When to Drop a Bombshell

Gabriele Gratton, Richard Holden, and Anton Kolotilin*

18th May 2017

Abstract

Sender, who is either good or bad, wishes to look good at an exogenous deadline. Sender privately observes if and when she can release a public flow of information about her private type. Releasing information earlier exposes it to greater scrutiny, but signals credibility. In equilibrium bad Sender releases information later than good Sender. We find empirical support for the dynamic predictions of our model using data on the timing of US presidential scandals and US initial public offerings. In the context of elections, our results suggest that "October Surprises" are driven by the strategic behavior of bad Sender.

Keywords: information disclosure, strategic timing, Bayesian learning, credibility vs. scrutiny.

JEL Classification Numbers: D72, D82, D83, L82.

*School of Economics, UNSW Business School, Sydney, NSW 2052, Australia. Email: g.gratton@unsw.edu.au, richard.holden@unsw.edu.au, a.kolotilin@unsw.edu.au. We are grateful to Alessandro Bonatti, Steven Callander, Yeon-Koo Che, Wouter Dessein, William Fuchs, Drew Fudenberg, Robert Gibbons, Navin Kartik, Keiichi Kawai, Hongyi Li, Jin Li, Brendan Nyhan, Carlos Pimienta, Andrea Prat, and participants at various seminars and conferences for helpful comments. Valentino Dardanoni, Antonio Forcina, and Jutta Roosen kindly shared their codes for the likelihood ratio order test. We especially thank Aleksandra Balyanova and Barton Lee for excellent research assistance. The authors acknowledge financial support from the Australian Research Council.

1 Introduction

Election campaigns consist of promises, allegations, and scandals. While most of them are inconsequential, some are pivotal events that can sway elections. Rather than settling existing issues, these *bombshells* typically start new debates that, in time, provide voters with new information. When bombshells are dropped, their timing is hotly debated. Was the bombshell intentionally timed to sway the election? What else did media and politicians know, when they dropped the bombshell, that voters might only discover after the election?

The 2016 US presidential campaign between Democrat Hillary Clinton and Republican Donald Trump provides several examples. Just eleven days before the election, FBI director James Comey announced that his agency was reopening its investigation into Secretary Clinton's emails. The announcement reignited claims that Clinton was not fit to be commander in chief because of her mishandling of classified information. Paul Ryan, the Republican Speaker of the House, went as far as to demand an end to classified intelligence briefings to Clinton. Some commentators maintain that Comey's announcement cost Clinton the election.[1]

While the announcement conveyed the impression of an emerging scandal, Clinton was confident that no actual wrongdoing would be revealed by the new investigation —there would be no *real* scandal. Comey's letter to Congress stated that "the FBI cannot yet assess whether or not this material may be significant, and I cannot predict how long it will take us to complete this additional work." The Clinton campaign—and Democrats generally—were furious, accusing Comey of interfering with the election. Comey wrote that he was briefed on the new material only the day before the announcement. But his critics maintained that the FBI had accessed the new emails weeks before the announcement and speculated about how long Comey sat on the new material and what he knew about it.[2]

Similarly, one month before the election, the Washington Post released a video featuring Donald Trump talking about women.[3] The video triggered a heated public debate about whether Trump was fit to be president. It revived allegations that he had assaulted women and even prominent Republicans called for Trump to end his campaign.[4] Within

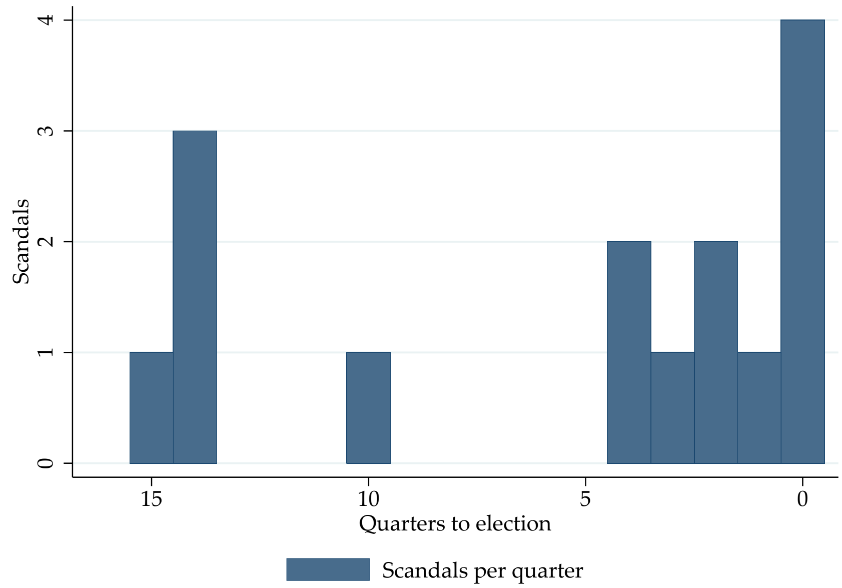
[1]For example, Paul Krugman wrote that the announcement "very probably installed Donald Trump in the White House" (*New York Times*, Jan. 13, 2017).

[2]Matt Zapotosky, Ellen Nakashima, and Rosalind S. Helderman, *Washington Post*, Oct. 30, 2016.

[3]Although the video was filmed 11 years prior to the release, raising the question of whether it was strategically timed, the Washington Post maintains it obtained the unedited video only a few hours before its online release (Farhi, Paul, *Washington Post*, Oct. 7, 2016).

[4]Aaron Blake, *Washington Post*, Oct. 9, 2016.

Figure 1: Distribution of scandals implicating US presidents running up for reelection, from 1977 to 2008. Data from Nyhan (2015).



a week of the video's release, five women came forward accusing Trump of sexual assault. Trump himself denied all accusations and dismissed the video as "locker-room banter," and "nothing more than a distraction from the important issues we are facing today."[5] Others echoed his statement that real scandals are about "actions" and not "words," and took the media coverage of the video as proof of a conspiracy against Trump.[6]

The concentration of scandals in the last months of the 2016 campaign is far from an exception. Such *October surprises* are commonplace in US presidential elections, as shown in Figure 1. Political commentators argue that scandals may be strategically released close to elections so that voters have not enough time to tell real from fake scandals. Yet, if all ill-founded scandals were released just before an election, then voters may rationally discount October surprises as fake. Voters may not do so fully, however, since while some scandals may be strategically timed, others are simply discovered close to the election.

In this paper we analyze a Sender-Receiver game which connects the timing of bombshells with voters' beliefs on the day of the election. On the one hand, dropping the bombshell earlier is more credible, in that it signals that Sender has nothing to hide. On the other hand, it exposes the bombshell to scrutiny for a longer period of time—possibly revealing that the bombshell is a fake. In equilibrium, while real scandals are released as they are discovered, fake scandals are strategically delayed and concentrated towards the

[5]*Los Angeles Times*, transcript of Trump's video statement, Oct. 7, 2016.

[6]Susan Page and Karina Shedrofsky, *USA TODAY*, Oct. 26, 2016

end of the campaign. In other words, our credibility-scrutiny tradeoff predicts that the October surprise phenomenon is driven by fake scandals.

The same tradeoff drives the timing of announcements about candidacy, running mates, cabinet members, and details of policy platforms. An early announcement exposes the background of the candidate or her team to more scrutiny, but boosts credibility. Beyond announcements, the tradeoff can also determine the timing of policy implementation. For example, presidents without a good long term plan for the economy would implement popular policies, such as pork-barrel spending and tax cuts, towards the end of their term and just before their bid for reelection, because the possible adverse effects of such policies on inflation and public debt are likely to realize after the election. This provides a potential rational-agent explanation for the so-called "political business cycle."[7] Finally, the same tradeoff between credibility and scrutiny is likely to drive the timing of information release in other contexts outside the political sphere. For instance, a company going public can provide more or less time for the market to evaluate its prospectus before the company's shares are traded.

In all these situations, (i) biased Sender has private information and (ii) she cares about Receiver's opinion at a given date. Crucially, (iii) Sender privately knows the earliest date at which she can drop a bombshell, but she can choose to do so later. In this paper we introduce and analyze a formal model of precisely these types of dynamic information release problems. In our model of Section 2, (i) Sender privately knows her binary type, *good* or *bad*, and (ii) wants Receiver to believe that she is good at an exogenous *deadline*; (iii) Sender privately observes whether and when an opportunity to start a public flow of information about her type arrives and chooses when to seize this opportunity. We call this opportunity an *arm* and say that Sender chooses when to *pull* the arm.[8]

In Section 3.1, we characterize the set of perfect Bayesian equilibria. In all equilibria, bad Sender delays pulling the arm relative to good Sender, despite the fact that pulling the arm has a positive instantaneous effect on Receiver's belief. An immediate implication is that, pulling the arm earlier is more credible in that it instantaneously induces a higher Receiver's belief that Sender is good. Moreover, bad Sender chooses not to pull the arm with strictly positive probability.

We prove that there exists an essentially unique divine equilibrium (Cho and Kreps,

[7]See Nordhaus (1975) for an early contribution, and Alesina (1987) for the first formal analysis of the phenomenon.

[8]In Section 3.2, we generalize the model in several directions allowing for more general utility functions, for Sender to be imperfectly informed, for Sender's type to affect when the arm arrives, and for the deadline to be stochastic.

1987).[9] In this equilibrium, good Sender *immediately* pulls the arm when it arrives and bad Sender is indifferent between pulling the arm at any time and not pulling it at all. Uniqueness allows us to analyze comparative statics in a tractable way in a special case of our model where the arm arrives according to a Poisson process and pulling the arm starts an exponential learning process in the sense of Keller, Rady and Cripps (2005).

We do this in Section 4 and show that the comparative static properties of this equilibrium are very intuitive. Both good and bad Sender gain from a higher Receiver's prior belief that Sender is good. Instead, while good Sender gains from a faster learning process and a faster arrival of the arm, bad Sender loses from these.

When learning is faster and when the arm arrives more slowly, bad Sender delays pulling the arm for longer and pulls it with lower probability. In this case, the total probability that (good and bad) Sender pulls the arm is also lower. When Receiver's prior belief is higher, withholding information is less damning, so bad Sender strategically pulls the arm with lower probability, but the probability that good Sender pulls the arm is mechanically higher. We show that the strategic effect dominates the mechanical effect if and only if Receiver's prior belief is sufficiently low.

We show that the probability density with which bad Sender pulls the arm is single-peaked in time, and derive the conditions under which it monotonically increases with time. We also characterize the shape of the probability density with which (good and bad) Sender pulls the arm, and show it has at most two peaks—an earlier peak driven by good Sender and a later peak driven by bad Sender.

In Section 5, we apply our model to the strategic release of political scandals in US presidential campaigns. Using data from Nyhan (2015), we show that fake scandals are released closer to the election and that the October surprise effect is mostly driven by fake scandals. To the best of our knowledge, this is the first empirical evidence about the strategic timing of political scandals relative to the date of elections and the first direct evidence of an October Surprise effect.

Finally, we apply our model to financial data on the timing of initial public offerings (IPOs). We link a stock's long-run performance to the time gap between the announcement of an IPO and the initial trading date. Our model predicts that firms with higher long-run returns should choose longer time gaps in the likelihood ratio order. We verify this prediction using an approach developed by Dardanoni and Forcina (1998).

[9] The equilibrium is essentially unique in the sense that the probability with which each type of Sender pulls the arm at any time is uniquely determined.

1.1 Related Literature

Grossman and Hart (1980), Grossman (1981), and Milgrom (1981) pioneered the study of verifiable information disclosure and established the *unraveling result*: if Sender's preferences are common knowledge and monotonic in Receiver's action (for all types of Sender) then Receiver learns Sender's type in any sequential equilibrium. Dye (1985) first pointed out that the unraveling result fails if Receiver is uncertain about Sender's *information endowment*.[10] When Sender does not disclose information, Receiver is unsure as to why, and thus cannot conclude that the non-disclosure was strategic, and hence does not "assume the worst" about Sender's type.

Acharya, DeMarzo and Kremer (2011) and Guttman, Kremer and Skrzypacz (2013) explore the strategic timing of information disclosure in a dynamic version of Dye (1985).[11] Acharya et al. (2011) focus on the interaction between the timing of disclosure of private information relative to the arrival of external news, and clustering of the timing of announcements across firms. Guttman et al. (2013) analyze a setting with two periods and two signals and show that, in equilibrium, both *what* is disclosed and *when* it is disclosed matters. Strikingly, the authors show that later disclosures are received more positively.

All these models are unsuited to study either the credibility or the scrutiny sides of our tradeoff, because information in these models is verified instantly and with certainty once disclosed. In our motivating examples, information is not immediately verifiable: when Sender releases the information, Receiver only knows that "time will tell" whether the information released is reliable. To capture this notion of partial verifiability, we model information as being verified stochastically over time in the sense that releasing information starts a learning process for Receiver akin to processes in Bolton and Harris (1999), Keller, Rady and Cripps (2005), and Brocas and Carrillo (2007). In contrast to these papers, in our model Sender is privately informed and she chooses when to start rather than stop the process.[12]

Our application to US presidential scandals also contributes to the literature on the effect of biased media on voters' behavior (e.g., Mullainathan and Shleifer, 2005; Gentzkow

[10] See also Jung and Kwon (1988), Shin (1994), and Dziuda (2011). The unraveling result migh also fail if disclosure is costly (Jovanovic, 1982) or information acquisition is costly (Shavell, 1994).

[11] Shin (2003, 2006) also study dynamic verifiable information disclosure, but Sender there does not strategically time disclosure. A series of recent papers consider dynamic information disclosure with different focuses to us, including: Che and Hörner (2015); Ely, Frankel and Kamenica (2015); Bizzotto, Rüdiger and Vigier (2016); Ely (2016); Grenadier, Malenko and Malenko (2016); Horner and Skrzypacz (2016); Orlov, Skrzypacz and Zryumov (2016).

[12] In our model Sender can influence only the starting time of the experimentation process, but not the design of the process itself. Instead, in the Bayesian persuasion literature (e.g., Rayo and Segal, 2010; Kamenica and Gentzkow, 2011), Sender fully controls the design of the experimentation process.

and Shapiro, 2006; Duggan and Martinelli, 2011).[13] DellaVigna and Kaplan (2007) provide evidence that biased media have a significant effect on the vote share in US presidential elections. We focus on when a biased source chooses to release information and show that voters respond differently to information released at different times in the election campaign.

2 The Model

In our model, Sender's payoff depends on Receiver's posterior belief about Sender's type at a deadline. We begin with a benchmark model in which (i) Sender's payoff is equal to Receiver's posterior belief, (ii) Sender is perfectly informed, (iii) Sender's type does not affect when the arm arrives, and (iv) the deadline is deterministic. Section 3.2 relaxes each of these assumptions and shows that our main results continue to hold.

2.1 Benchmark Model

There are two players: Sender (she) and Receiver (he). Sender is one of two types $\theta \in \{G, B\}$: good ($\theta = G$) or bad ($\theta = B$). Let $\pi \in (0, 1)$ be the common *prior belief* that Sender is good.

Time is discrete and indexed by $t \in \{1, 2, \ldots, T+1\}$. Sender is concerned about being perceived as good at a deadline $t = T$. In particular, the expected payoff of type $\theta \in \{G, B\}$ is given by $v_\theta(s) = s$, where s is Receiver's posterior belief at $t = T$ that $\theta = G$. Time $T+1$ combines all future dates after the deadline, including *never*.

An *arm* arrives to Sender at a random time according to distribution F with support $\{1, 2, \ldots, T+1\}$. If the arm has arrived, Sender privately observes her type and can pull the arm immediately or at any time after its arrival, including time $T+1$. Because Sender moves only after the arrival of the arm, it is immaterial for the analysis whether Sender learns her type when the arm arrives or when the game starts.

Pulling the arm starts a learning process for Receiver. Specifically, if the arm is pulled at a time τ before the deadline ($\tau \leq T$), Receiver observes realizations of a stochastic process

$$L = \{L_\theta(t; \tau), \tau \leq t \leq T\}.$$

The process L can be viewed as a sequence of signals, one per each time from τ to T with the precision of the signal at time t possibly depending on τ, t, and all previous signals.

[13] See also Prat and Stromberg (2013) for a review of this literature in the broader context of the relationship between media and politics.

Notice that if the arm is pulled at $\tau = T$, Receiver observes the realization $L_\theta(T,T)$ before taking his action. For notational convenience, we assume that L is either discrete or atomless.

It is more convenient to work directly with the distribution of beliefs induced by the process L rather than with the process itself. Recall that s is Receiver's *posterior belief* that Sender is good after observing all realizations of the process from τ to T. Let m denote Receiver's *interim belief* that Sender is good upon observing that she pulls the arm at time τ and before observing any realization of L. Given τ and m, the process L generates a distribution $H(. \mid \tau, m)$ over Receiver's posterior beliefs s; given τ, m, and θ, the process L generates a distribution $H_\theta(. \mid \tau, m)$ over s. Notice that if the arm is pulled after the deadline ($\tau = T+1$), then the distributions $H_\theta(. \mid \tau, m)$ and $H(. \mid \tau, m)$ assign probability one to $s = m$.

Assumption 1 says that (i) pulling the arm later reveals strictly less information about Sender's type in Blackwell (1953)'s sense and (ii) the learning process never fully reveals Sender's type.

Assumption 1. *(i) For all $\tau, \tau' \in \{1, 2, \ldots, T+1\}$ such that $\tau < \tau'$, $H(. \mid \tau, \pi)$ is a strict mean-preserving spread of $H(. \mid \tau', \pi)$. (ii) The support of $H(. \mid 1, \pi)$ is a subset of $(0,1)$.*

For example, consider a set of (imperfectly informative) signals $\mathcal{S}$ with some joint distribution and suppose that pulling the arm at τ reveals to Receiver a set of signals $\mathcal{S}_\tau \subset \mathcal{S}$. Assumption 1 holds whenever $\mathcal{S}_{\tau'}$ is a proper subset of $\mathcal{S}_\tau$ for all $\tau < \tau'$.

We characterize the set of perfect Bayesian equilibria, henceforth equilibria. Let $\mu(\tau)$ be Receiver's equilibrium interim belief that Sender is good given that Sender pulls the arm at time $\tau \in \{1, 2, \ldots, T+1\}$. Also, let P_θ denote an equilibrium distribution of pulling time τ given Sender's type θ (with the convention that $P_\theta(0) = 0$).

2.2 Discussion

We now pause to interpret key ingredients of our model using our main application—the timing of US presidential scandals in the lead-up to elections. Receiver is the median voter and Sender is an opposition member or organization wishing to reduce a candidate's chances of being elected. The candidate is either fit ($\theta = B$) or unfit ($\theta = G$) to run the country. The prior belief that the candidate is unfit is π. At a random time, the opposition may privately receive scandalous material against the candidate (arrival of the arm). The opposition can choose when and whether to release the material (pull the arm). After it is released, the material is subject to scrutiny, and the median voter gradually learns about the candidate's type. Crucially, the opposition has private information about what the expected outcome of scrutiny is. We say that the scandal is real (fake) if further scrutiny is likely to reveal that the candidate is unfit (fit) to run the country. If, at the time of the election (deadline), the median voter believes that the candidate is likely to be unfit to run the country, the candidate's chances of being elected are weak.

Notice that releasing a scandal might backfire. For example, before the FBI reopened its investigation over Secretary Clinton's emails, the median US voter had some belief π that Secretary Clinton had grossly mishandled classified information and was therefore unfit to be commander in chief. Further investigations could have revealed that her conduct was more than a mere procedural mistake. In this case, the median voter's posterior belief s would have been higher than π. On the contrary, the FBI might not have found any evidence of misconduct, despite investigating yet more emails. In this case, the median voter's posterior belief s would have been lower than π.

In this application, Sender's payoff depends on Receiver's belief at the deadline because this belief affects the probability that the median voter elects the candidate. Specifically, suppose that the opposition is uncertain about the ideological position r of the median voter, which is uniformly distributed on the unit interval. If the candidate is not elected, the median voter's payoff is normalized to 0. If the incumbent is elected, the median voter with position r gets payoff $r-1$ if the candidate is unfit and payoff r otherwise. The opposition gets payoff 0 if the candidate is elected and 1 otherwise. Therefore, Sender's expected payoff is given by

$$v_\theta(s) = \Pr(r \leq s) = s \text{ for } \theta \in \{G, B\}.$$

Furthermore, Receiver's expected payoff $u(s)$ is given by

$$u(s) = \int_s^1 [s(r-1) + (1-s)r]\, dr = \frac{(1-s)^2}{2}.$$

The Receiver's ex-ante expected payoff is therefore given by

$$\mathbb{E}[u(s)] = \frac{(1-\mathbb{E}[s])^2 + \mathbb{E}\left[(s-\mathbb{E}[s])^2\right]}{2} = \frac{(1-\pi)^2 + \mathrm{Var}[\mathrm{s}]}{2}. \tag{1}$$

3 Analysis

3.1 Equilibrium

We begin our analysis by deriving statistical properties of the model that rely only on players being Bayesian. These properties link the pulling time and Receiver's interim belief to the expectation of Receiver's posterior belief. First, from (good and bad) Sender's perspective, keeping the pulling time constant, a higher interim belief results in a higher expected posterior belief. Furthermore, pulling the arm earlier reveals more information about Sender's type. Therefore, from bad (good) Sender's perspective, pulling the arm earlier decreases (increases) the expected posterior belief that Sender is good. In short, Lemma 1 says that credibility is beneficial for both types of Sender, whereas scrutiny is detrimental for bad Sender but beneficial for good Sender.

Lemma 1 (Statistical Properties). *Let $\mathbb{E}[s \mid \tau, m, \theta]$ be the expectation of Receiver's posterior belief s conditional on the pulling time τ, Receiver's interim belief m, and Sender's type θ. For all $\tau, \tau' \in \{1, \ldots, T+1\}$ such that $\tau < \tau'$, and all $m, m' \in (0, 1]$ such that $m < m'$,*

1. $\mathbb{E}[s \mid \tau, m', \theta] > \mathbb{E}[s \mid \tau, m, \theta]$ *for* $\theta \in \{G, B\}$;
2. $\mathbb{E}[s \mid \tau', m, B] > \mathbb{E}[s \mid \tau, m, B]$;
3. $\mathbb{E}[s \mid \tau, m, G] > \mathbb{E}[s \mid \tau', m, G]$.

Proof. In Appendix A. $\square$

We now show that in any equilibrium, (i) good Sender strictly prefers to pull the arm whenever bad Sender weakly prefers to do so, and therefore (ii) if the arm has arrived, good Sender pulls it with certainty whenever bad Sender pulls it with positive probability.

Lemma 2 (Good Sender's Behavior). *In any equilibrium:*

1. *For all $\tau, \tau' \in \{1, \ldots, T+1\}$ such that $\tau < \tau'$ and $\mu(\tau), \mu(\tau') \in (0, 1)$, if bad Sender weakly prefers to pull the arm at τ than at τ', then $\mu(\tau) > \mu(\tau')$ and good Sender strictly prefers to pull the arm at τ than at τ';*
2. *For all $\tau \in \{1, \ldots, T\}$ in the support of P_B, we have $P_G(\tau) = F(\tau)$.*

Proof. In Appendix B. $\square$

The proof relies on the three statistical properties from Lemma 1. The key to Lemma 2 is that if bad Sender weakly prefers to pull the arm at some time τ than at $\tau' > \tau$, then Receiver's interim belief $\mu(\tau)$ must be greater than $\mu(\tau')$. Intuitively, bad Sender is willing to endure more scrutiny only if pulling the arm earlier boosts her credibility. Since $\mu(\tau) > \mu(\tau')$, good Sender strictly prefers to pull the arm at the earlier time τ, as she benefits from *both* scrutiny and credibility.

Next, we show that bad Sender pulls the arm with positive probability whenever good Sender does, but bad Sender pulls the arm later than good Sender in the first-order stochastic dominance sense. Moreover, bad sender pulls the arm strictly later unless no type pulls the arm. An immediate implication is that bad Sender always withholds the arm with positive probability.

Lemma 3 (Bad Sender's Behavior). *In any equilibrium, P_G and P_B have the same supports and, for all $\tau \in \{1, \ldots, T\}$ with $P_G(\tau) > 0$, we have $P_B(\tau) < P_G(\tau)$. Therefore, in any equilibrium, $P_B(T) < F(T)$.*

Proof. In Appendix B. □

Intuitively, if there were a time $\tau \in \{1, \ldots, T\}$ at which only good Sender pulled the arm with positive probability, then, upon observing that the arm was pulled at τ, Receiver would conclude that Sender was good. But then, to achieve this perfect credibility,[14] bad Sender would want to mimic good Sender and therefore strictly prefer to pull the arm at τ, contradicting that only good Sender pulled the arm at τ. Nevertheless, bad Sender always delays relative to good Sender. Indeed, if bad and good Sender were to pull the arm at the same time, then Sender's credibility would not depend on the pulling time. But with constant credibility, bad Sender would never pull the arm to avoid scrutiny. Therefore, good Sender must necessarily pull the arm earlier than bad Sender.

We now show that, at any time when good Sender pulls the arm, bad Sender is indifferent between pulling and not pulling the arm. That is, in equilibrium, pulling the arm earlier boosts Sender's credibility as much as to exactly offset the expected cost of longer scrutiny for bad Sender. Thus, Receiver's interim beliefs are determined by bad Sender's indifference condition (2) and the consistency condition (3). The consistency condition follows from Receiver's interim beliefs being determined by Bayes's rule and Sender's equilibrium strategy. Roughly, it says that a weighted average of interim beliefs is equal to the prior belief.

[14]By part (ii) of Assumption 1, such perfect credibility can never be dented: $H_\theta(. \mid \tau, 1)$ assigns probability 1 to $s = 1$ for all θ and τ.

Lemma 4 (Receiver's Beliefs). *In any equilibrium,*

$$\int v_B(s)\, dH_B(s|\tau, \mu(\tau)) = v_B(\mu(T+1)) \text{ for all } \tau \text{ in the support of } P_G, \tag{2}$$

$$\sum_{\tau \in supp(P_G)} \frac{1-\mu(\tau)}{\mu(\tau)} \left(P_G(\tau) - P_G(\tau-1)\right) = \frac{1-\pi}{\pi}. \tag{3}$$

Proof. In Appendix B. □

We now characterize the set of equilibria. Part 1 of Proposition 1 states that, for any set of times, there exists an equilibrium in which good Sender pulls the arm only at times in this set. Moreover, in any equilibrium, at any time when good Sender pulls the arm, she pulls it with probability 1 and bad Sender pulls it with strictly positive probability. The probability with which bad Sender pulls the arm at any time is determined by the condition that the induced interim beliefs keep bad Sender exactly indifferent between pulling the arm then and not pulling it at all. Part 2 of Proposition 1 characterizes the set of *divine equilibria* of Banks and Sobel (1987) and Cho and Kreps (1987).[15] In such equilibria, good Sender pulls the arm as soon as it arrives.

Proposition 1 (Equilibrium).

1. *For any $\mathcal{T} \subseteq \{1, \ldots, T+1\}$ with $T+1 \in \mathcal{T}$, there exists an equilibrium in which the support of P_G is $\mathcal{T}$. In any equilibrium, P_G and P_B have the same supports, and for all τ in the support of P_G, $P_G(\tau) = F(\tau)$ and*

$$P_B(\tau) = \frac{\pi}{1-\pi} \sum_{t \in supp(P_G) \text{ s.t. } t \leq \tau} \frac{1-\mu(t)}{\mu(t)} \left(P_G(t) - P_G(t-1)\right), \tag{4}$$

   where $\mu(\tau) \in (0,1)$ is uniquely determined by (2) *and* (3).

2. *There exists a divine equilibrium. In any divine equilibrium, for all $\tau \in \{1, \ldots, T+1\}$, $P_G(\tau) = F(\tau)$.*

Proof. In Appendix B. □

[15] Divinity is a standard refinement used by the signalling literature. It requires Receiver to attribute a deviation to those types of Sender who would choose it for the widest range of Receiver's interim beliefs. In our setting, the set of divine equilibria coincides with the set of *monotone* equilibria in which Receiver's interim belief about Sender is non-increasing in the pulling time. Specifically, divinity rules out all equilibria in which both types of Sender do not pull the arm at some times, because Receiver's out-of-equilibrium beliefs for those times are sufficiently unfavorable.

Although there exist a plethora of divine equilibria, in all such equilibria, pulling probabilities of good and bad Sender, as well as Receiver's beliefs, are uniquely determined by $P_G = F$ and (2)-(4). In this sense, there exists an essentially unique divine equilibrium.

Our main testable prediction is that bad Sender pulls the arm strictly later than good Sender in the likelihood ratio order sense.

Corollary 1 (Equilibrium Dynamics). *In the divine equilibrium,*

$$\frac{P_B(\tau) - P_B(\tau-1)}{P_G(\tau) - P_G(\tau-1)} < \frac{P_B(\tau+1) - P_B(\tau)}{P_G(\tau+1) - P_G(\tau)} \text{ for all } \tau \in \{1, \ldots, T\}.$$

Proof. In Appendix B. □

Corollary 1 implies that, conditional on pulling time τ being between any two times τ' and τ'', bad Sender pulls the arm strictly later than good Sender in the first-order stochastic dominance sense (Theorem 1.C.5, Shaked and Shanthikumar, 2007):

$$\frac{P_B(\tau) - P_B(\tau')}{P_B(\tau'') - P_B(\tau')} < \frac{P_G(\tau) - P_G(\tau')}{P_G(\tau'') - P_G(\tau')} \text{ for all } \tau' < \tau < \tau''.$$

Our model also gives predictions about the evolution of Receiver's beliefs. Pulling the arm earlier is more credible as Receiver's interim beliefs $\mu(\tau)$ decrease over time. Moreover, pulling the arm instantaneously boosts credibility in the sense that Receiver's belief at any time τ about Sender's type is higher if Sender pulls the arm than if she does not.

Corollary 2 (Belief Dynamics). *Let $\tilde{\mu}(\tau)$ denote Receiver's interim belief that Sender is good given that she has not pulled the arm before or at τ. In the divine equilibrium,*

$$\mu(\tau-1) > \mu(\tau) > \tilde{\mu}(\tau-1) > \tilde{\mu}(\tau) \text{ for all } \tau \in \{2, \ldots, T\}.$$

Proof. In Appendix B. □

3.2 Discussion of Model Assumptions

We now discuss how our results change (or do not change) if we relax several of the assumptions made in our benchmark model. We discuss each assumption in a separate subsection. The reader may skip this section without any loss of understanding of subsequent sections.

3.2.1 Nonlinear Sender's Payoff

In the benchmark model, we assume that Sender's payoff is linear in Receiver's posterior belief: $v_G(s) = v_B(s) = s$ for all s. In our motivating example, this linearity arises because the opposition is uncertain about the ideological position r of the median voter. If there is no such uncertainty, then the median voter reelects the incumbent whenever s is below r, where $r \in (0,1)$ is a constant. Therefore, Sender's payoff is a step function:

$$v_\theta(s) = v(s) = \begin{cases} 0 & \text{if } s < r; \\ 1 & \text{if } s > r. \end{cases} \tag{5}$$

We now allow for Sender's payoff to be non-linear in Receiver's posterior belief and even type dependent. To understand how the shapes of the payoff functions v_G and v_B affect our analysis, we extend the statistical properties of Lemma 1, which describe the evolution of Receiver's posterior belief from Sender's perspective. First and not surprisingly, a more favorable interim belief results in more favorable posterior beliefs for all types of Sender and for all realizations of the process. Moreover, Receiver's posterior belief follows a supermartingale (submartingale) process from bad (good) Sender's perspective. Lemma 1′ formalizes these statistical properties, using standard stochastic orders (see, e.g., Shaked and Shanthikumar, 2007). Distribution Z_2 strictly dominates distribution Z_1 in the increasing convex (concave) order if there exists a distribution Z such that Z_2 strictly first-order stochastically dominates Z and Z is a mean-preserving spread (reduction) of Z_1.

Lemma 1′ (Statistical Properties). *For all $\tau, \tau' \in \{1, \ldots, T+1\}$ such that $\tau < \tau'$, and all $m, m' \in (0,1]$ such that $m < m'$,*

1. *$H_\theta(. \mid \tau, m')$ strictly first-order stochastically dominates $H_\theta(. \mid \tau, m)$ for $\theta \in \{G, B\}$;*
2. *$H_B(. \mid \tau', m)$ strictly dominates $H_B(. \mid \tau, m)$ in the increasing concave order;*
3. *$H_G(. \mid \tau, m)$ strictly dominates $H_G(. \mid \tau', m)$ in the increasing convex order.*

Proof. In Appendix A. □

To interpret Lemma 1′, we assume that the payoff of both types of Sender is a continuous strictly increasing function of Receiver's posterior belief, so that both types of Sender want to look good.[16] Part 1 says that credibility is beneficial for both types of Sender, regardless of the shape of their payoff functions. Part 2 (part 3) says that from bad (good) Sender's perspective, pulling the arm earlier results in more spread out and less (more) favorable posteriors provided that the interim belief does not depend on the pulling time. So scrutiny is detrimental for bad Sender if her payoff is not too convex but beneficial for good Sender if her payoff is not too concave. Therefore, for a given process satisfying Assumption 1, Proposition 1 continues to hold if bad Sender is not too risk-loving and good Sender is not too risk-averse. In fact, Proposition 1 continues to hold verbatim if bad Sender's payoff is weakly concave and good Sender's payoff is weakly convex (the proof in Appendix B explicitly allows for this possibility).[17]

Much less can be said in general if the payoff functions v_G and v_B have an arbitrary shape. For example, if v_G is sufficiently concave, then good Sender can prefer to delay pulling the arm to reduce the spread in posterior beliefs. Likewise, if v_B is sufficiently convex, then bad Sender can prefer to pull the arm earlier than good Sender to increase the spread in posterior beliefs. These effects work against our credibility-scrutiny tradeoff and Proposition 1 no longer holds.[18] Nevertheless, bad Sender weakly delays pulling the arm relative to good Sender under the following single crossing assumption.

Assumption 2. *For all $\tau, \tau' \in \{1, \dots, T+1\}$ such that $\tau < \tau'$ and $\mu(\tau), \mu(\tau') \in (0,1)$, if bad Sender weakly prefers to pull the arm at τ than at τ', then good Sender strictly prefers to pull the arm at τ than at τ'.*

This assumption holds in the benchmark model by Lemma 2. This assumption also holds if Sender's payoff is the step function in (5) whenever pulling the arm later reveals strictly less *useful* information about Sender's type, in the sense that Receiver is strictly worse off.[19]

[16] It is sufficient for our results to assume that Sender's payoff is an upper hemicontinuous correspondence (rather than a continuous function) of Receiver's posterior belief. For example, this is the case if Sender's and Receiver's payoffs depend on Receiver's action and Sender's type, and Receiver's action set is finite. In the above example with constant ideological position, Sender's payoff in (5) is a correspondence with $v(r) = [0,1]$, because it is optimal for Receiver to randomize between the two actions when $s = r$.

[17] More generally, Proposition 1 holds whenever $sv_G(s)$ is strictly convex and $(1-s)\,v_B(s)$ is strictly concave, i.e., for all s, Sender's Arrow-Prat coefficient of absolute risk aversion $-v''_\theta(s)/v'_\theta(s)$ is less than $2/s$ for good Sender and more than $-2/(1-s)$ for bad Sender. For the Poisson model of Section 4, Proposition 1 continues to hold for any risk attitude of good Sender and only relies on bad Sender being not too risk-loving.

[18] These effects are common in the Bayesian persuasion literature (Kamenica and Gentzkow, 2011). In this literature, Sender is uninformed. Therefore, from her perspective, Receiver's beliefs follow a martingale process (Ely et al., 2015), so only convexity properties of Sender's payoff affect the time at which she pulls the arm.

[19] The inequality (6) holds if and only if $\int_x^1 H(s \mid \tau', \pi)\,ds > \int_x^1 H(s \mid \tau, \pi)\,ds$ for all $x \in (0,1)$. In com-

Lemma 2′ (Good Sender's Behavior). *Let v_θ be given by* (5). *If for all $\tau, \tau' \in \{1, \ldots, T+1\}$ such that $\tau < \tau'$*

$$\int_r^1 H\left(s \mid \tau', m\right) ds > \int_r^1 H\left(s \mid \tau, m\right) ds \text{ for all } m \in (0,1), \tag{6}$$

then Assumption 2 holds.

Proof. In Appendix B. □

If Assumption 2 holds and v_θ is strictly increasing, then in the unique divine equilibrium, good Sender pulls the arm as soon as it arrives and bad Sender pulls the arm *weakly* later than good Sender—there may exist an equilibrium in which both good and bad Sender pull the arm as soon as it arrives.[20]

3.2.2 Imperfectly Informed Sender

In many applications, Sender does not know with certainty whether pulling the arm would start a good or bad learning process for Receiver. For example, when announcing the reopening of the Clinton investigation, Director Comey could not know for certain what the results of the investigation would eventually be.

We generalize our model to allow for Sender to only observe a signal $\sigma \in \{\sigma_B, \sigma_G\}$ about an underlying binary state θ, with normalization

$$\sigma_G = \Pr\left(\theta = G \mid \sigma_G\right) > \pi > \Pr\left(\theta = G \mid \sigma_B\right) = \sigma_B.$$

The statistical properties of Lemma 1 still hold.

Lemma 1″ (Statistical Properties). *Let $\mathbb{E}\left[s \mid \tau, m, \sigma\right]$ be the expectation of Receiver's posterior belief s conditional on the pulling time τ, Receiver's interim belief m, and Sender's signal σ. For all $\tau, \tau' \in \{1, \ldots, T+1\}$ such that $\tau < \tau'$, and all $m, m' \in (0,1]$ such that $m < m'$,*

1. $\mathbb{E}\left[s \mid \tau, m', \sigma\right] > \mathbb{E}\left[s \mid \tau, m, \sigma\right]$;
2. $\mathbb{E}\left[s \mid \tau', m, \sigma_B\right] > \mathbb{E}\left[s \mid \tau, m, \sigma_B\right]$;

parison, part (i) of Assumption 1 holds if and only if $\int_x^1 H\left(s \mid \tau', \pi\right) ds \geq \int_x^1 H\left(s \mid \tau, \pi\right) ds$ for all $x \in (0,1)$ with strict inequality for some $x \in (0,1)$.

[20] For v_θ given by (5), we can show that bad Sender withholds the arm with strictly positive probability, $P_B(T) < F(T)$, in all divine equilibria, if $\pi > r$. In this case, however, v_θ is not *strictly* increasing, and there exist divine equilibria in which good Sender does not always pull the arm as soon as it arrives. For example, there exists a divine equilibrium in which bad and good Sender never pull the arm by the deadline: $P_G(T) = P_B(T) = 0$. In this equilibrium, both bad and good Sender enjoy the highest possible payoff, 1.

3. $\mathbb{E}[s \mid \tau, m, \sigma_G] > \mathbb{E}[s \mid \tau', m, \sigma_G]$.

Proof. In Appendix A. □

These statistical results ensure that credibility is always beneficial for Sender, whereas scrutiny is detrimental for Sender with signal σ_B but beneficial for Sender with signal σ_G. Therefore, all our results carry over.

Moreover, we can extend our analysis to allow for signal σ to be continuously distributed on the interval $[\underline{\sigma}, \bar{\sigma})$, with normalization $\sigma = \Pr(\theta = G \mid \sigma)$. In particular, in this case, there exists a partition equilibrium with $\bar{\sigma} = \sigma_0 > \sigma_1 > \cdots > \sigma_{T+1} = \underline{\sigma}$ such that Sender $\sigma \in [\sigma_t, \sigma_{t-1})$ pulls the arm as soon as it arrives unless it arrives before time $t \in \{1, \ldots, T+1\}$ (and pulls the arm at time t if it arrives before t).

3.2.3 Type-Dependent Arrival of the Arm

In many applications, it is more reasonable to assume that the distribution of the arrival of the arm differs for good and bad Sender. For example, fake scandals may be easy to fabricate, whereas real scandals need time to be discovered.

We generalize the model to allow for different distributions of the arrival of the arm for good and bad Sender. In particular, the arm arrives at a random time according to distributions $F_G = F$ for good Sender and F_B for bad Sender.

The proof of Proposition 1 (in Appendix B) explicitly allows for the arm to arrive (weakly) earlier to bad Sender than to good Sender in the first-order stochastic dominance sense: $F_B(t) \geq F_G(t)$ for all t. This assumption is clearly satisfied if bad Sender has the arm from the outset or if bad and good Sender receive the arm at the same time.

More generally, Proposition 1 continues to hold verbatim unless the arm arrives sufficiently later to bad Sender than to good Sender such that $F_B(t) < P_B(t)$ for some t, where $P_B(t)$ is given by (4). But even then, Corollary 1 still holds. That is, bad Sender pulls the arm strictly later than good Sender. Yet, bad Sender may do so for the simple mechanical (rather than strategic) reason that the arm arrives to her later than to good Sender.[21]

3.2.4 Stochastic Deadline

In the benchmark model, we assume that the deadline T is fixed and common knowledge. In some applications, the deadline T may be stochastic. In particular, suppose that T is a random variable distributed on $\{1, \ldots, \bar{T}\}$ where time runs from 1 to $\bar{T}+1$. Now the

[21] In this case, there exist some time τ at which bad Sender strictly prefers to pull the arm and (2) no longer holds for τ.

process L has T as a random variable rather than a constant. For this process, we can define the ex-ante distribution H of posteriors at T, where H depends only on pulling time τ and interim belief m. Notice that Assumption 1 still holds for this ex-ante distribution of posteriors for any $\tau, \tau' \in \{1, \ldots, \bar{T}+1\}$. Therefore, from the ex-ante perspective, Sender's problem is identical to the problem with a deterministic deadline and all results carry over.

4 Poisson Model

To get more precise predictions about the strategic timing of information release, we now assume that the arrival of the arm and Receiver's learning follow Poisson processes. In this Poisson model, time is continuous $t \in [0, T]$.[22] The arm arrives to Sender at Poisson rate α, so that $F(t) = 1 - e^{-\alpha t}$. Once Sender pulls the arm, a *breakdown* occurs at Poisson rate λ if Sender is bad, but never occurs if Sender is good, so that $H(. \mid \tau, m)$ puts probability $(1-m)\left(1 - e^{-\lambda(T-\tau)}\right)$ on $s = 0$ and the complementary probability on

$$s = \frac{m}{m + (1-m)\, e^{-\lambda(T-\tau)}}.$$

Returning to our main application, the Poisson model assumes that scandals can be conclusively debunked, but cannot be proven real. It also assumes that the opposition receives the scandalous material against the president at a constant rate, independent of whether they are real or fake. As discussed in Section 3.2.3, the results would not change if real documents take more time to be discovered than fake document take to be fabricated.

Our benchmark model does not completely nest the Poisson model. In fact, part (ii) of Assumption 1, that the learning process never fully reveals Sender's type, fails in the Poisson model because a breakdown fully reveals that Sender is bad. Nevertheless, a version of Proposition 1 continues to hold without this assumption, with the difference that bad Sender never pulls the arm before some time $\bar{t}$. Specifically, Proposition 1 holds for all $\tau \geq \bar{t}$, whereas $\mu(\tau) = 1$ and $P_B(\tau) = 0$ for all $\tau < \bar{t}$. Intuitively, even if Receiver believes that only good Sender pulls the arm before $\bar{t}$, bad Sender strictly prefers to pull the arm after $\bar{t}$ to reduce the risk that Receiver fully learns that Sender is bad.

We can, therefore, explicitly characterize the divine equilibrium of the Poisson model. First, good Sender pulls the arm as soon as it arrives.[23] Second, bad Sender is indifferent

[22] Technically, we use the results from Section 3.1 by treating continuous time as an appropriate limit of discrete time.

[23] In every divine equilibrium, $P_G(t) = F(t)$ for all $t \in [\bar{t}, T]$ and $P_B(t) = 0$ for all $t \in [0, \bar{t}]$. But for each

between pulling the arm at any time $t \geq \bar{t} \geq 0$ and not pulling it at all. Third, bad Sender strictly prefers to delay pulling the arm if $t < \bar{t}$.

In the divine equilibrium of the Poisson model, $\mu(t) = 1$ for all $t < \bar{t}$, and equations (2) and (3) become

$$\begin{aligned}
\frac{\mu(t) e^{-\lambda(T-t)}}{\mu(t) + (1-\mu(t)) e^{-\lambda(T-t)}} &= \mu(T) \text{ for all } t \geq \bar{t}, \\
\int_0^T \alpha \frac{1-\mu(t)}{\mu(t)} e^{-\alpha t} dt + \frac{1-\mu(T)}{\mu(T)} e^{-\alpha T} &= \frac{1-\pi}{\pi}.
\end{aligned}$$

Adding the boundary condition $\lim_{t \downarrow \bar{t}} \mu(t) = 1$ yields the explicit solution $\mu(t)$ and uniquely determines $\bar{t}$.

Proposition 2. *In the divine equilibrium, good Sender pulls the arm as soon as it arrives and Receiver's interim belief that Sender is good given pulling time t is:*

$$\mu(t) = \begin{cases} \frac{\mu(T)}{1-\mu(T)\left(e^{\lambda(T-t)}-1\right)} & \text{if } t \geq \bar{t}; \\ 1 & \text{otherwise,} \end{cases}$$

where $\mu(T)$ is Receiver's posterior belief if the arm is never pulled and

$$\bar{t} = \begin{cases} 0 & \text{if } \pi < \bar{\pi}; \\ T - \frac{1}{\lambda} \ln \frac{1}{\mu(T)} & \text{otherwise,} \end{cases}$$

$$\begin{aligned}
\mu(T) &= \begin{cases} \left[\frac{\alpha e^{\lambda T} + \lambda e^{-\alpha T}}{\alpha+\lambda} + \frac{1-\pi}{\pi}\right]^{-1} & \text{if } \pi < \bar{\pi}; \\ \left[\frac{(\alpha+\lambda)(1-\pi)}{\lambda\pi} e^{\alpha T} + 1\right]^{-\frac{\lambda}{\alpha+\lambda}} & \text{otherwise,} \end{cases} \\
\bar{\pi} &= \left[1 + \frac{\lambda}{\alpha+\lambda}\left(e^{\lambda T} - e^{-\alpha T}\right)\right]^{-1}.
\end{aligned}$$

The parameters of the model affect welfare directly and through Sender's equilibrium behavior. Proposition 3 says that, in the divine equilibrium, direct effects dominate. Specifically, a higher prior belief π results in higher posterior beliefs, which increases both bad and good Sender's welfare. Moreover, a higher breakdown rate λ or a higher arrival rate α allows Receiver to learn more about Sender, which decreases (increases)

distribution $\hat{P}$ such that $\hat{P}(t) \leq F(t)$ for all $t \in [0, \bar{t})$ and $\hat{P}(t) = F(t)$ for all $t \in [\bar{t}, T]$, there exists a divine equilibrium with $P_G = \hat{P}$. For ease of exposition, we focus on the divine equilibrium in which $P_G(t) = F(t)$ for all $t \in [0, T]$.

bad (good) Sender's welfare. Proposition 3 also derives comparative statics on Receiver's welfare given by (1).[24]

Proposition 3. *In the divine equilibrium,*

1. *the expected payoff of bad Sender increases with π but decreases with λ and α;*
2. *the expected payoff of good Sender increases with π, λ, and α;*
3. *the expected payoff of Receiver decreases with π but increases with λ and α.*

Proof. In Appendix C. □

4.1 Static Analysis

We now explore how the parameters of the model affect the probability that Sender releases information. The probability that bad Sender pulls the arm is

$$P_B(T) = 1 - \frac{\pi}{1-\pi}\frac{1-\mu(T)}{\mu(T)}e^{-\alpha T}, \tag{7}$$

which follows from

$$\mu(T) = \frac{\pi e^{-\alpha T}}{\pi e^{-\alpha T} + (1-\pi)(1-P_B(T))}. \tag{8}$$

Proposition 4 says that bad Sender pulls the arm with a higher probability if the prior belief π is lower, if the breakdown rate λ is lower, or if the arrival rate α is higher.

Proposition 4. *In the divine equilibrium, the probability that bad Sender pulls the arm decreases with π and λ but increases with α.*

Proof. In Appendix C. □

Intuitively, if the prior belief π is higher, bad Sender has more to lose in case of a breakdown. Similarly, if the breakdown rate λ is higher, pulling the arm is more likely to reveal that Sender is bad. In both cases, bad Sender is more reluctant to pull the arm. In contrast, if the arrival rate α is higher, good Sender is more likely to pull the arm and

[24] For an arbitrary Receiver's Bernoulli payoff function, which depends on Receiver's action and Sender's type, the second-order Taylor approximation of Receiver's expected payoff increases with the variance of his posterior belief, and therefore with λ and α. In contrast, the comparative statics with respect to π are less robust. For example, if Receiver's Bernoulli payoff function is $-(a-\theta)^2$, where $a \in \mathbb{R}$ is Receiver's action and $\theta \in \{0,1\}$ is Sender's type, then Receiver's expected payoff decreases with π for $\pi < 1/2$ and increases with π for $\pi > 1/2$.

Receiver will believe that Sender is bad with higher probability if she does not pull the arm. In this case, bad Sender is more willing to pull the arm.

The total probability that Sender pulls the arm is given by the weighted sum of the probabilities $P_B(T)$ and $P_G(T)$ that bad Sender and good Sender pull the arm:

$$P_B(T) = \pi P_G(T) + (1-\pi) P_B(T) = 1 - \frac{\pi e^{-\alpha T}}{\mu(T)}. \tag{9}$$

A change in λ affects $P_B(T)$, but not $P_G(T)$; a change in α affects both $P_B(T)$ and $P_G(T)$ in the same direction. Therefore, Sender pulls the arm with a higher total probability if the breakdown rate is lower or if the arrival rate is higher. The prior belief π has a direct and an indirect effect on the total probability that Sender pulls the arm. On the one hand, holding $P_B(T)$ constant, $P(T)$ directly increases with π, because $P_G(T) > P_B(T)$. On the other hand, $P(T)$ indirectly decreases with π, because $P_B(T)$ decreases with π. Proposition 5 says that the indirect effect dominates the direct effect when π is sufficiently low.

Proposition 5. *In the divine equilibrium, the total probability that Sender pulls the arm decreases with λ, increases with α, and is quasiconvex in π: decreases with π if*

$$\pi < \frac{\alpha e^{\alpha T}}{\alpha e^{\alpha T} + \lambda\left(e^{\alpha T} - 1\right)} \in (0,1)$$

and increases with π otherwise.

Proof. In Appendix C. □

The probabilities $P_G(T) = 1 - e^{-\alpha T}$ and $P_B(T)$ that good Sender and bad Sender pull the arm also determine Receiver's posterior belief $\mu(T)$. By (8), $\mu(T)$ decreases with the breakdown rate λ, because $P_B(T)$ decreases with λ. Equation (8) also suggests that there are direct and indirect effects of the prior belief π and the arrival rate α on $\mu(T)$. On the one hand, holding $P_B(T)$ constant, $\mu(T)$ directly increases with π and decreases with α. On the other hand, $\mu(T)$ indirectly decreases with π and increases with α, because $P_B(T)$ decreases with π and increases with α. Proposition 6 says that the direct effect always dominates the indirect effect in the Poisson model.

Proposition 6. *In the divine equilibrium, Receiver's posterior belief if the arm is never pulled increases with π but decreases with λ and α.*

Proof. In Appendix C. □

4.2 Dynamic Analysis

The Poisson model also allows for a more detailed analysis of the strategic timing of information release. By Proposition 2, bad Sender begins to pull the arm at time $\bar{t}$. In the spirit of Proposition 4, bad Sender begins to pull the arm later if the prior belief π is higher, if the breakdown rate λ is higher, or if the arrival rate α is lower.

Proposition 7. *In the divine equilibrium, $\bar{t}$ increases with π and λ but decreases with α.*

At each time t after $\bar{t}$, bad Sender pulls the arm with a strictly positive probability density $p_B(t)$ (see Figure 2a). Proposition 8 says that $p_B(t)$ first increases and then decreases with time.

Proposition 8. *In the divine equilibrium, the probability density that bad Sender pulls the arm at time t is quasiconcave: increases with t if*

$$t < t_b \equiv T - \frac{1}{\lambda} \ln\left(\frac{\alpha}{\alpha+\lambda} \frac{1}{\mu(T)}\right)$$

and decreases with t otherwise.

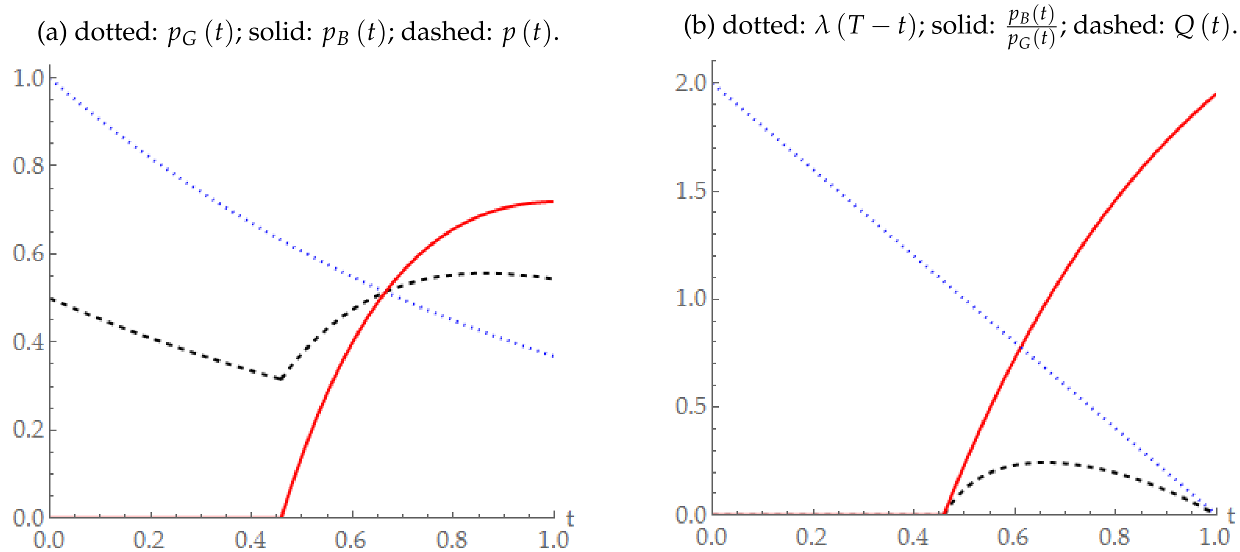
Proof. In Appendix C. □

Intuitively, the dynamics of $p_B(t)$ are driven by a strategic and a mechanic force. Strategically, as in Corollary 1, bad Sender delays pulling the arm with respect to good Sender, so that the likelihood ratio $p_B(t)/p_G(t)$ increases with time, where $p_G(t) = \alpha e^{-\alpha t}$ is the probability density that good Sender pulls the arm at time t. Mechanically, $p_B(t)$ roughly follows the dynamics of $p_G(t)$. If the arrival rate α is sufficiently small, so that the density $p_G(t)$ barely changes over time, the strategic force dominates and the probability that bad sender pulls the arm monotonically increases with time ($t_b > T$). Instead, if the arrival rate α is sufficiently large, so that $p_G(t)$ rapidly decreases over time, the mechanic force dominates and the probability that bad sender pulls the arm monotonically decreases with time ($t_b < T$).

The total probability density $p(t)$ that Sender pulls the arm is a weighted sum of $p_G(t)$ and $p_B(t)$, so that $p(t) = \pi p_G(t) + (1-\pi) p_B(t)$ (see Figure 2a). Therefore, until $\bar{t}$, $p(t) = \pi p_G(t)$, and thereafter, as in Proposition 8, $p(t)$ first increases and then decreases with time.

Proposition 9. *In the divine equilibrium, the total probability density that Sender pulls the arm at time t decreases with t from 0 to $\bar{t}$ and is quasiconcave in t on the interval $[\bar{t}, T]$: increases with*

Figure 2: Pulling density and breakdown probability; $\alpha = 1$, $\lambda = 2$, $\pi = .5$, $T = 1$.



t if

$$\bar{t} < t < t_s \equiv T - \frac{1}{\lambda} \ln\left(\frac{\alpha}{\alpha+\lambda} \frac{1+\mu(T)}{\mu(T)}\right)$$

and decreases with t if $t > t_s$.

Proof. In Appendix C. □

Let the *breakdown probability* $Q(t)$ be the probability that a breakdown occurs before the deadline given that the arm is pulled at time t (see Figure 2b). Proposition 10 says that, as time passes, the breakdown probability first increases and then decreases.[25]

Proposition 10. *In the divine equilibrium, the breakdown probability is quasiconcave: increases with t if*

$$t < t_b \equiv T - \frac{1}{\lambda} \ln\left(\frac{1+\mu(T)}{2\mu(T)}\right) < T$$

and decreases with t otherwise.

Proof. In Appendix C. □

Intuitively, the breakdown probability increases with the amount of scrutiny and with the likelihood ratio $p_B(t)/p_G(t)$. Obviously, Sender is exposed to more scrutiny if she pulls the arm earlier. But the likelihood ratio $p_B(t)/p_G(t)$ is lower earlier, because bad

[25]If the arrival rate α is sufficiently small, then t_b is negative and hence the breakdown probability monotonically decreases with time.

Sender strategically delays pulling the arm. Proposition 10 says that this strategic effect dominates for earlier times.

5 Applications

5.1 US Presidential Scandals

Returning to our presidential scandals example, the main prediction of our model is that fake scandals are released later than real scandals. We explore this prediction using Nyhan's (2015) data on US presidential scandals from 1977 to 2008. For each week, the data report whether a new scandal involving the current US president was first mentioned in the Washington Post. Although scandals might have first appeared on other outlets, we agree with Nyhan that the Washington Post is likely to have mentioned such scandals immediately thereafter. As our model concerns scandals involving the incumbent in view of his possible reelection, we focus on all the presidential elections in which the incumbent was a candidate. Therefore we consider only the first term of each president from 1977 to 2008, beginning on the first week of January after the president's election.[26,27] In all cases, the election was held on the 201st week after this date. We construct the variable *weeks to election* as the difference between the election week at the end of the term and the release week of the scandal.

For each scandal,[28] we locate the original Washington Post article as well as other contemporary articles on The New York Times and the Los Angeles Times. We then search for subsequent articles on the same scandal in following years until 2016, as well as court decisions and scholarly books when possible. We check whether factual evidence of wrongdoing or otherwise reputationally damaging conduct was conclusively verified at a later time. If so, we check whether the evidence involved the president directly or close family members or political collaborators chosen or appointed by the president or his administration. We code these scandals as *real*. For the remaining scandals, we check whether

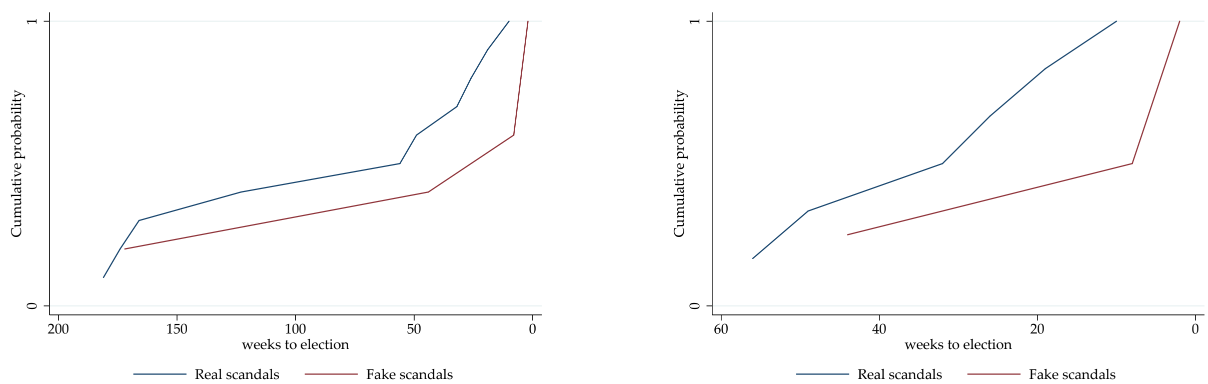
[26]This corresponds to the first terms of five presidents: Jimmy Carter (1976-1980), Ronald Reagan (1980-1984), George H. W. Bush (1988-1992), Bill Clinton (1992-1996), and George W. Bush (2000-2004). Each president run for reelection and three (Reagan, Clinton, and Bush) served two full terms.

[27]Nyhan (2015) does not provide data on scandals involving the president-elect between Election Day and the first week of January of the following year, but it contains data on scandals involving the president-elect between the first week of January and the date of his inauguration: there are no such scandals.

[28]We omit from our sample the "GSA corruption" scandal during Jimmy Carter's presidency as the allegations, explicit and implicit, of the scandal, while involving the federal administration, did not involve any of the member of Carter's administration or their collaborators (if anything, as Carter run with the promise to end corruption in the GSA, the scandal might have actually reinforced his position). In any case, we check in Additional Material A that our qualitative results are robust to the inclusion of this scandal.

Figure 3: US presidential scandals and weeks to election. Distribution of real and fake scandals.

(a) Whole term. (b) Last 60 weeks only.

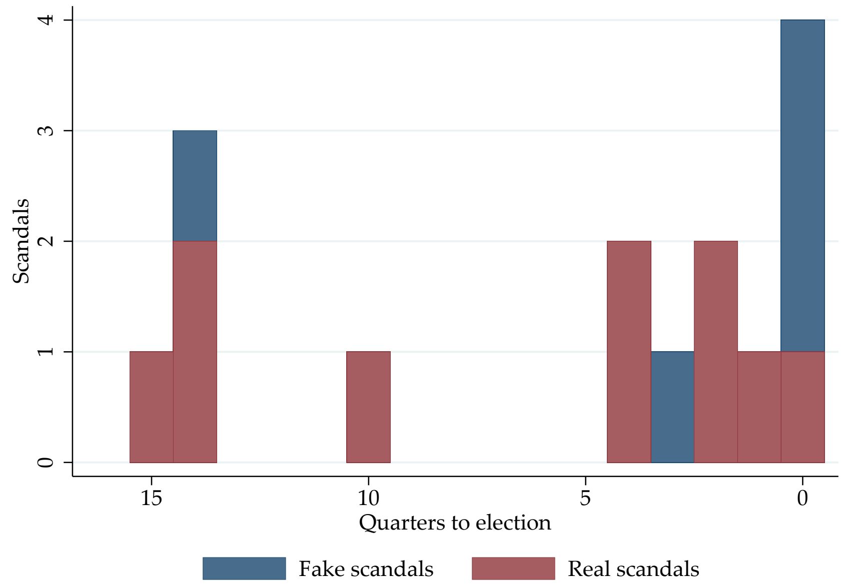


a case for libel was successful or all political actors linked to the scandal were cleared of wrongdoings. We code these scandals as *fake*. The only scandal we were not able to code by this procedure is the "Banca Nazionale del Lavoro" scandal (also known as "Iraq-gate"). We code this scandal as real, but we check in Additional Material A that all our qualitative results are robust to coding it as fake. In Additional Material A, we report the complete list of scandals and a summary motivation of our coding decisions.

Figure 3 shows the empirical distributions of the first mention of real and fake presidential scandals in the Washington Post as a function of *weeks to election*. Although we do not observe scandals released after the election ($t = T + 1$ in our model) and cannot pinpoint the date at which the campaign begins ($t = 1$ in our model), Corollary 1 implies that fake scandals are released later than real scandals conditional on any given time interval. The left panel covers the whole presidential term; the right panel focuses on the election campaign period only, which we identify with the last 60 weeks before the election. Both figures suggest that fake scandals are released later than real scandals. Because of the small sample size (only 15 scandals), formal tests have low power. Nevertheless, using the Dardanoni and Forcina (1998) test for the likelihood ratio order (which implies first order stochastic dominance), we almost reject the hypothesis that the two distributions are equal in favor of the alternative hypothesis that fake scandals are released later (p-value 0.114); we cannot reject the hypothesis that fake scandals are released later in favor of the unrestricted hypothesis at all standard statistical significance levels (p-value: 0.834).[29]

[29] We discuss this test in greater detail in the context of the next application. For this application, we use $k = 3$ equiprobable time intervals. For the election campaign period only (10 scandals), the p-values are

Figure 4: Distribution of real and fake scandals.



Our Poisson model offers a novel perspective over the October surprise concentration of scandals towards the end of the presidential election campaign (see Figure 1). In equilibrium, real scandals are released as they are discovered by the media. Unless real scandals are more likely to be discovered towards the end of the first term of a president, then we should not expect their release to be concentrated towards the end of the campaign (see $p_G(t)$ in Figure 2a). Instead, fake scandals are strategically delayed, and so they should be concentrated towards the end of the first term of the president and just before the election (see $p_B(t)$ in Figure 2a). In other words, our model predicts that the October surprise effect is driven by fake scandals. In contrast, were the October surprise effect driven by the desire to release scandals when they are most salient, then the timing of release of real and fake scandals would be similar. Figure 4 is a replica of Figure 1, but with scandals coded as real and fake. Fake scandals are concentrated close to the election, with a majority of them released in the last quarter before the election. In contrast, real scandals appear to be scattered across the entire presidential term.

Our Poisson model also predicts how different parameters affect the release of a US presidential scandal. We now illustrate how Nyhan's (2015) empirical findings may be interpreted using our model. Nyhan (2015) finds that scandals are more likely to appear when the president's opposition approval rate is low. In our model, the approval rate is most naturally captured by the prior belief $1-\pi$ (the belief that the president is fit to run

0.003 and 0.839, respectively.

the country). In our Poisson model, a higher π has a direct and an indirect effect on the probability of release of a scandal. On the one hand, a higher π means that the president is more likely to be involved in a real scandal, thus directly increasing the probability that such a scandal is released. On the other hand, the opposition optimally resorts to fake scandals more when the president is so popular that only a scandal could prevent the president's reelection. Therefore, a higher π reduces the incentive for the opposition to release fake scandals, indirectly reducing the probability that a scandal is released. We can then interpret Nyhan's finding as suggesting that the direct effect on average dominates the indirect effect.

But the president's opposition approval rate also measures opposition voters' hostility towards the president, which might be captured by the rate λ at which voters learn that a scandal is fake.[30] Indeed, Nyhan conjectures that when opposition voters are more hostile to the president, then they are "supportive of scandal allegations against the president and less sensitive to the evidentiary basis for these claims [and] opposition elites will be more likely to pursue scandal allegations" (p. 6). Consistently, in our Poisson model, when voters take more time to tell real and fake scandals apart, the opposition optimally resorts to fake scandals.

Nyhan (2015) also finds that fewer scandals involving the president are released when the news agenda is more congested. Such media congestion may have the following two effects. First, when the news agenda is congested, the opposition media has less time to devote to investigate the president. In our Poisson model, this is captured by a lower arrival rate α, which in turn reduces the probability that a scandal is released. Second, when the news agenda is congested, public scrutiny of the scandal is slower as the attention of media, politicians, and voters are captured by other events. In our Poisson model, this is captured by a lower breakdown rate λ, which in turn increases the probability that a scandal is released. We can interpret Nyhan's finding as suggesting that the media congestion effect through the arrival rate α dominates the effect through the breakdown rate λ.[31]

[30] The rate of learning λ might also be related to the verifiability of information, which may depend on the scandal's type (e.g., infidelity vs. corruption).

[31] Another possible explanation (not captured by our model) for Nyhan's finding is that media organizations strategically avoid releasing scandals when voters' attention is captured by other media events and scandals may be less effective (see Durante and Zhuravskaya, 2016).

5.2 Initial Public Offerings

We now apply our model to the timing of Initial Public Offerings (IPOs). Sender is a firm that needs liquidity in a particular *time frame*, and this time frame is private information of the firm. The need for liquidity could arise from the desire to grow the firm, expand into new products or markets, or because of operating expenses outstripping revenues. It could also arise because of investors having so-called "drag-along rights," where they can force founders and other shareholders to vote in favor of a liquidity event.

When announcing the IPO, firms have private information regarding their prospective long-run performance. *Good* firms expect their business to out-perform the market's prior expectation; *bad* firms expect their business to under-perform the market's prior expectation. After a firm announces an IPO, the market scrutinizes the firm's prospectus, and learns about the firm's prospective performance. The initial trading price of the stock is determined by the market's posterior belief at the initial trading date.[32] Therefore, after the initial trading date, as the firms' potential is gradually revealed to the market, good firms' stocks out-perform bad firms' stocks.

Since the true time frame is private information, the firm can "pretend" to need liquidity faster than it actually does, and it has significant control over the *time gap* between the announcement of the IPO and the initial trading date. A shorter time gap decreases the amount of scrutiny the firm undergoes before going public, but also reduces credibility. Therefore, our model predicts that bad firms should choose a shorter time gap than good firms.[33]

We test this prediction using data on US IPOs from 1983 to 2016. For each IPO, we record the time gap and calculate the cumulative monthly returns of the stock, starting from the first trading date. We measure the stock's performance as its *excess returns* over the value-weighted return of all CRSP firms incorporated in the US and listed on the NYSE, AMEX or NASDAQ, over the same period. As Ritter (2003) highlights, it is customary to look at five years post IPO to assess long-run performance. We measure the performance of IPOs after $y \in \{5, \dots, 10\}$ years since the launch date. For each value of t

[32] In our empirical analysis, we use the closing price on the first day of trading of the stock, as in Ritter (2003).

[33] One way to map this application into our model is as follows. Suppose that a firm learns at date t_ℓ that it needs liquidity in a time frame Δ_F, meaning that the latest possible initial trading date is $t_\ell + \Delta_F$. Both t_ℓ and Δ_F are privately known by the firm. Date t_ℓ is drawn according to the (improper) uniform distribution on the set of integers $\mathbb{Z}$. The time frame is $\Delta_F \equiv T - t$, where t has a distribution F on $\{1, \dots, T+1\}$. The firm chooses a time gap $\Delta_G \equiv T - \tau$ subject to $\tau \in \{t, \dots, T+1\}$, meaning that it announces an IPO at a date $t_a \in \{t_\ell, \dots, t_\ell + (\Delta_F - \Delta_G)\}$ with the initial trading date at $t_a + \Delta_G \leq t_\ell + \Delta_F$. Announcing an IPO at date t_a with the initial trading date at $t_a - 1$ means that the firm accesses liquidity through other channels than an IPO. With this mapping, all our results hold exactly with P_G and P_B being the distributions of $\tau = T - \Delta_G$ for good and bad firms, respectively.

Figure 5: US IPOs and time gap. Distributions for good and bad IPOs.

(a) 6 years. (b) 8 years. (c) 10 years.

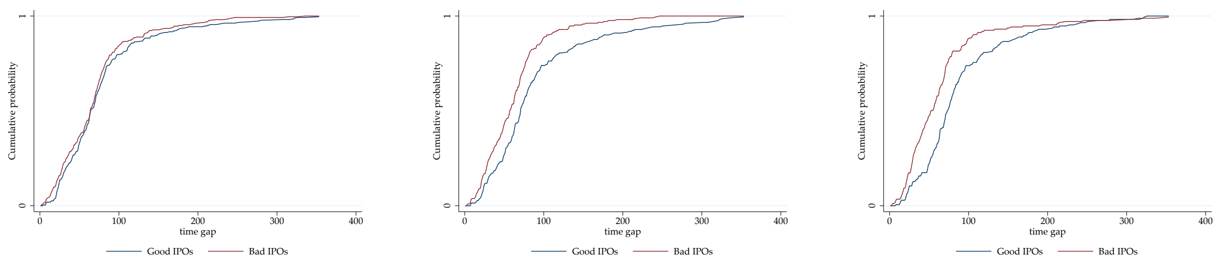


Table 1: Dardanoni and Forcina test for likelihood ratio order (p-values).

	5 years	6 years	7 year	8 years	9 years	10 years
H0 vs H1	0.679	0.351	0.002	0.000	0.000	0.000
H1 vs H2	0.811	0.973	0.957	0.997	0.899	0.877
obs.	567	528	472	422	384	345

we code as *good* (*bad*) those IPOs who performed above (below) the median IPO. We offer a more complete description of the data and our empirical strategy in Additional Material B.

Figure 5 shows the empirical distributions of *time gap* for good and bad IPOs for $y = 6, 8, 10$. All figures suggest that bad firms choose a shorter time gap in the first order stochastic dominance sense, but with the effect being more clearly visible for larger y.

We carry out a formal test of our main prediction. Corollary 1 predicts that the distribution of *time gap* for good IPOs dominates the distribution of *time gap* for bad IPOs in the likelihood ratio order. We test this prediction using an approach developed by Dardanoni and Forcina (1998). This approach tests (i) the hypothesis H_0 that the distributions are identical against the alternative H_1 that the distributions are ordered in the likelihood ratio order; as well as (ii) the hypothesis H_1 against an unrestricted alternative H_2. The hypothesis of interest H_1 is accepted if the first test rejects H_0 and the second test fails to reject H_1. Following Roosen and Hennessy (2004), we partition the variable *time gap* into k intervals that are equiprobable according to the empirical distribution of *time gap*. We report in Table 1 the p-values of the two statistics for the case of $k = 7$.

For all years, we cannot reject the hypothesis H_1 in favor of H_2 at all standard significance levels. For all years $y > 6$, we reject the hypothesis H_0 in favor of H_1 at the 1 percent significance level. However, we cannot reject H_0 at standard significance levels for $y = 5$

and $y = 6$. This pattern is consistent with our idea that that firms' private information is only gradually (and slowly) revealed to the market once the period of intense scrutiny of the IPO ends.

Our analysis may suffer from three potential pitfalls, which we address in Additional Material B. First, not all stocks are listed for 10 years. Therefore, selection, rather than performance, could explain why the effect is significant only for $y > 6$. To address this potential issue, we run the same test on the subsample of firms which were listed for at least 10 years. All our results are unchanged. Second, the variance of the long-run performance of firms going public could differ systematically across years. Therefore, our dummy for good firms could capture firms which went public earlier or later in the period, which may correlate with the choice of time gap. To address this potential issue, we run the same test with firms coded as good and bad on a yearly basis, rather than over the whole period. All our results are unchanged. Finally, we check that all our results are robust to different choices of the number of equiprobable intervals k.

6 Concluding Remarks

This paper analyzes a model in which the timing of information release is driven by the tradeoff between credibility and scrutiny. The analysis yields novel predictions about the dynamics of information release. We find supporting evidence for these predictions using available data on the timing of US presidential scandals and IPO announcements.

Our model can also be used to deliver normative implications for the design of a variety of institutions. In the context of election campaigns, our model could be employed to evaluate laws that limit the period in which candidates can announce new policies in their platforms or media can cover candidates. For example, more than a third of the world's countries mandate a *blackout period* before elections: a ban on political campaigns or, in some cases, on any mention of a candidate's name, for one or more days immediately preceding elections.[34]

Beyond election campaigns, a firm's management can give more or less time to the board to examine a plan that must be approved or rejected by a given deadline. Similarly, expert opinions and depositions from interested parties, in a court as well as in a parliament, might be given closer or farther from the date at which a decision must be taken. In all these applications, the credibility-scrutiny tradeoff may be important, and we hope our model will serve as a useful framework for studying them in the future.

[34]The 1992 US Supreme Court sentence *Burson v. Freeman*, 504 US 191, forbids such practices as violations of freedom of speech.

Appendix

A Statistical Properties

Proof of Lemma 1. Follows from Lemma 1′. □

Proof of Lemma 1′*.* *Part 1.* By Blackwell (1953), Assumption 1 with $\tau' = T + 1$ implies that pulling the arm at τ is the same as releasing an informative signal y. By Bayes's rule, posterior s is given by:

$$s = \frac{mq\left(y \mid G\right)}{mq\left(y \mid G\right) + \left(1 - m\right) q\left(y \mid B\right)}$$

where $q\left(y \mid \theta\right)$ is the density of y given θ. (If L is discrete, then $q\left(y \mid \theta\right)$ is the discrete density of y given θ.) Therefore,

$$\frac{q\left(y \mid G\right)}{q\left(y \mid B\right)} = \frac{1 - m}{m} \frac{s}{1 - s}. \tag{10}$$

Writing (10) for interim beliefs m and m', we obtain the following relation for corresponding posterior beliefs s and s':

$$\frac{1 - m'}{m'} \frac{s'}{1 - s'} = \frac{1 - m}{m} \frac{s}{1 - s},$$

which implies that

$$s' = \frac{\frac{m's}{m}}{\frac{m's}{m} + \frac{(1-m')(1-s)}{1-m}} \tag{11}$$

Therefore, $s' > s$ for $m' > m$; so part 1 follows.

Part 2. By Blackwell (1953), Assumption 1 implies that pulling the arm at τ is the same as pulling the arm at τ' and then releasing an additional informative signal y with conditional density $q\left(y \mid \theta\right)$. Part 2 holds because for any strictly increasing concave v_B, we have

$$\begin{aligned}
\mathbb{E}\left[v_B(s) \mid \tau, m, B\right] &= \mathbb{E}\left[v_B\left(\frac{sq(y \mid G)}{sq(y \mid G)+(1-s)\,q(y \mid B)}\right) \mid \tau', m, B\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[v_B\left(\frac{sq(y \mid G)}{sq(y \mid G)+(1-s)\,q(y \mid B)}\right) \mid \tau', s, B\right] \mid \tau', m, B\right] \\
&\leq \mathbb{E}\left[v_B\left(\mathbb{E}\left[\frac{sq(y \mid G)}{sq(y \mid G)+(1-s)\,q(y \mid B)} \mid \tau', s, B\right]\right) \mid \tau', m, B\right] \\
&< \mathbb{E}\left[v_B\left(\frac{s\mathbb{E}\left[\frac{q(y|G)}{q(y|B)} \mid \tau', s, B\right]}{s\mathbb{E}\left[\frac{q(y|G)}{q(y|B)} \mid \tau', s, B\right]+(1-s)}\right) \mid \tau', m, B\right] \\
&= \mathbb{E}\left[v_B(s) \mid \tau', m, B\right],
\end{aligned}$$

where the first line holds by Bayes's rule, the second by the law of iterated expectations, the third by Jensen's inequality applied to concave v_B, the fourth by strict monotonicity of v_B and Jensen's inequality applied to function $sz/\,(sz+1-s)$ which is strictly concave in z, and the last by definition of expectations.

Part 3. Analogously to Part 2, Part 3 holds because for any strictly increasing convex v_G, we have

$$\begin{aligned}
\mathbb{E}\left[v_G(s) \mid \tau, m, G\right] &= \mathbb{E}\left[v_G\left(\frac{sq(y \mid G)}{sq(y \mid G)+(1-s)\,q(y \mid B)}\right) \mid \tau', m, G\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[v_G\left(\frac{sq(y \mid G)}{sq(y \mid G)+(1-s)\,q(y \mid B)}\right) \mid \tau', s, G\right] \mid \tau', m, G\right] \\
&\geq \mathbb{E}\left[v_G\left(\mathbb{E}\left[\frac{sq(y \mid G)}{sq(y \mid G)+(1-s)\,q(y \mid B)} \mid \tau', s, G\right]\right) \mid \tau', m, G\right] \\
&> \mathbb{E}\left[v_G\left(\frac{s}{s+(1-s)\,\mathbb{E}\left[\frac{q(y|B)}{q(y|G)} \mid \tau', s, G\right]}\right) \mid \tau', m, G\right] \\
&= \mathbb{E}\left[v_G(s) \mid \tau', m, G\right].
\end{aligned}$$

□

Proof of Lemma 1′′. The proof of part 1 is the same as in Lemma 1′. As noted before, pulling the arm at τ is the same as pulling the arm at τ' and then releasing an additional informative signal y with conditional density $q(y \mid \theta)$. Let s_σ be the probability that Sender is good given that Receiver's posterior is s and Sender's signal is σ. By (11),

$$s_\sigma = \frac{\frac{\sigma s}{\pi}}{\frac{\sigma s}{m} + \frac{(1-\sigma)(1-s)}{1-m}}.$$

We have,

$$\begin{aligned}
\mathbb{E}\left[s \mid \tau, m, \sigma\right] &= \mathbb{E}\left[\frac{sq\left(y \mid G\right)}{sq\left(y \mid G\right) + \left(1-s\right)q\left(y \mid B\right)} \mid \tau', m, \sigma\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\frac{sq\left(y \mid G\right)}{sq\left(y \mid G\right) + \left(1-s\right)q\left(y \mid B\right)} \mid \tau', s, \sigma\right] \mid \tau', m, \sigma\right] \\
&= \mathbb{E}\left[\begin{array}{c} s_\sigma \mathbb{E}\left[\frac{sq(y|G)}{sq(y|G)+(1-s)q(y|B)} \mid \tau', s, G\right] + \\ +\left(1-s_\sigma\right)\mathbb{E}\left[\frac{sq(y|G)}{sq(y|G)+(1-s)q(y|B)} \mid \tau', s, B\right] \end{array} \middle| \tau', m, \sigma\right] \\
&= \mathbb{E}\left[\begin{array}{c} s_\sigma \mathbb{E}\left[\frac{sq(y|G)}{sq(y|G)+(1-s)q(y|B)} \mid \tau', s, G\right] + \\ +\left(1-s_\sigma\right)\mathbb{E}\left[\frac{sq(y|B)}{sq(y|G)+(1-s)q(y|B)} \mid \tau', s, G\right] \end{array} \middle| \tau', m, \sigma\right] \\
&= \mathbb{E}\left[s\mathbb{E}\left[\frac{s_\sigma + \left(1-s_\sigma\right)\frac{q(y|B)}{q(y|G)}}{s + \left(1-s\right)\frac{q(y|B)}{q(y|G)}} \mid \tau', s, G\right] \mid \tau', m, \sigma\right] \\
&\gtreqless \mathbb{E}\left[s \mid \tau', m, \sigma\right] \text{ whenever } s_\sigma \gtreqless s,
\end{aligned}$$

where the last line holds by Jensen's inequality applied to function $\left(s_\sigma + \left(1-s_\sigma\right)z\right) / \left(s + \left(1-s\right)z\right)$, which is strictly convex (concave) in z whenever $s_\sigma > s$ ($s_\sigma < s$). Because $\sigma_G > \pi > \sigma_B$, we have $s_{\sigma_G} > s > s_{\sigma_B}$, so parts 2 and 3 follow. □

B Benchmark Model

To facilitate our discussion in Section 3.2, we prove our results under more general assumptions than in our benchmark model. First, we assume that $v_G\left(s\right)$ is continuous, strictly increasing, and (weakly) convex, and $v_B\left(s\right)$ is continuous, strictly increasing, and (weakly) concave. Second, we assume that the arm arrives at a random time according to distributions $F_G = F$ for good Sender and F_B for bad Sender, where $F_B\left(t\right) \geq F_G\left(t\right)$ for all t.

Proof of Lemma 2. *Part 1.* Suppose, on the contrary, that $\mu(\tau) \leq \mu(\tau')$. Then

$$\begin{aligned}
\int v_B(s)\, dH_B(s|\tau, \mu(\tau)) &< \int v_B(s)\, dH_B(s|\tau', \mu(\tau)) \\
&\leq \int v_B(s)\, dH_B(s|\tau', \mu(\tau')),
\end{aligned}$$

where the first inequality holds by part 2 of Lemma $1'$ and the second by part 1 of Lemma $1'$. Therefore, bad Sender strictly prefers to pull the arm at τ' than at τ. A contradiction.

Good Sender strictly prefers to pull the arm at τ because

$$\begin{aligned}
\int v_G(s)\, dH_G(s|\tau, \mu(\tau)) &> \int v_G(s)\, dH_G(s|\tau', \mu(\tau)) \\
&> \int v_G(s)\, dH_G(s|\tau', \mu(\tau')),
\end{aligned}$$

where the first inequality holds by part 3 of Lemma $1'$ and the second by $\mu(\tau) > \mu(\tau')$ and part 1 of Lemma $1'$.

Part 2. By part 1 of this lemma applied to τ and $\tau' = T+1$, it suffices to show that $\mu(\tau), \mu(T+1) \in (0,1)$ for all τ in the support of P_B. First, by Bayes's rule, τ being in the support of P_B implies $\mu(\tau) < 1$. Second, by Bayes's rule, $F_G(T) < 1$ implies $\mu(T+1) > 0$. Third, $\mu(T+1) > 0$ implies $\mu(\tau) > 0$, otherwise τ could not be in the support of P_B because $v_B(\mu(T+1)) > v_B(0) = \mathbb{E}[v_B(s) \mid \tau, 0, B]$. Finally, $\mu(\tau) < 1$ implies $\mu(T+1) < 1$, otherwise τ could not be in the support of P_B because $v_B(1) > \mathbb{E}[v_B(s) \mid \tau, \mu(\tau), B]$. □

Proof of Lemma 3. By part 2 of Lemma 2, each t' in the support of P_B is also in the support of P_G. We show that each t' in the support of P_G is also in the support of P_B by contradiction. Suppose that there exists t' in the support of P_G but not in the support of P_B. Then, by Bayes's rule $\mu(t') = 1$; so bad Sender who receives the arm at $t \leq t'$ gets the highest possible equilibrium payoff $v_B(1)$, because she can pull the arm at time t' and get payoff $v_B(1)$ (recall that, for all t, the support of $H(.|t, \pi)$ does not contain $s = 0$ by part (ii) of Assumption (1)). Because bad Sender receives the arm at or before t' with a positive probability (recall that, for all t, $F_B(t) \geq F_G(t) > 0$ by assumption), there exists time τ at which bad Sender pulls the arm with a positive probability and gets payoff $v_B(1)$. But then $\mu(\tau) = 1$, contradicting that bad Sender pulls the arm at τ with a positive probability.

Suppose, on the contrary, that there exists τ such that $P_G(\tau) > 0$ and $P_B(\tau) \geq P_G(\tau)$. Because $P_\theta(\tau) = \sum_{t=1}^{\tau} (P_\theta(t) - P_\theta(t-1))$, there exists $\tau' \leq \tau$ in the support of P_B

such that $P_B(\tau') - P_B(\tau'-1) \geq P_G(\tau') - P_G(\tau'-1)$. Similarly, because $1 - P_\theta(\tau) = \sum_{t=\tau+1}^{T+1} (P_\theta(t) - P_\theta(t-1))$ and $1 - P_G(\tau) > 0$ (recall that $P_G(T) \leq F_G(T) < 1$), there exists $\tau'' > \tau$ in the support of P_G such that $P_G(\tau'') - P_G(\tau''-1) \geq P_B(\tau'') - P_B(\tau''-1)$. By Bayes's rule,

$$\begin{aligned} \mu(\tau') &= \frac{\pi (P_G(\tau') - P_G(\tau'-1))}{\pi (P_G(\tau') - P_G(\tau'-1)) + (1-\pi)(P_B(\tau') - P_B(\tau'-1))} \leq \pi \\ &\leq \frac{\pi (P_G(\tau'') - P_G(\tau''-1))}{\pi (P_G(\tau'') - P_G(\tau''-1)) + (1-\pi)(P_B(\tau'') - P_B(\tau''-1))} = \mu(\tau'') . \end{aligned}$$

Therefore, by Lemma 2, bad Sender strictly prefers to pull the arm at τ'' than at τ', contradicting that τ' is in the support of P_B. □

Proof of Lemma 4. By Lemma 3, P_G and P_B have the same supports and therefore $\mu(\tau) \in (0,1)$. Let the support of P_G be $\{\tau_1, ..., \tau_n\}$. Notice that $\tau_n = T+1$ because $P_G(T) \leq F_G(T) < 1$. Since τ_{n-1} is in the support of P_B and

$$P_B(\tau_{n-1}) < P_G(\tau_{n-1}) = F_G(\tau_{n-1}) \leq F_B(\tau_{n-1}),$$

where the first inequality holds by Lemma 3, the equality by part 2 of Lemma 2, and the last inequality by assumption $F_B(t) \geq F_G(t)$. Therefore, bad Sender who receives the arm at τ_{n-1} must be indifferent between pulling the arm at τ_{n-1} or at τ_n. Analogously, bad Sender who receives the arm at τ_{n-k-1} must be indifferent between pulling it at τ_{n-k-1} and at some $\tau \in \{\tau_{n-k}, \ldots, \tau_n\}$. Thus, by mathematical induction on k, bad Sender is indifferent between pulling the arm at any τ in the support of P_G and at $T+1$, which proves (2).

By Bayes's rule, for all τ in the support of P_G,

$$\frac{1-\pi}{\pi} (P_B(\tau) - P_B(\tau-1)) = \frac{1-\mu(\tau)}{\mu(\tau)} (P_G(\tau) - P_G(\tau-1)) . \tag{12}$$

Summing up over τ yields (3). □

Proof of Proposition 1. *Part 1.* We first show that, for each $\mathcal{T} \subseteq \{1, \ldots, T+1\}$ with $T+1 \in \mathcal{T}$and each $\tau \in \mathcal{T}$, there exist unique $P_G(\tau)$, $P_B(\tau)$, and $\mu(\tau)$ given by part 1 of this proposition. It suffices to show that there exists a unique $\{\mu(\tau)\}_{\tau \in \mathcal{T}} \in [0,1]^{|\mathcal{T}|}$ that solves (2) and (3). Using (11) with $m = \pi$ and $m' = \mu(\tau)$, the left hand side of (2) can be rewritten as

$$V_B(\mu(\tau)) \equiv \int v_B\left(\frac{\frac{\mu(\tau)s}{\pi}}{\frac{\mu(\tau)s}{\pi}+\frac{(1-\mu(\tau))(1-s)}{1-\pi}}\right) dH_B(s|\tau,\pi).$$

Because v_B is continuous and strictly increasing, V_B is also continuous and strictly increasing. Furthermore, $V_B(0) = v_B(0)$ and $V_B(1) = v_B(1)$. Therefore, for all $\mu(T+1) \in [0,1]$ and all $\tau \in \mathcal{T}$, there exists a unique $\mu(\tau)$ that solves (2). Moreover, for all $\tau \in \mathcal{T}$, $\mu(\tau)$ is continuous and strictly increasing in $\mu(T+1)$, is equal to 0 if $\mu(T+1) = 0$, and is equal to 1 if $\mu(T+1) = 1$. The left hand side of (3) is continuous and strictly decreasing in $\mu(\tau)$ for all $\tau \in \mathcal{T}$. Moreover, the left hand side of (3) is 0 when $\mu(\tau) = 1$ for all $\tau \in \mathcal{T}$, and it approaches infinity when $\mu(\tau)$ approaches 0 for all $\tau \in \mathcal{T}$. Therefore, substituting each $\mu(\tau)$ in (3) with a function of $\mu(T+1)$ obtained from (2), we conclude that there exists a unique $\mu(T+1)$ that solves (3).

We now construct an equilibrium for each $\mathcal{T} \subseteq \{1,\ldots,T+1\}$ with $T+1 \in \mathcal{T}$. Let $P_G(\tau)$ and $P_B(\tau)$ be given by part 1 of this proposition for all $\tau \in \{1,\ldots,T+1\}$. Let $\mu(\tau)$ be given by part 1 of this proposition for all $\tau \in \mathcal{T}$ and $\mu(\tau) = 0$ otherwise. Notice that, so constructed, P_G, P_B, and μ exist and are unique. P_G is clearly a distribution. P_B is also a distribution, because $P_B(\tau)$ increases with τ by (4) and $P_B(T+1) = 1$ by (3) and (4). Furthermore, μ is a consistent belief because (12) holds for all $\tau \in \mathcal{T}$ by (4). It remains to show that there exists an optimal strategy for Sender such that good and bad Sender's distributions of pulling time are given by P_G and P_B. First, both good and bad Sender strictly prefer not to pull the arm at any time $\tau \notin \mathcal{T}$, because, by part (ii) of Assumption (1), pulling the arm at τ gives Sender a payoff of $v_\theta(0) < v_\theta(\mu(T+1))$. Second, by (2), pulling the arm at any time $\tau \in \mathcal{T}$ gives bad Sender the same expected payoff $v_\theta(\mu(T+1))$. Last, by part 1 of Lemma 2, good Sender strictly prefers to pull the arm at time $\tau \in \mathcal{T}$ than at any other time $\tau' > \tau$.

Finally, in any equilibrium, P_G and P_B have the same supports by Lemma 3. Moreover, for all τ in the support of P_G, $P_G(\tau) = F(\tau)$ by part 2 of Lemma 2, $P_B(\tau)$ satisfies (4) by (12), $\mu(\tau) \in (0,1)$ by Lemma 3, and $\mu(\tau)$ satisfies (3) and (4) by Lemma 4.

Part 2. First, we notice that, by part 1 of Proposition 1, there exists an equilibrium with $\mathcal{T} = \{1,\ldots,T+1\}$. In this equilibrium, there are no out of equilibrium events and therefore it is divine.

Adopting Cho and Kreps (1987)'s definition to our setting (see e.g., Maskin and Tirole, 1992), we say that an equilibrium is divine if $\mu(\tau) = 1$ for any $\tau \notin supp(P_G)$ at which condition D1 holds. D1 holds at τ if for all $m \in [0,1]$ that satisfy

$$\int v_B(s)\, dH_B(s|\tau, m) \geq \max_{t \in supp(P_G), t > \tau} \int v_B(s)\, dH_B(s|t, \mu(t)) \tag{13}$$

the following inequality holds:

$$\int v_G(s)\, dH_G(s|\tau, m) > \max_{t \in supp(P_G), t > \tau} \int v_G(s)\, dH_G(s|t, \mu(t)) \,. \tag{14}$$

Suppose, on the contrary, that there exists a divine equilibrium in which $P_G(\tau) < F_G(\tau)$ for some $\tau \in \{1, \ldots, T\}$. By part 1 of Proposition 1, $\tau \notin supp(P_G)$. Let t^* denote t that maximizes the right hand side of (14). By Lemma 3, $\mu(t^*) < 1$, and, by Lemma 4, t^* maximizes the right hand side of (13). Therefore, by part 1 of Lemma 2, D1 holds at τ; so $\mu(\tau) = 1$. But then $\tau \notin supp(P_G)$ cannot hold, because

$$\int v_G(s)\, dH_G(s|\tau, 1) = v_G(1) > \max_{t \in supp(P_G)} \int v_G(s)\, dH_G(s|t, \mu(t)) \,.$$

$\square$

Proof of Corollary 2. By Lemma 4 and part 2 of Proposition 1, bad Sender is indifferent between pulling the arm at any time before the deadline and not pulling the arm at all. Then, by Lemma 2, $\mu(\tau - 1) > \mu(\tau)$ for all τ.

Using (4) with $P_G = F_G$, we have that for all $\tau < T$,

$$\begin{aligned} \frac{1 - \tilde{\mu}(\tau)}{\tilde{\mu}(\tau)} &= \frac{1-\pi}{\pi} \frac{1 - P_B(\tau)}{1 - P_G(\tau)} \\ &= \frac{\sum_{t=\tau+1}^{T+1} \frac{1-\mu(t)}{\mu(t)} \left(F_G(t) - F_G(t-1)\right)}{1 - F_G(\tau)} \\ &= \mathbb{E}_F \left[\frac{1 - \mu(t)}{\mu(t)} \mid t \geq \tau + 1 \right]. \end{aligned} \tag{15}$$

Since $\mu(\tau - 1) > \mu(\tau)$ for all τ, (15) implies that $\tilde{\mu}(\tau - 1) > \tilde{\mu}(\tau)$ and $\mu(\tau) > \tilde{\mu}(\tau - 1)$ for all τ. $\square$

Proof of Corollary 1. Using (4) with $P_G = F_G$, we have

$$\frac{1 - \mu(\tau)}{\mu(\tau)} = \frac{1-\pi}{\pi} \frac{P_B(\tau) - P_B(\tau - 1)}{P_G(\tau) - P_G(\tau - 1)}.$$

To complete the proof, notice that, by Corollary 2, $\mu(\tau) > \mu(\tau')$ whenever $\tau < \tau'$. $\square$

Proof of Lemma 2'. Given Receiver's interim belief m and pulling times τ and τ', we write H and H' for distributions $H(. \mid \tau, m)$ and $H(. \mid \tau', m)$ of Receiver's posterior belief s from Receiver's perspective, and we write H_θ and H'_θ for distributions $H_\theta(. \mid \tau, m)$ and $H_\theta(. \mid \tau', m)$ of Receiver's posterior belief s from type-θ Sender's perspective.

For any interim belief $m \in (0,1)$ and pulling times τ, τ', by Bayes's rule, we have

$$\begin{aligned} dH(s) &= m dH_G(s) + (1-m) dH_B(s), \\ s &= \frac{m dH_G(s)}{m dH_G(s) + (1-m) dH_B(s)}, \end{aligned}$$

so that $dH_G(s) = \frac{s}{m} dH(s)$ and $dH_B(s) = \frac{1-s}{1-m} dH(s)$. Likewise, $dH'_G(s) = \frac{s}{m} dH'(s)$ and $dH'_B(s) = \frac{1-s}{1-m} dH'(s)$.

For any pulling time τ and interim beliefs $m, m' \in (0,1)$, each posterior belief s under interim belief m transforms into the posterior belief s' given by (11) under interim belief m'.

Let $m = \mu(\tau)$ and $m' = \mu(\tau')$. Bad Sender weakly prefers to pull the arm at τ than at τ' if and only if

$$\int_0^1 v_B(s) \, dH_B(s) \geq \int_0^1 v_B\left(\frac{\frac{m's}{m}}{\frac{m's}{m} + \frac{(1-m')(1-s)}{1-m}}\right) dH'_B(s),$$

which is equivalent to

$$\int_0^1 v_B(s)(1-s) \, dH(s) \geq \int_0^1 v_B\left(\frac{\frac{m's}{m}}{\frac{m's}{m} + \frac{(1-m')(1-s)}{1-m}}\right)(1-s) \, dH'(s). \tag{16}$$

Similarly, good Sender strictly prefers to pull the arm at τ than at τ' if and only if

$$\int_0^1 v_G(s) s dH(s) > \int_0^1 v_G\left(\frac{\frac{m's}{m}}{\frac{m's}{m} + \frac{(1-m')(1-s)}{1-m}}\right) s dH'(s). \tag{17}$$

Because s and s' are in $(0,1)$, and s' is strictly increasing in s, for any $r \in (0,1)$, we have that $s' > r$ if and only if $s > r'$ for some $r' \in (0,1)$, which depends on m and m'.

Thus, for v_θ given by (5), the inequalities (16) and (17) can be rewritten as

$$\int_r^1 (1-s)\, dH(s) \;\geq\; \int_{r'}^1 (1-s)\, dH'(s), \tag{18}$$

$$\int_r^1 s dH(s) \;>\; \int_{r'}^1 s dH'(s), \tag{19}$$

where $\int_r^1$ and $\int_{r'}^1$ stand for the Lebesgue integrals over the sets $(r,1]$ and $(r',1]$. Notice also that we are using a selection from Receiver's best response correspondence for which $v(r)=0$. The proof goes through for other selections, after adding appropriate terms on both sides of (18) and (19). Integrating by parts, we can rewrite (18) and (19) as

$$-(1-r)\, H(r) + \int_r^1 H(s)\, ds \;\geq\; -\left(1-r'\right) H'\left(r'\right) + \int_{r'}^1 H'(s)\, ds, \tag{20}$$

$$-rH(r) - \int_r^1 H(s)\, ds \;>\; -r'H'\left(r'\right) - \int_{r'}^1 H'(s)\, ds, \tag{21}$$

Suppose that (20) and (6) hold and let us show that (21) holds. We have $H'(r') > H(r)$, because

$$\begin{aligned}
(1-r)\left(H'\left(r'\right) - H(r)\right) &= \left(1-r'\right) H'\left(r'\right) + \left(r'-r\right) H'\left(r'\right) - (1-r)\, H(r) \\
&\geq \left(1-r'\right) H'\left(r'\right) + \int_r^{r'} H'(s)\, ds - (1-r)\, H(r) \\
&\geq \int_r^1 H'(s)\, ds - \int_r^1 H(s)\, ds > 0,
\end{aligned}$$

where the equality holds by rearrangement, the first inequality holds by monotonicity of H, the second by (20), and the last by (6). The inequality (21) then holds because

$$\begin{aligned}
r'H'\left(r'\right) - rH(r) + \int_{r'}^1 H'(s)\, ds - \int_r^1 H(s)\, ds &> r'H'\left(r'\right) - rH(r) - \int_r^{r'} H'(s)\, ds \\
&\geq r'H'\left(r'\right) - rH(r) - H'\left(r'\right)\left(r'-r\right) \\
&= r\left(H'\left(r'\right) - H(r)\right) > 0,
\end{aligned}$$

where the first inequality holds by (6), the second by monotonicity of H, and the last by the established inequality $H'(r') > H(r)$. □

C Poisson Model

Proof of Proposition 6. For π: Differentiating $\mu(T)$ in Proposition 2 with respect to π, we have

$$\frac{d\mu(T)}{d\pi} = \left\{ \begin{array}{l} \frac{1}{\pi^2}\mu(T)^2 \text{ if } \pi < \bar{\pi}, \\ \frac{e^{\alpha T}}{\pi^2}\mu(T)^{2+\frac{\alpha}{\lambda}} \text{ otherwise,} \end{array} \right\} > 0.$$

For λ: First, when $\pi < \bar{\pi}$, $\frac{d\mu(T)}{d\lambda} < 0$ since $e^{-(\alpha+\lambda)T} > 1 - (\alpha+\lambda)T$ for all $\alpha, \lambda, T > 0$. Second, when $\pi > \bar{\pi}$,

$$\begin{aligned} \frac{d\mu(T)}{d\lambda} &= \frac{d}{d\lambda} e^{-\frac{\lambda}{\alpha+\lambda}\ln(1+\phi(\lambda))}, \\ \phi &\equiv \frac{\alpha+\lambda}{\lambda}\frac{1-\pi}{\pi}e^{\alpha T} > 0. \end{aligned}$$

Thus, $\frac{d\mu(T)}{d\lambda} < 0$, because

$$\frac{d}{d\lambda}\frac{\lambda}{\alpha+\lambda}\ln(1+\phi) = \frac{\alpha}{\alpha+\lambda}\left[\frac{\ln(1+\phi)}{\alpha+\lambda} - \frac{1}{\lambda}\frac{1-\pi}{\pi}\frac{1}{1+\phi}\right] > 0$$

where the inequality follows from $(1+\phi)\ln(1+\phi) > \phi$.

For α: First, when $\pi < \bar{\pi}$,

$$\begin{aligned} \frac{d\mu(T)}{d\alpha} &= -(\mu(T))^2 \frac{\chi}{(\alpha+\lambda)^2} < 0, \\ \chi &\equiv \lambda\left\{e^{\lambda T} - [1 + (\alpha+\lambda)T]e^{-\alpha T}\right\} > 0, \end{aligned}$$

where the last passage follows from $e^{(\alpha+\lambda)T} > 1 + (\alpha+\lambda)T$ for all $\alpha, \lambda, T > 0$. Second, when $\pi \geq \bar{\pi}$, by log-differentiation,

$$\frac{d\mu(T)}{d\alpha} = \mu(T)\frac{\lambda}{\alpha+\lambda}\left[\frac{\ln(1+\phi)}{\alpha+\lambda} - \frac{1}{1+\phi}\frac{d\phi}{d\alpha}\right].$$

Thus,

$$\frac{d\mu(T)}{d\alpha} < 0 \iff \frac{(1+\phi)\ln(1+\phi)}{\phi} < 1 + T(\alpha+\lambda). \tag{22}$$

For $\pi = \bar{\pi}$, $\phi = e^{(\alpha+\lambda)T} - 1 > 0$; so

$$\frac{d\mu(T)}{d\alpha} < 0 \iff \ln(1+\phi) < \phi,$$

which is true for all $\phi > 0$. Then $\frac{d\mu(T)}{d\alpha} < 0$ for $\pi \geq \bar{\pi}$ follows because $\phi < e^{(\alpha+\lambda)T} - 1$ for $\pi > \bar{\pi}$ and the left hand side of (22) increases with ϕ for $\phi > 0$. □

Proof of Proposition 3. *Part 1.* Recall that (i) Sender's payoff equals Receiver's posterior belief about Sender at $t = T$ and (ii) in equilibrium, bad Sender (weakly) prefers not to pull the arm at all than pulling it at any time $t \in [0, T]$. Therefore, bad Sender's expected payoff equals Receiver's belief about Sender at $t = T$ if the arm has not been pulled:

$$\mathbb{E}\left[v_B\right] = \mu\left(T\right). \tag{23}$$

Part 1 then follows from Proposition 6.

Part 2. By the law of iterated expectations,

$$\begin{aligned} \mathbb{E}\left[s\right] &= \pi\mathbb{E}\left[v_G\right] + \left(1-\pi\right)\mathbb{E}\left[v_B\right] = \pi \\ \Rightarrow \mathbb{E}\left[v_G\right] &= 1 - \frac{1-\pi}{\pi}\mu\left(T\right) \end{aligned} \tag{24}$$

where s is Receiver's posterior belief about Sender at $t = T$ and we use (23) in the last passage. Thus, good Sender's expected payoff increases with α and λ by Proposition 6. Finally, it is easy to see that $\mathbb{E}\left[v_G\right]$ increases in π after substituting $\mu\left(T\right)$ in $\mathbb{E}\left[v_G\right]$.

Part 3. We shall show that in the divine equilibrium

$$\mathbb{E}\left[u\right] = \frac{\left(1-\pi\right)\left(1-\mu\left(T\right)\right)}{2}. \tag{25}$$

Part 3 then follows from Proposition 6.

Since $\mathbb{E}\left[s\right] = \pi$, by (1) and (24), it is sufficient to prove that $\mathbb{E}\left[s^2\right] = \pi\mathbb{E}\left[v_G\right]$. We divide the proof in two cases: $\pi \leq \bar{\pi}$ and $\pi > \bar{\pi}$. If $\pi \leq \bar{\pi}$, Receiver's expected payoff is given by the sum of four terms: (i) Sender is good and the arm does not arrive; (ii) Sender is good and the arm arrives; (iii) Sender is bad and she does not pull the arm; and (iv) Sender is bad and she pulls the arm. Thus,

$$\begin{aligned} \mathbb{E}\left[s^2\right] &= \pi e^{-\alpha T}\left(\mu\left(T\right)\right)^2 \\ &\quad + \pi\int_0^T \left(e^{\lambda(T-t)}\mu\left(T\right)\right)^2 \alpha e^{-\alpha t} dt \\ &\quad + \left(1-\pi\right)\left(1-P_B\left(T\right)\right)\left(\mu\left(T\right)\right)^2 \\ &\quad + \left(1-\pi\right)\int_0^T e^{-\lambda(T-t)}\left(e^{\lambda(T-t)}\mu\left(T\right)\right)^2 \frac{\pi}{1-\pi}\left(\frac{1-\mu\left(t\right)}{\mu\left(t\right)}\right)\alpha e^{-\alpha t} dt. \end{aligned}$$

Solving all integrals and rearranging all common terms we get

$$\mathbb{E}\left[s^2\right] = \pi \mathbb{E}\left[v_G\right].$$

If $\pi > \bar{\pi}$, Receiver's expected payoff is given by the sum of five terms: (i) Sender is good and the arm does not arrive; (ii) Sender is good and the arm arrives before $\bar{t}$; (iii) Sender is good and the arm arrives between $\bar{t}$ and T; (iv) Sender is bad and she does not pull the arm; (v) Sender is bad and she pulls the arm. Thus,

$$\begin{aligned}
\mathbb{E}\left[s^2\right] &= \pi e^{-\alpha T}\left(\mu\left(T\right)\right)^2 \\
&\quad +\pi\left(1-e^{-\alpha\bar{t}}\right) \\
&\quad +\pi\int_{\bar{t}}^{T}\left(e^{\lambda(T-t)}\mu\left(T\right)\right)^2 \alpha e^{-\alpha t}dt + \\
&\quad +\left(1-\pi\right)\left(1-P_B\left(T\right)\right)\left(\mu\left(T\right)\right)^2 \\
&\quad +\left(1-\pi\right)\int_{\bar{t}}^{T} e^{-\lambda(T-t)}\left(e^{\lambda(T-t)}\mu\left(T\right)\right)^2 \frac{\pi}{1-\pi}\left(\frac{1-\mu\left(t\right)}{\mu\left(t\right)}\right)\alpha e^{-\alpha t}dt.
\end{aligned}$$

Solving all integrals and rearranging all common terms we again get

$$\mathbb{E}\left[s^2\right] = \pi \mathbb{E}\left[v_G\right].$$

□

Proof of Proposition 4. For π: Differentiating $P_B\left(T\right)$ in (7) with respect to π, we have

$$\begin{aligned}
\frac{dP_B\left(T\right)}{d\pi} &= \frac{e^{-\alpha T}}{\mu\left(T\right)\left(1-\pi\right)} \times \left[\frac{\pi}{\mu\left(T\right)}\frac{d\mu\left(T\right)}{d\pi} - \frac{1-\mu\left(T\right)}{1-\pi}\right] \\
&= \frac{e^{-\alpha T}}{\mu\left(T\right)\left(1-\pi\right)} \times \left\{ \begin{array}{l} \left[\frac{\mu(T)}{\pi} - \frac{1-\mu(T)}{1-\pi}\right] \text{ if } \pi < \bar{\pi}, \\ \left[e^{\alpha T}\frac{\mu(T)^{1+\frac{\alpha}{\lambda}}}{\pi} - \frac{1-\mu(T)}{1-\pi}\right] \text{ otherwise.} \end{array} \right\}
\end{aligned}$$

First, when $\pi < \bar{\pi}$, $\frac{dP_B(T)}{d\pi} < 0$ because $\mu\left(T\right) < \pi$. Second, when$\pi \geq \bar{\pi}$, $\frac{dP_B(T)}{d\pi} < 0$ if and only if

$$1 + \frac{\alpha}{\alpha+\lambda}\phi > \left(1+\phi\right)^{\frac{\alpha}{\alpha+\lambda}},$$

$$\phi \equiv \frac{\alpha+\lambda}{\lambda}\frac{1-\pi}{\pi}e^{\alpha T} > 0.$$

Thus, $\frac{dP_B(T)}{d\pi} < 0$, because $1 + x\phi > (1+\phi)^x$ for all $\phi > 0$ and $x \in (0,1)$.

For λ: Differentiating $P_B(T)$ in (7) with respect to λ, we have

$$\frac{dP_B(T)}{d\lambda} = \frac{\pi}{1-\pi}\frac{e^{-\alpha T}}{\mu(T)^2}\frac{d\mu(T)}{d\lambda} < 0,$$

where the inequality follows from Proposition 6.

For α: Without loss of generality we can set $T = 1$. Differentiating $P_B(T)$ in (7) with respect to α, we have

$$\frac{dP_B(T)}{d\alpha} = \frac{\pi}{1-\pi}e^{-\alpha}\left[\frac{1-\mu(T)}{\mu(T)} + \frac{1}{(\mu(T))^2}\frac{d\mu(T)}{d\alpha}\right].$$

First, when $\pi < \bar{\pi}$,

$$\begin{aligned}
\frac{1-\pi}{\pi}e^{2\alpha}\frac{dP_B(T)}{d\alpha} &= \left(\frac{1}{\pi}-2\right)e^{\alpha} + \frac{(\alpha(\alpha+\lambda)-\lambda)e^{\alpha+\lambda} + \lambda(1+2(\alpha+\lambda))}{(\alpha+\lambda)^2} \\
&> \left(\frac{1}{\bar{\pi}}-2\right) + \frac{(\alpha(\alpha+\lambda)-\lambda)e^{\alpha+\lambda} + \lambda(1+2(\alpha+\lambda))}{(\alpha+\lambda)^2} \\
&= \frac{1}{(\alpha+\lambda)^2}\left(\lambda(1+(\alpha+\lambda)) + \left((\alpha+\lambda)^2-\lambda\right)e^{(\alpha+\lambda)} - (\alpha+\lambda)^2 e^{\alpha}\right) \\
&= \sum_{k=3}^{\infty}\left[\frac{(\alpha+\lambda)^k}{(k-2)!} - \lambda\frac{(\alpha+\lambda)^{k-1}}{(k-1)!} - (\alpha+\lambda)^2\frac{\alpha^{k-2}}{(k-2)!}\right] \equiv \sum_{k=3}^{\infty} c_k > 0,
\end{aligned}$$

where the inequality holds because each term c_k in the sum is positive:

$$\begin{aligned}
c_k &= \frac{(\alpha+\lambda)^2\left((\alpha+\lambda)^{k-2}-\alpha^{k-2}\right)}{(k-2)!} - \frac{(\alpha+\lambda)^2\lambda(\alpha+\lambda)^{k-3}}{(k-1)!} \\
&= \frac{(\alpha+\lambda)^2\lambda\left(\sum_{n=0}^{k-3}(\alpha+\lambda)^{k-3-n}\alpha^n\right)}{(k-2)!} - \frac{(\alpha+\lambda)^2\lambda(\alpha+\lambda)^{k-3}}{(k-1)!} \\
&> \frac{(\alpha+\lambda)^2\lambda(\alpha+\lambda)^{k-3}}{(k-2)!} - \frac{(\alpha+\lambda)^2\lambda(\alpha+\lambda)^{k-3}}{(k-1)!} > 0.
\end{aligned}$$

Second, when $\pi \geq \bar{\pi}$, $\frac{dP_B(T)}{d\alpha} > 0$ if and only if

$$\frac{1+\phi}{\phi}\left[\ln(1+\phi) + \frac{(\alpha+\lambda)^2}{\lambda}\left(1-(1+\phi)^{-\frac{\alpha+\lambda}{\lambda}}\right)\right] - 1 - \alpha - \lambda > 0$$

$$\phi \equiv \frac{\alpha+\lambda}{\lambda}\frac{1-\pi}{\pi}e^{\alpha T}.$$

The left hand side increases with α, treating ϕ as a constant. Then the inequality holds because it holds for $\alpha \to 0$:

$$\begin{aligned}
\frac{1+\phi}{\phi}\left[\ln(1+\phi)+\lambda\left(1-(1+\phi)^{-1}\right)\right]-1-\lambda &> 0 \\
\iff \frac{1+\phi}{\phi}\ln(1+\phi) &> 1.
\end{aligned}$$

$\square$

Proof of Proposition 5. For λ: Differentiating $P(T)$ in (9) with respect to λ, we have

$$\frac{dP(T)}{d\lambda} = (1-\pi)\frac{dP_B(T)}{d\lambda} < 0,$$

where the inequality follows from Proposition 4.

For α: Differentiating $P(T)$ in (9) with respect to α, we have

$$\frac{dP(T)}{d\alpha} > (1-\pi)\frac{dP_B(T)}{d\alpha} > 0,$$

where the last inequality follows from Proposition 4.

For π: Differentiating $P(T)$ in (9) with respect to π, we have

$$\frac{dP(T)}{d\pi} = \frac{\pi e^{-\alpha T}}{\mu(T)^2}\left(\frac{d\mu(T)}{d\pi}-\frac{\mu(T)}{\pi}\right).$$

We now show that

$$\frac{dP(T)}{d\pi} \geq 0 \iff \pi \geq \frac{\alpha e^{\alpha T}}{(\alpha+\lambda)e^{\alpha T}-1}.$$

First, when $\underline{\pi < \bar{\pi}}$, we have $dP(T)/d\pi < 0$ because $\mu(T) < \pi$ and

$$\frac{d\mu(T)}{d\pi} = \frac{\mu(T)^2}{\pi^2} < \frac{\mu(T)}{\pi}.$$

Second, when $\underline{\pi \geq \bar{\pi}}$, we have $dP(T)/d\pi < 0$ if and only if

$$\frac{d\mu(T)}{d\pi} = e^{\alpha T}\frac{\mu(T)^{2+\frac{\alpha}{\lambda}}}{\pi^2} < \frac{\mu(T)}{\pi}.$$

Substituting $\mu(T)$, we get that this inequality is equivalent to

$$\pi < \frac{\alpha e^{\alpha T}}{\alpha e^{\alpha T} + \lambda\left(e^{\alpha T} - 1\right)}.$$

It remains to show that

$$\frac{\alpha e^{\alpha T}}{\alpha e^{\alpha T} + \lambda\left(e^{\alpha T} - 1\right)} > \bar{\pi}.$$

Substituting $\bar{\pi}$, we get that this inequality is equivalent to

$$\frac{e^{(\alpha+\lambda)T} - 1}{\alpha + \lambda} > \frac{e^{\alpha T} - 1}{\alpha},$$

which is satisfied because function $\left(e^{x} - 1\right)/x$ increases with x. □

Proof of Proposition 7. First, for $\pi < \bar{\pi}$, $\bar{t} = 0$. Second, for $\pi \geq \bar{\pi}$, $\bar{t}$ increases with π and decreases with α because $\mu(T)$ increases with π and decreasing with α. Furthermore,

$$\begin{aligned} \frac{d\bar{t}}{d\lambda} &= \frac{1}{\alpha+\lambda}\left(\frac{1}{\alpha+\lambda}\ln(1+\phi) + \frac{\alpha}{\lambda^{2}}\frac{1-\pi}{\pi}e^{\alpha T}\frac{1}{1+\phi}\right) > 0, \\ \phi &\equiv \frac{\alpha+\lambda}{\lambda}\frac{1-\pi}{\pi}e^{\alpha T}. \end{aligned}$$

□

Proof of Proposition 8. The density $p_B(t)$ is equal to 0 for $t \leq \bar{t}$ and is given by

$$p_B(t) \equiv \frac{dP_B(t)}{dt} = \frac{\pi}{1-\pi}\frac{\alpha e^{-\alpha t}\left(1 - \mu(T)\, e^{\lambda(T-t)}\right)}{\mu(T)}$$

for $t > \bar{t}$. Differentiating $p_B(t)$ with respect to t for $t > \bar{t}$, we get

$$\frac{dp_B(t)}{dt} = \frac{\pi}{1-\pi}\frac{\alpha e^{-\alpha t}}{\mu(T)}\left[(\alpha+\lambda)\,\mu(T)\, e^{\lambda(T-t)} - \alpha\right] > 0$$

if and only if

$$t < T - \frac{1}{\lambda}\ln\left(\frac{\alpha}{\alpha+\lambda}\frac{1}{\mu(T)}\right).$$

We can therefore conclude that $p_B(t)$ is quasiconcave on the interval $[\bar{t}, T]$. □

Proof of Proposition 9. The density $p(t)$ is given by

$$p(t) = \begin{cases} \pi\alpha e^{-\alpha t} & \text{if } t < \bar{t} \\ \pi\alpha e^{-\alpha t} + \pi\alpha e^{-\alpha t} \frac{1-\mu(T)e^{\lambda(T-t)}}{\mu(T)} & \text{if } t \geq \bar{t}. \end{cases}$$

Obviously, for $t \leq \bar{t}$, $p(t)$ is decreasing in t. For $t > \bar{t}$, differentiating $p(t)$ with respect to t, we get

$$\frac{dp(t)}{dt} = \pi\alpha e^{-\alpha t}\left[(\alpha+\lambda)e^{\lambda(T-t)} - \alpha\frac{1+\mu(T)}{\mu(T)}\right] > 0$$

if and only if

$$t < T - \frac{1}{\lambda}\ln\left(\frac{\alpha}{\alpha+\lambda}\frac{1+\mu(T)}{\mu(T)}\right).$$

□

Proof of Proposition 10. The breakdown probability at t is given by

$$Q(t) \equiv \left(1 - e^{-\lambda(T-t)}\right)[1-\mu(t)].$$

Notice that $Q(t)$ is continuous in t because $\mu(t)$ is continuous in t. Also, $Q(t)$ equals 0 for $t \leq \bar{t}$, is strictly positive for all $t \in (\bar{t}, T)$, and equals 0 for $t = T$. Substituting $\mu(t)$ and differentiating $Q(t)$ with respect to t for $t \geq \bar{t}$, we get

$$\frac{dQ(t)}{dt} = -\lambda\frac{e^{-\lambda(T-t)}(1+\mu(T)) - 2\mu(T)}{\left[1-\mu(T)\left(1-e^{\lambda(T-t)}\right)\right]^2} > 0$$

if and only if

$$t < T - \frac{1}{\lambda}\ln\left(\frac{1+\mu(T)}{2\mu(T)}\right).$$

□

References

Acharya, Viral V, Peter DeMarzo, and Ilan Kremer, "Endogenous Information Flows and the Clustering of Announcements," *American Economic Review*, 2011, *101* (7), 2955–79.

Banks, Jeffrey S. and Joel Sobel, "Equilibrium Selection in Signaling Games," *Econometrica*, 1987, *55* (3), 647–661.

Bizzotto, Jacopo, Jesper Rüdiger, and Adrien Vigier, "Delayed Persuasion," *University of Oxford Working Paper*, 2016.

Blackwell, David, "Equivalent Comparisons of Experiments," *Annals of Mathematical Statistics*, 1953, *24* (2), 262–272.

Bolton, Patrick and Christopher Harris, "Strategic Experimentation," *Econometrica*, 1999, *67* (2), 349–374.

Brocas, Isabelle and Juan D Carrillo, "Influence through Ignorance," *RAND Journal of Economics*, 2007, *38* (4), 931–947.

Che, Yeon-Koo and Johannes Hörner, "Optimal Design for Social Learning," *Yale University Working Paper*, 2015.

Cho, In-Koo and David M Kreps, "Signaling Games and Stable Equilibria," *Quarterly Journal of Economics*, 1987, *102* (2), 179–221.

Dardanoni, Valentino and Antonio Forcina, "A unified approach to likelihood inference on stochastic orderings in a nonparametric context," *Journal of the American Statistical Association*, 1998, *93* (443), 1112–1123.

DellaVigna, Stefano and Ethan Kaplan, "The Fox News Effect: Media Bias and Voting," *Quarterly Journal of Economics*, 2007, *122* (3), 1187–1234.

Duggan, John and Cesar Martinelli, "A Spatial Theory of Media Slant and Voter Choice," *Review of Economic Studies*, 2011, *78* (2), 640–666.

Durante, Ruben and Ekaterina Zhuravskaya, "Attack when the World is Not Watching?: International Media and the Israeli-Palestinian Conflict," *Journal of Political Economy*, 2016, *forthcoming.*

Dye, Ronald A, "Disclosure of Nonproprietary Information," *Journal of Accounting Research*, 1985, *23* (1), 123–145.

Dziuda, Wioletta, "Strategic Argumentation," *Journal of Economic Theory*, 2011, *146* (4), 1362–1397.

Ely, Jeffrey, Alexander Frankel, and Emir Kamenica, "Suspense and Surprise," *Journal of Political Economy*, 2015, *123* (1), 215–260.

Ely, Jeffrey C, "Beeps," *American Economic Review*, 2016, *107* (1), 31–53.

Gentzkow, Matthew and Jesse M Shapiro, "Media Bias and Reputation," *Journal of Political Economy*, 2006, *114* (2), 280–316.

Grenadier, Steven R, Andrey Malenko, and Nadya Malenko, "Timing Decisions in Organizations: Communication and Authority in a Dynamic Environment," *American Economic Review*, 2016, *106* (9), 2552–2581.

Grossman, Sanford J., "The Informational Role of Warranties and Private Disclosures about Product Quality," *Journal of Law and Economics*, 1981, *24*, 461–483.

_ **and Oliver D. Hart**, "Takeover Bids, the Free-Rider Problem, and the Theory of the Corporation," *Bell Journal of Economics*, 1980, *11* (1), 42–64.

Guttman, Ilan, Ilan Kremer, and Andrej Skrzypacz, "Not Only What but also When: A Theory of Dynamic Voluntary Disclosure," *American Economic Review*, 2013, *104* (8), 2400–2420.

Horner, Johannes and Andrzej Skrzypacz, "Selling Information," *Journal of Political Economy*, 2016, *forthcoming.*

Jovanovic, Boyan, "Truthful Disclosure of Information," *Bell Journal of Economics*, 1982, *13* (1), 36–44.

Jung, Woon-Oh and Young K Kwon, "Disclosure when the Market is Unsure of Information Endowment of Managers," *Journal of Accounting Research*, 1988, *26* (1), 146–153.

Kamenica, Emir and Matthew Gentzkow, "Bayesian Persuasion," *American Economic Review*, 2011, *101*, 2590–2615.

Keller, Godfrey, Sven Rady, and Martin Cripps, "Strategic Experimentation with Exponential Bandits," *Econometrica*, 2005, *73* (1), 39–68.

Maskin, Eric S and Jean Tirole, "The Principal-Agent Relationship with an Informed Principal, II: Common Values," *Econometrica*, 1992, *60* (1), 1–42.

Milgrom, Paul R., "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 1981, *12*, 350–391.

Mullainathan, Sendhil and Andrei Shleifer, "The market for news," *American Economic Review*, 2005, *95* (4), 1031–1053.

Nyhan, Brendan, "Scandal Potential: How Political Context and News Congestion Affect the President's Vulnerability to Media Scandal," *British Journal of Political Science*, 2015, *45* (2), 435–466.

Orlov, Dmitry, Andrzej Skrzypacz, and Pavel Zryumov, "Persuading the Regulator To Wait," *Stanford University Working Paper*, 2016.

Prat, Andrea and David Stromberg, "The Political Economy of Mass Media," in "Advances in Economics and Econometrics: Volume 2, Applied Economics: Tenth World Congress," Vol. 50 Cambridge University Press 2013, p. 135.

Rayo, Luis and Ilya Segal, "Optimal Information Disclosure," *Journal of Political Economy*, 2010, *118* (5), 949–987.

Ritter, Jay R., "Investment Banking and Securities Issuance," in George M. Constantinides, Milton Harris, and Rene M. Stulz, eds., *Handbook of the Economics of Finance*, Vol. 1A: Corporate Finance, Elsevier, 2003, chapter 5, pp. 255–306.

Roosen, Jutta and David A Hennessy, "Testing for the monotone likelihood ratio assumption," *Journal of Business & Economic Statistics*, 2004, *22* (3), 358–366.

Shaked, Moshe and George Shanthikumar, *Stochastic Orders*, New York, NY: Springer, 2007.

Shavell, Steven, "Acquisition and Discolsure of Information Prior to Sale," *RAND Journal of Economics*, 1994, *25* (1), 20–36.

Shin, Hyun-Song, "News Management and the Value of Firms," *RAND Journal of Economics*, 1994, *25* (1), 58–71.

_ , "Disclosure and Asset Returns," *Econometrica*, 2003, *71* (1), 105–133.

_ , "Disclosure Risk and Price Drift," *Journal of Accounting Research*, 2006, *44*, 351–379.