_t\alpha_t\Big(- \mu_t^2 + \delta V_T(1)\Big) + \Big(p_t(1-\alpha_t) + 1 - p_t\Big)\Big(- \frac{1}{4} + \delta V_T(p_{t+1})\Big)$$
$$U_T(m=0|\theta=1) = p_t\alpha_t\Big(- (1-\mu_t)^2 + \delta V_T(1)\Big) + \Big(p_t(1-\alpha_t) + (1 - p_t)\Big)\Big(- \frac{1}{4} + \delta V_T(p_{t+1})\Big) - c.$$

Thus, $U_T(m=1|\theta=1) \geq U_T(m=0|\theta=1)$ iff $c \geq \alpha_t(2\mu_t - 1)p_t$, which always holds if $c \geq p_t$.

Note that if T chooses $a_t = \theta_t$ with probability 1, then S must act. This completes the proof. □

This result provides important information about qualitative features of equilibrium play. Specifically, if $p_t \leq c$ then (i) T becomes completely "honest," choosing $m_t = \theta_t$ and (ii) the listening type of S acts on their information with probability 1. Given this, the path of play in future periods is predictable and falls into one of two categories depending on S's type. If S does not act in a period in which $p_t \leq c$, then T updates its belief to $p_{t+1} = 0$. Consequently, in all future periods, as T knows S is not listening it sets $m_t = \theta_t$. If S does act on information then T knows for sure that it is the listening type, and in all future periods behavior is as described in the proof of Lemma 1.

Our first two results showed that S acts on information with probability 1 if $p_t \leq c$ or if $p_t = 1$. Our next result demonstrates that for intermediate values of p_t S always acts on information with positive probability.

Lemma 3. *In equilibrium, (i) $\alpha(p_t) > 0$, (ii) $\mu(p_t) > 1/2$, and (iii) $p_t > c$ implies $\mu(p_t) < 1$.*

Proof. First, we prove (i). To derive a contradiction, suppose $\alpha(p_t) = 0$. Note that T's best response to $\alpha(p_t)$ is $\mu(p_t) = 1$. However, given that $\mu(p_t) = 1$, S may profitably deviate to $\alpha'(p_t) = 1$, a contradiction.

Next, we prove (ii). To start, note that $\mu < 1/2$ cannot be optimal since it induces the same variance in beliefs for S as $1/2 + |1/2 - \mu|$ but incurs greater lying costs for T. Next, suppose $\mu(p_t) = 1/2$. In this case, S's best response is to choose $x = 1/2$ with probability 1. However, then T could profitably deviate $\mu = 1$.

Finally, we prove (iii). To derive a contradiction, suppose $p_t > c$ and $\mu(p_t) = 1$. By earlier lemma, for T to be honest requires $c \geq \alpha_t(p_t)$. If $\mu_t = 1$ then S's best response is to always act, $\alpha_t = 1$. Thus, the condition for T to be honest becomes $c \geq p_t$, a contradiction. □

The proof of Lemma 3 illustrates an important strategic tension that arises in equilibrium. If there is a reputation at which S will not ever act on information, then T wants to use its information for sure. But then S will learn perfectly for T's action and will want to always act on that information. Due to this tension, there is always a chance that S acts on its information.

Additionally, T is always strictly more likely to choose the action that matches the state. It has to favor one direction, because S would otherwise always want to choose $x = \frac{1}{2}$ and, in turn, T would want to just match the state. Furthermore, it has no incentive to favor the action that does not match the state, since it randomizes in order to create risk for S.

Next, we establish a lower bound on how much T's belief changes between periods after seeing $x = \frac{1}{2}$. It is a useful property for the equilibrium existence argument

Lemma 4. *In equilibrium, $p_t - p_{t+1} > p_t - \frac{p_t - c}{1-c} > 0$ for any belief p_t.*

Proof. Fix p_t. By previous lemmas, at any p_t we must have either $\alpha_t = 1$ or $\alpha_t \in (0,1)$. In the first case, Bayes rule immediately yields $p_{t+1} = 0$, thus, $p_t - p_{t+1} = p_t > p_t - \frac{p_t - c}{1-c} > 0$.

Next, assume $\alpha_t \in (0,1)$. This also implies that $\mu_t < 1$, since $\mu_t = 1$ iff $p_t < c$ and if $p_t < c$ then $\alpha_t = 1$. Thus, T must be indifferent in equilibrium, which implies the following must hold

$$
\begin{aligned}
c &= \alpha_t(2\mu_t - 1)p_t \\
\Leftrightarrow \alpha_t &= \frac{c}{p_t(2\mu_t - 1)}.
\end{aligned}
$$

In any equilibrium, by Bayes rule we have

$$
p_t - p_{t+1} = p_t - \frac{(1-\alpha_t)p_t}{(1-\alpha_t)p_t + 1 - p_t} = p_t - \frac{p_t - \frac{c}{2\mu_t - 1}}{1 - \frac{c}{2\mu_t - 1}}
$$

Note, $p_t - p_{t+1}$ is decreasing in μ_t. Thus, $p_t - p_{t+1}$ is minimized at $\mu_t = 1$, which yields the result. □

Using Lemma 4, we now characterize equilibrium behavior when T is fairly certain S is not listening, but still suspicious enough to not always follow her information. Specifically, we show that there is an interval of beliefs just above c for which T mixes but S always acts on its information. It does so because T is very likely to act according to its information. Therefore, even though observing T's action is not fully informative, S have enough information to act boldly in the information it acquires from its secret intel. Additionally, we show that in this region of beliefs T's action is more informative as p_t decreases towards c and converges to being fully informative.

Lemma 5. *If $p_t \in (c, c/\sqrt{\delta(1-c^2)}]$, then $\mu(p_t) = \frac{1}{2} + \frac{c}{2p_t}$ and $\alpha(p_t) = 1$ in equilibrium.*

Proof. We start by showing that such an equilibrium exists. We show that no player has a profitable deviation from the proposed strategies given their beliefs, and that beliefs are derived via Bayes rule when possible. Fix $p_t \in (c, c/\sqrt{\delta(1-c^2)}]$, $\mu(p_t) = 1/2 + c/2p_t$, and $\alpha(p_t) = 1$.

First, in order for T to mix, they must be indifferent. Using the indifference condition from the proof of Lemma 2 and substituting for $\alpha(p_t) = 1$, we find that T is indifferent if and only if

$$\mu(p_t) = \frac{1}{2} + \frac{c}{2p_t},$$

which we have assumed.

Next, note that in period $t+1$, if player S does not act in period t, Bayes rule yields $p_{t+1} = 0$. Further, if player S does act in period t, it follows from Bayes rule that $p_{t+1} = 1$. Therefore, by Lemmas 1 and 2, the continuation values for each player in period $t+1$ are pinned down.

Using these continuation values, we see that S cannot profitably deviate from acting in period t if and only if

$$-\mu_t(1-\mu_t) + \delta V_S(1) \geq -\frac{1}{4} + \delta^2 V_S(1).$$

Substituting for μ_t, this condition holds if and only if

$$p_t \leq \frac{c}{\sqrt{\delta(1-c^2)}}.$$

We now prove uniqueness. The above proof implies there cannot exist another equilibrium in which $p_{t+1} < c$. Thus, there cannot be any other equilibrium in which $\alpha_t \geq \frac{p-c}{p(1-c)}$, since by Bayes' rule $p_{t+1} = \frac{(1-\alpha_t)p_t}{(1-\alpha)p_t+1-p_t} \leq c$ for all $\alpha_t \geq \frac{p-c}{p(1-c)}$.

To obtain uniqueness, suppose there exists an equilibrium in which $\alpha_t < \frac{p-c}{p(1-c)}$, which implies $p_{t+1} > c$. Note, we must have $p_t > c + (1-c)c$, otherwise by Lemma 4 $p_{t+1} < c$.

By Lemma 3, $\alpha_t > 0$, so S must be mixing and we must also have T mixing. This implies $\alpha_t = \frac{c}{p_t(2\mu_t-1)}$. Thus, in such an equilibrium we have

$$\frac{c}{p_t(2\mu_t-1)} \leq \frac{p-c}{p_t(1-c)}$$
$$\mu_t \geq \frac{1}{2} + \frac{c(1-c)}{2(p_t-c)}.$$

Since both players are mixing, we have that $p_{t+1} = \frac{p_t - \frac{c}{2\mu_t-1}}{1-\frac{c}{2\mu_t-1}}$. Since this term is increasing in μ and $\mu \geq \frac{1}{2} + \frac{c(1-c)}{2(p_t-c)}$, substituting in we have that

$$p_{t+1} \geq \frac{p_t - \frac{c}{(\frac{c(1-c)}{p-c})}}{1 - \frac{c}{(\frac{c(1-c)}{p-c})}}.$$

Additionally, Lemma 4 implies

$$p_{t+1} \leq \frac{p_t - c}{1-c}.$$

Thus, a necessary condition for such an equilibrium to exist is

$$\frac{p_t - \frac{c}{(\frac{c(1-c)}{p-c})}}{1 - \frac{c}{(\frac{c(1-c)}{p-c})}} \leq \frac{p_t - c}{1-c}$$
$$\Leftrightarrow p_t \leq c + (1-c)c,$$

a contradiction. $\square$

In light of Lemma 4, we have pinned down equilibrium behavior for $p_t \in [0, c/\sqrt{\delta(1-c^2)}] \cup \{1\}$. By using an inductive argument and leveraging Lemma 4, we can complete the characterization for all reputations. Essentially, we can partition the space of reputations into intervals that are distinguished primarily by the probability that S acts on its information. Moreover, we show that the partition is unique and, in turn, equilibrium behavior is too.

The equilibrium has the property that S will be "biding time". In each period, either (i) S acts and is "found out" or (ii) S does not act and play proceeds with T slightly less suspicious. Each additional period in which $a_t = 1/2$ makes T more "honest." In turn, this makes acting more attractive for the L type of S. Thus, equilibrium behavior is probabilistic and repeated rounds of no action from S do not imply that they are N type.

We provide the details below in Lemma 6 and Proposition 1.

Lemma 6. *There exists an equilibrium.*

Proof. We construct an equilibrium by using an induction argument. First, we prove the base case. Take any $p \leq \frac{c}{\sqrt{\delta(1-c^2)}}$. Define $\overline{\mu}(p)$ as the solution to

$$\mu(1-\mu) = \frac{1}{4} + \delta\Big[V(1) - V(p)\Big].$$

Note from previous lemmas we have $V(p)$ is weakly decreasing in p for $p \leq \frac{c}{\sqrt{\delta(1-c^2)}}$, which implies $\overline{\mu}(p)$ is decreasing in p. Next, define the function $\overline{p}(p)$ as

$$\overline{p}(p) = p(1 - \frac{c}{2\overline{\mu}(p)-1}) + \frac{c}{2\overline{\mu}(p)-1}.$$

By construction, if $p_t = \overline{p}(p)$ there exists an equilibrium in which $\mu_t = \overline{\mu}(p)$, $\alpha_t = \frac{c}{p(2\overline{\mu}(p)-1)}$, and $p_{t+1} = p$.

Finally, we show that $\overline{p}(p)$ is strictly increasing in p. Let $p' > p''$ then

$$\begin{aligned}\overline{p}(p') - \overline{p}(p'') &= p'(1 - \frac{c}{2\overline{\mu}(p')-1}) + \frac{c}{2\overline{\mu}(p')-1} - \Big[p''(1 - \frac{c}{2\overline{\mu}(p'')-1}) + \frac{c}{2\overline{\mu}(p'')-1}\Big] \\ &= p' - p'' + \frac{c}{2\overline{\mu}(p')-1}(1-p') - \frac{c}{2\overline{\mu}(p'')-1}(1-p'').\end{aligned}$$

Since $\overline{\mu}(p)$ is decreasing in p, $\frac{c}{2\overline{\mu}(p)-1}$ is increasing in p. Thus,

$$\overline{p}(p') - \overline{p}(p'') \geq p' - p'' + \frac{c}{2\overline{\mu}(p')-1}(1-p') - \frac{c}{2\overline{\mu}(p')-1}(1-p'')$$
$$\Rightarrow \overline{p}(p') - \overline{p}(p'') \geq (p'-p'')(1 - \frac{c}{2\overline{\mu}(p')-1}) > 0$$

Since $\overline{p}(p)$ is strictly increasing and continuous in p, this implies an equilibrium exists for all $p \in \left(\frac{c}{\sqrt{\delta(1-c^2)}}, \overline{p}\left(\frac{c}{\sqrt{\delta(1-c^2)}}\right)\right]$.

Finally, note that $V(p) = -1/4 + \delta V(p_{t+1})$. Recall that p_{t+1} is increasing in p, which implies that $V(p_{t+1})$ is decreasing in p_{t+1}. Therefore, $V(p)$ is decreasing in p.

This completes the proof of the base case.

For the induction step, suppose that there exists an equilibrium for $p \in (\overline{p}_{k-1}, \overline{p}_k]$ and that $V(p)$ is decreasing in p for $p \in (\overline{p}_{k-1}, \overline{p}_k]$. We will show that there is an equilibrium for $p \in (\overline{p}_k, \overline{p}_{k+1}]$ and that $V(p)$ is decreasing in p for $p \in (\overline{p}_k, \overline{p}_{k+1}]$.

Fix $p \in (\overline{p}_{k-1}, \overline{p}_k]$. The fact that $V(p)$ is decreasing for $p \in (\overline{p}_{k-1}, \overline{p}_k]$ implies that $\overline{\mu}(p)$ is decreasing for $p \in (\overline{p}_{k-1}, \overline{p}_k]$. As above, by construction of $\overline{p}(p)$, if $p_t = \overline{p}(p)$ then there exists an equilibrium in which $\mu_t = \overline{\mu}(p)$, $\alpha_t = \frac{c}{p(2\overline{\mu}(p)-1}$, and $p_{t+1} = p$. Finally, note that by the arguments above, $\overline{p}(p)$ is strictly increasing in p. This implies that an equilibrium exists for all $p \in (\overline{p}_k, \overline{p}_{k+1}]$. As above, note that $V(p) = -1/4 + \delta V(p_{t+1})$. Recall that p_{t+1} is increasing in p, which implies that $V(p_{t+1})$ is decreasing in p_{t+1}. Therefore, $V(p)$ is decreasing in p.

Finally, we show that as $\lim_{k\to\infty} p_k = 1$. First, note that $p_k \leq 1$ for all k by construction. To complete the proof, for a contradiction suppose that $\lim_{k\to\infty} p_k = p < 1$. This implies that for all $\epsilon > 0$, $\overline{p}(p-\epsilon) \leq p$. However, for sufficiently small $\epsilon > 0$, this cannot hold, as for all $p < 1$, $p_t - p_{t+1} > 0$ by Lemma 6.

□

Proposition 1. *There is a unique SSPBE.*

Step 1: Lemma [unique below $\overline{p}_0$. implies uniqueness below for $p \leq \overline{p}_0$.]

Step 2: Suppose our equilibrium is the unique SSPBE for $\overline{p}_{k-1}$. We show that if $p \in [\overline{p}_{k-1}, \overline{p}_k]$, then the equilibrium characterized in Proposition [existence] is unique. This step has two parts.

Part 1: In our equilibrium, must get below $\overline{p}_0$ in k steps. To show a contradiction, suppose there exists an equilibrium σ' that takes $k+1$ steps.

Claim: In σ', we must have $p'_{t+1} > \overline{p}_{k-1}$. If not, then $p'_{t+1} < \overline{p}_{k-1}$ and by the induction assumption there is a unique equilibrium that gets below $\overline{p}_0$ in weakly less than $k-1$ steps, but that would contradict σ' taking $k+1$ steps.

So, suppose we have σ' such that $p'_{t+1} > \overline{p}_{k-1}$. The continuation value in σ satisfies

$$V_S(p;\sigma) \geq \delta^k[\delta V_S(1) - \mu(\overline{p}_0)(1-\mu(\overline{p}_0)] + \sum_{t=1}^{k}(-\frac{1}{4})\,\delta^{t-1} \equiv \underline{V}_S(p;\sigma).$$

The continuation value in σ' satisfies

$$V_S(p;\sigma') \leq \delta^{k+1}[\delta V_S(1)] + \sum_{t=1}^{k+1}(-\frac{1}{4})\,\delta^{t-1} \equiv \overline{V}_S(p;\sigma').$$

We have $\underline{V}_S(p;\sigma) \geq \overline{V}_S(p;\sigma')$ if and only if

$$\delta^k[\delta\, V_S(1) - \mu(\overline{p}_0)(1-\mu(\overline{p}_0)] + \frac{\delta^k}{4} - \delta^{k+2}\, V_S(1) \geq 0 \tag{1}$$

$$\frac{\delta}{4}(1-c^2) - \frac{\delta}{4}(1-c^2) \geq 0, \tag{2}$$

where (2) follows from the definitions of $\mu(\overline{p}_0)$ and $V_S(1)$. Thus, we have shown that $V_S(p;\sigma) \geq V_S(p;\sigma')$.

Furthermore, $\overline{V}_S(p;\sigma') \geq V_S(p;\sigma'')$ for any equilibrium σ'' that takes $k+\ell$ steps to get below $\overline{p}_0$.

Part 2: A similar argument shows that we cannot have an equilibrium σ' that takes fewer than k steps to get below $\overline{p}_0$. That would require that $p'_{t+1} < \overline{p}_{k-2}$, which in turn implies that $p_{t+1} < \overline{p}_{k-2}$, a contradiction.

Step 3: Assume $p \in [p_{k-1}, p_k]$. Let σ_n denote the equilibrium that takes n steps and σ_k the k step equilibrium, with $n > k$. T's equilibrium mixing probability

$\mu(\sigma_n)$ in σ_n solves

$$\mu(1-\mu) = \frac{1}{4} + \delta\Big[V(1) - V(p_{t+1};\sigma_n)\Big], \tag{3}$$

while T's mixing probability $\mu(\sigma_k)$ in σ_k solves

$$\mu(1-\mu) = \frac{1}{4} + \delta\Big[V(1) - V(p_{t+1};\sigma_n)\Big], \tag{4}$$

Note that by step 2 we have $V_S(p;\sigma_n) \leq V_S(p;\sigma_k)$. Thus, RHS(3) $\geq$ RHS(4), which implies that $\mu(\sigma_k) \geq \mu(\sigma_n)$. In either equilibrium the belief in the next period is given by

$$p_{t+1} = \frac{p_t - \frac{c}{2\mu-1}}{1 - \frac{c}{2\mu-1}}.$$

Since p_{t+1} is increasing μ and $\mu(\sigma_k) \geq \mu(\sigma_n)$ it must be that $p_{t+1}(\sigma_k) > p_{t+1}(\sigma_n)$, but this contradicts that $p_{t+1}(\sigma_n) > \overline{p}_{k-1} > p_{t+1}(\sigma_k)$. □

Discussion of Next Steps

We see two clear next steps. First, relax assumption that θ_t is redrawn in each period. Second, relax assumption that "acting" brings belief to 1. Each allow us to address a wider range of applications. Additionally, we will characterize benchmarks with commitment.

Extension idea 1: In baseline model, θ_t is redrawn i.i.d. in each period. Alternative: let θ_t be "sticky." Baseline model applies better to situations where T's desired outcome changes day to day, e.g. battlefield plans. Sticky θ_t applies to settings where T's desired outcome day to day has some persistence. Criminal investigation is one possibility, seems natural to have $\theta_t = \theta_{t+1}$ in this setting.

Extension idea 2: In baseline, $a_t \neq 1/2$ *always* comes from L type. In many applications, this is not true (e.g. cracking of Enigma code). Several ways to model this. Equilibrium path should be more nuanced, involve fewer "absorbing" states.

References

Bar-Isaac, Heski, "Reputation and Survival: Learning in a Dynamic Signalling Model," *Review of Economic Studies*, 2003, *70* (2), 231–251.

_ **and Steven Tadelis**, "Seller Reputation," *Foundations and Trends® in Microeconomics*, 2008, *4* (4), 273–351.

Benabou, Roland and Guy Laroque, "Using Privileged Information to Manipulate Markets: Insiders, Gurus, and Credibility," *Quarterly Journal of Economics*, 1992, pp. 921–958.

Diamond, Douglas, "Monitoring and Reputation: The Choice between Bank Loans and Directly Placed Debt," *Journal of Political Economy*, 1991, *99* (4), 689–721.

Dilme, Francesc and Daniel Garrett, "Residual Deterrence," *Journal of the European Economic Association*, 2019, *17* (5), 1654–1686.

Dye, Ronald A., "Auditing Standards, Legal Liability, and Auditor Wealth," *Journal of Political Economy*, 1993, *101* (5), 887–914.

Ely, Jeffrey C. and Juuso Välimäki, "Bad Reputation," *The Quarterly Journal of Economics*, 2003, pp. 785–814.

Graetz, Michael J., Jennifer F. Reinganum, and Louis L. Wilde, "The Tax Compliance game: Toward an Interactive Theory of Law Enforcement," *Journal of Law, Economics, and Organization*, 1986, *2* (1), 1–32.

Halac, Marina and Andrea Prat, "Managerial Attention and Worker Performance," *American Economic Review*, 2016, *106* (10), 3104–32.

Kreps, David M. and Robert Wilson, "Reputation and Imperfect Information," *Journal of Economic Theory*, 1982, *27* (2), 253–279.

Marinovic, Iván and Martin Szydlowski, "Monitoring with Career Concerns," *RAND Journal of Economics*, 2022, *53* (2), 404–428.

— **and** — , "Monitor Reputation and Transparency," *American Economic Journal: Microeconomics*, 2023, *15* (4), 1–67.

Mathis, Jerome, James McAndrews, and Jean-Charles Rochet, "Rating the Raters: Are Reputation Concerns Powerful Enough to Discipline Rating Agencies?," *Journal of Monetary Economics*, 2009, *56* (5), 657–674.

Milgrom, Paul and John Roberts, "Predation, Reputation, and Entry Deterrence," *Journal of Economic Theory*, 1982, *27* (2), 280–312.

Morris, Stephen, "Political Correctness," *Journal of Political Economy*, 2001, *109* (2), 231–265.

Rajan, Raghuram G., "Insiders and Outsiders: The Choice between Informed and Arm's-length Debt," *Journal of Finance*, 1992, *47* (4), 1367–1400.

Sobel, Joel, "A Theory of Credibility," *The Review of Economic Studies*, 1985, *52* (4), 557–573.

Solan, Eilon and Leeat Yariv, "Games with Espionage," *Games and Economic Behavior*, 2004, *47* (1), 172–199.

Stolper, Anno, "Regulation of Credit Rating Agencies," *Journal of Banking & Finance*, 2009, *33* (7), 1266–1273.

Strausz, Roland, "Delegation of Monitoring in a Principal-Agent Relationship," *Review of Economic Studies*, 1997, pp. 337–357.

— , "Timing of Verification Procedures: Monitoring versus Auditing," *Journal of Economic Behavior & Organization*, 2006, *59* (1), 89–107.

Tan, Teck Yong, "Optimal Transparency of Monitoring Capability," *Journal of Economic Theory*, 2023, *209*, 105620.