Panics and Early Warnings*

Deepal Basak[†] Zhen Zhou[‡]

Sep 2022

Abstract

This paper studies dynamic persuasion under adversarial selection in regime change games where the agents choose when to attack, and delay is costly. We construct a simple optimal disclosure policy called "timely disaster alert," which promptly warns agents about the impending crisis. This policy induces perfect coordination on the right course of action — the unique rationalizable strategy is to wait for and follow the alert, regardless of private signals. Thus, the optimal policy eliminates panic in the robust sense. We relate this optimal policy to practical early warning systems, such as forward-looking bank stress tests and debt sustainability analysis.

JEL Classification Numbers: *D02, D82, D83, G28*

Keywords: *Dynamic Persuasion, Dynamic Coordination, Adversarial Selection, First-mover Advantage, Regime Change, Early Warning, Panic*

*This paper was previously circulated under the title "Timely Persuasion." We thank Emiliano Catonini, Christophe Chamley, Joyee Deb, Douglas Gale, Ju Hu, Nicolas Inostroza, Aaron Kolb, Stephen Morris, Alessandro Pavan, David Pearce, Antonio Penta, Edouard Schaal, Martin Szydlowski, Ennio Stacchetti, Colin Stewart, Debraj Ray, Laura Veldkamp, Xingye Wu, and Ming Yang, as well as the seminar participants at the Econometric Society meetings, Stony Brook, Peking University, Tsinghua University, Fudan, NYU Shanghai, NUS, UPF, SMU, University of Washington, Toronto, Kansas, Indiana Kelley, Rochester Simon, UIUC, Utah Eccles, Ashoka, Michigan State, and SET (online) for their helpful comments and suggestions. We thank Samyak Jain and Dandan Zhao for their excellent research assistance, and Tsinghua University Spring Breeze Fund for financial support.

[†]Kelley School of Business, Indiana University. Email: dbasak@iu.edu;

[‡]PBC School of Finance, Tsinghua University. Email: zhouzh@pbcsf.tsinghua.edu.cn

Introduction

Economic agents often fail to coordinate on the right course of action, leading to undesirable outcomes. Consider a bank run, for example, where creditors may withdraw if they believe others will do the same. Accordingly, a bank may fail even when it is solvent. We refer to such events as *panic*. Economists have long been interested in understanding how a planner can design policies to avoid such undesirable outcomes. The early literature proposes intervention through a "big push." One can imagine that such interventions through payoff incentives can be exceedingly costly. More recently, Sakovics and Steiner (2012) designs the optimal payoff incentives when the planner has limited resources. This paper designs the optimal dynamic information disclosure policy to persuade agents to coordinate on the right course of action.

The existing literature studies a static persuasion problem under different coordination environments (see, for instance, Goldstein and Huang (2016), Inostroza and Pavan (2020), Li, Song and Zhao (2019), and Morris, Oyama and Takahashi (2020)). Unlike in standard information design problems, this literature adopts the approach of an adversarial information design under which the principal expects the agents to play the worst possible equilibrium. The literature has established a uniform result: the optimal disclosure policy should perfectly coordinate agents' actions (conformity of action), thereby removing all strategic uncertainty while retaining the fundamental uncertainty (non-conformity of belief). However, the literature also shows that, under adversarial selection, even the optimal information structure cannot persuade agents to coordinate perfectly on the desirable action. For instance, in our bank run example, under the optimal static disclosure, in some states where the bank is solvent, all the creditors coordinate but do so on the undesirable action and run. In other words, panic cannot be averted under the static adversarial information design.

In this paper, we augment the canonical regime change game with a timing dimension and show that dynamic persuasion — namely, allowing the disclosure policy to be scheduled at a future date and to be dependent on the past actions taken by the agents — averts panic under adversarial information design. We construct a simple dynamic information disclosure policy that not only perfectly coordinates agents' actions but also ensures that they coordinate on the right course of action. Therefore, dynamic persuasion expands the set of outcomes that the principal can achieve through static persuasion under adversarial selection. More importantly, the optimal policy completely eliminates undesirable coordination outcomes.

The timing dimension we introduce is natural for many applications. For instance, consider the following classic examples.

Example 1 (sovereign debt default) *A country has a significant amount of outstanding sovereign debt that matures in, say, six months, and the country may default. We interpret this example as a future shock that arrives at a fixed maturity date. Foreign investors who are wary about the forthcoming crisis and resulting currency depreciation may relocate their investments before the maturity date. These decisions* *worsen the country's fiscal condition, which contributes to its inability to repay its debt when the debt matures.*

Example 2 (bank run) *Consider creditors who anticipate that the bank will be hit by an adverse shock in the future (for instance, borrowers in a certain industry will start defaulting on their loans). The creditors may run on the bank before the shock actually arrives, which weakens the bank and contributes to its failure. The only difference compared with the first example is that the date on which the shock arrives and the bank becomes vulnerable is stochastic rather than fixed.*

Motivated by such examples, we introduce a finite time window $[0, T]$. A shock arrives at a future date (possibly stochastic) within this window, and once it arrives, it may make the regime vulnerable. The agents can wait at a cost (standard discounting) or attack at any time. If they attack, they get their payoff instantaneously depending on whether or not the regime has changed and the regime change depends on the aggregate attack until then. If the agents attack before the regime changes, they get significantly higher payoff than if they attack afterwards. In this sense, our dynamic setup features a *first-mover advantage*.

Introducing a time dimension to the game can change the dynamic incentives of the agents in many different ways.[1] Readers may wonder whether this gain in the power of persuasion stems from the principal's ability to disclose information over time or from the dynamic game being different from the static game. In this paper, we focus on the first source. We only consider a dynamic regime change game, where in the absence of dynamic disclosure, the outcome will be exactly as in the canonical static regime-change game. However, as the above examples show, there can be a *time gap* between when the agents start attacking and when the regime changes. This gap provides an opportunity for dynamic information disclosure.[2]

We call our proposed policy a *timely disaster alert*. A disaster alert at a given date τ gets triggered if, based on the fundamental of the regime as well as on what agents do before the disclosure, it becomes evident that the regime will change (if it has not already). However, if the alert is not triggered, it means that the regime will survive provided that no further attack occurs. We call this alert "timely" when $\tau \in (0, \hat{\tau})$ for some small but positive $\hat{\tau}$.

The disclosure policy is simple. However, it is important to note that (1) it is an *endogenous disclosure*: if agents attack without waiting for the alert, their actions may trigger the alert; and (2) it *forecasts* an impending regime change: the alert gets triggered even though the regime has not changed yet but is in no condition to survive when the shock arrives.

As is standard in the literature, we assume that before the game begins, the agents receive some noisy

[1] See Dasgupta, Steiner and Stewart (2012) for some interesting examples.

[2] In the above examples, the gap arises (at least, with high probability) because the shock arrives in the future, and the regime becomes vulnerable only after the shock arrives. However, this gap could also arise for many other reasons. For instance, it may take time to move physical capital. Also, for some applications, the payoff from attacking is not collected instantaneously but, in the end, is dependent on the ultimate regime outcome, thereby removing the first-mover advantage. See the discussions in Section 3.1 for details.

private signals about the fundamental. The principal does not observe these signals. We consider a fairly flexible exogenous information environment under which the private signals need not be conditionally independent but can be arbitrarily correlated. We assume only that regardless of their signals and others' actions, an agent will always have some *doubt* about whether his action is the right one.

We show that, under a timely disaster alert, the unique rationalizable strategy for any agent of any type (private signal) is to wait for and then follow the disaster alert.[3] Therefore, when the fundamental does not warrant a run, because all the agents wait for the alert, the alert does not get triggered, which implies that the agents do not attack afterward. Thus, panic is eliminated under our proposed policy in a robust sense.

We prove this result using three steps. First, consider an agent who plays a strategy that involves waiting for the disaster alert. We show that, regardless of others' strategy, this agent must follow the disaster alert — that is, attack if and only if the alert is triggered. If the alert is triggered, then the regime must change, and so, he must plan to attack when he learns this. If he also plans to attack when the alert is not triggered, then this means that he chooses to wait for the alert and then attack, regardless of whether or not the alert is triggered. As such, waiting has no option value. Since waiting is costly, such strategies are strictly dominated by attacking right away and not waiting for the alert. Thus, the only rationalizable strategies are attacking right away or waiting and then following the alert. We refer to this argument as the *option value* argument (see Lemma 1).[4]

Second, the disaster alert is designed such that when the alert is not triggered, regardless of how many agents choose to attack (instead of waiting), as long as the agents who choose to wait will not attack, the regime survives. Therefore, following the option value argument, if the alert is not triggered, no further attack will occur, and the regime never changes. This demonstrates the *perfect predictability* of the disaster alert (see Lemma 2).[5]

Third, it follows from the perfect predictability of the alert that there is no strategic uncertainty after the time of disclosure; that is, the action of the agents who wait for the disclosure will be perfectly coordinated. In this sense, their private signals become irrelevant. An agent can benefit from waiting for the alert because he avoids the mistake of attacking a regime that ultimately survives. This benefit is realized when the alert is not triggered ex post. Since an agent always has doubt, regardless of whether or not others wait, the benefit of waiting is positive and independent of the disclosure time τ.

However, this positive benefit does not mean that the agents will wait for the alert. Waiting for the alert may also result in attacking — but at a later date when the alert is triggered. A short wait is not very costly if the regime does not change in the meantime. However, given the first-mover advantage, if the regime

[3] A strategy for an agent of any type is a complete contingency plan. Formally, under a disaster alert, a strategy is when to attack at three possible histories: the initial history, the one after the alert is not triggered, and the one after the alert is triggered.

[4] It follows from the doubt assumption that agents always assign a positive probability both to the event that the alert gets triggered and to the event that the alert does not get triggered. This assumption makes the dominance strict. In Section 4.1, we discuss the essential features that drive Lemma 1.

[5] In Section 4.2 and 4.3, we discuss the essential features that drive Lemma 2.

changes while the agents wait for the alert, then the agent may miss the opportunity to attack a regime before it changes, which could result in a significant loss. Nonetheless, by setting the disclosure time τ close to 0, the principal ensures that the probability that an agent will bear such a cost is negligibly small. Therefore, there exists $\hat{\tau} > 0$ such that under $\tau \in (0, \hat{\tau})$, regardless of his signal and regardless of whether or not others will attack before the disclosure, the benefit outweighs the cost of waiting. This step shows the importance of providing information in a timely manner (see Lemma 3).

From the three above-mentioned lemmas, it follows that when the principal sets a timely disaster alert, under the unique rationalizable strategy, the agents ignore their private signals and wait for and then follow the alert. Thus, even though the agents receive different signals and have different beliefs about the fundamental, this simple policy persuades the agents to perfectly coordinate on taking the desirable action, which eliminates panic in the robust sense. It is worth noting that our iterated elimination arguments only require that the agents are rational and believe others are rational, whereas common knowledge of rationality is not needed.

The proposed optimal policy — timely disaster alert — resembles an early warning system (EWS). An EWS is common in practice. For instance, the International Monetary Fund (IMF) and the World Bank regularly conduct forward-looking debt sustainability analysis (DSA), which aims to provide the market with an early warning of sovereign debt distress. After the Great Recession, bank supervisors adopted forward-looking bank stress tests to assess banks' performance under some predicted future shock, and they publicly disclosed whether or not a bank would be able to sustain that adverse shock. However, it is important to note a crucial feature of our proposed disclosure policy: it not only depends on the fundamentals but also takes all historical market responses (up to the disclosure time) into account. This feature can be missing in many EWS in practice. We show in Section 4.2 that without this feature, the disclosure policy may fail to perfectly coordinate the agents' actions after the disclosure, and accordingly, the agents may not wait for such warnings.[6]

It is also important to note that the effectiveness of our policy hinges on the scheduled date of the disclosure being timely. This makes it unlikely that the regime changes while the agents wait for future disclosure. When the agents can collect their payoff instantaneously by attacking, and early movers get the advantage, they will not wait for the alert if they believe the regime is likely to change while they wait. It is the stochastic arrival of the shock that causes the cost of waiting to decrease continuously to zero as the disclosure date moves closer to time zero. When the payoffs from attacking are not realized instantaneously or when there is no first-mover advantage, the result will be stronger: the agents will wait and follow the alert even when the shock always arrives before the disclosure date (see Section 3.1). We

[6]In practice, an EWS serves many other purposes and is often designed in conjunction with other supplementary policies. See Goldstein and Leitner (2018) for how such disclosure policies can facilitate banks of heterogeneous quality to raise funds under adverse selection. See Orlov, Zryumov and Skrzypacz (2018) for how it helps to discover systemic risk and how it can be combined with capital regulations. See Inostroza (2019) for a comprehensive disclosure policy including banks' liquidity position and asset quality.

also show that the result is robust when we allow for the observability of regime change, the infrequent arrival of new information over time, and the possibility that only some (but not all) agents have doubts (see Section 3.2, 3.3 and 3.4).

Theoretically, this paper contributes to two strands of the literature: (1) dynamic coordination and (2) Bayesian persuasion or information design. See Angeletos and Lian (2017) for a recent survey on dynamic coordination. This literature studies how coordination can be affected when agents learn some information about the past (see, for instance, Chamley (1999, 2003), Angeletos, Hellwig and Pavan (2007), Dasgupta (2007), Chassang (2010)). Motivated by our examples, we build a dynamic regime change game in which a shock arrives in the future and a mass of privately informed agents decide when to attack. Unlike the above-mentioned papers, we study how a principal can manipulate the agents' beliefs by optimally choosing when and how the agents learn about the past attacks and the underlying state. The crucial feature of the game that we exploit is that the agents can choose to wait at a cost and attack at a later date. If, instead, agents move sequentially in an exogenous order, then our result does not hold. In this case, to dissuade an agent from attacking, the principal needs to assure him that other agents moving later in the exogenous order will also be dissuaded from attacking. Basak and Zhou (2020) consider such a setup and show how frequent disclosures can help.[7]

There is a large and growing literature on Bayesian persuasion or information design, which started with Kamenica and Gentzkow (2011). See Kamenica (2019) and Bergemann and Morris (2019) for recent surveys. Bergemann and Morris (2016) find the optimal Bayes correlated equilibrium (BCE) and then design a disclosure policy that will generate a Bayes Nash equilibrium (BNE), which will have the same outcome as in the optimal BCE. The authors assume that the agents will play the principal's preferred BNE. In our coordination game, a simple full disclosure can eliminate panics if agents play the "good" BNE in which no one attacks a regime unless an attack is warranted by its fundamental. However, there is also a "bad" BNE under full disclosure, which admits the undesirable outcome. To this end, we adopt an adversarial information design approach (as in Mathevet, Perego and Taneva (2020)), which assumes that the principal anticipates that the agents will play her least preferred equilibrium.

Assuming adversarial selection, in a regime change game with conditionally independent Gaussian private signals, Goldstein and Huang (2016) find the optimal monotone binary public disclosure policy. Inostroza and Pavan (2020) significantly generalize this result and establish that the optimal public disclosure should always perfectly coordinate the agents' actions. Li, Song and Zhao (2019) find the optimal private disclosure when the agents share a homogeneous belief. Morris, Oyama and Takahashi (2020)

[7] Notice that, in our motivating examples, the agents endogenously choose the timing of their attack. The timing of the attack is exogenous when, for instance, a borrower issues a portfolio of short-term debt with diversified maturity dates, forcing the creditors to make their rollover decisions at different maturity dates. Together, these two studies provide a broad picture of how the principal can adopt dynamic information disclosure policies to foster coordination. However, these two papers are fundamentally different in terms of their underlying mechanisms and their scope of application. We will discuss the relation between the two studies in detail in Section 4.

characterize the set of implementable adversarial equilibrium outcomes in all binary action super modular games, including regime-change games. As we have already mentioned, this literature shows that under adversarial selection, panic cannot be averted using static persuasion. Our study contributes to this literature by showing that when one adds a natural timing dimension to the problem, perfect coordination on the desirable action can be achieved using a simple dynamic information disclosure policy.

Our paper also contributes to the growing literature on dynamic persuasion (see, for instance, Au (2015), Ely (2017), Ball (2019), Ely and Szydlowski (2020), Orlov, Skrzypacz and Zryumov (2020), Smolin (2021), and Orlov, Skrzypacz and Zryumov (2020)). Broadly speaking, this literature uses future disclosure as a carrot to persuade an agent to do what the principal wants. We study adversarial information design in dynamic regime change games.[8] In our setup, the design of future information, on the one hand, can be made dependent on whether agents wait for the future disclosure; on the other hand, it determines whether agents would choose to wait. We find that the optimal policy features a delayed binary disclosure (or disaster alert), and, as long as the period of delay is sufficiently short, all agents wait for and then follow the alert.[9]

Outline The rest of this paper is organized as follows. Section 1 describes the model and solution concept. Section 2 shows how a timely disaster alert eliminates panic in a robust sense. Section 3 discusses the limitations of and extensions to this design. Section 4 highlights the essential features of our optimal design that drives this positive result, and Section 5 concludes.

1 Model

The economy is populated by a principal, a continuum of risk-neutral agents, indexed by $i \in [0, 1]$, and a regime. To understand the dynamic aspect of panic, we introduce two features to the canonical regime change game: (1) a shock arrives at a future (possibly stochastic) date, which makes the regime vulnerable; and (2) the agents can attack at any time within a time window, where attack is irreversible and waiting is costly. Below, we describe the details of this model.

The shock's arrival date There is a time window $[0, T]$. A shock arrives once at some date t_s within this time window, following a commonly known distribution $G : [0, T] \to [0, 1]$. The regime becomes

[8] Only a few recent studies look into optimal disclosure in dynamic games. See, for instance, Li, Szydlowski and Yu (2021) for the dynamic entry game or Ely et al. (2021) for the dynamic contest.

[9] Orlov, Skrzypacz and Zryumov (2020) investigate a problem in which a receiver chooses when to exercise a real option, and the sender is biased toward either late or early exercise. They find that, under commitment, the optimal policy to persuade the receiver to wait is a delayed disclosure of all information; and, if the sender lacks commitment, it is optimal to "pipette" information gradually over time. In our dynamic persuasion problem, the timely disaster alert policy remains optimal even if the principal lacks commitment. See Section 4.4 for a detailed discussion.

vulnerable and may change only after the shock arrives. We assume that there is no mass at time 0 (i.e., $G(0) = 0$) and that the shock is certain to arrive by time T (i.e., $G(T) = 1$).

The fundamental of the regime Before the game begins, nature chooses the fundamental state θ from a commonly known distribution $\Psi : \Theta \to [0, 1]$, where Θ is a subset of $\mathbb{R}$. For a country facing a potential sovereign default, θ can be interpreted as the fiscal capacity of the country; for a bank, θ captures the strength of the balance sheet.

Irreversible Attack The agents can attack at any time within the time window $[0, T]$. In different applications of a regime change game, attacking could mean the withdrawal of early investment, exiting from a market, or shorting a currency. Since attack is irreversible, an agent simply chooses when to attack, $t_i \in [0, T]$, if at all. We represent not attacking in the given time window as $\mathbb{T}$, where $T < \mathbb{T} < \infty$. Therefore, the action space of any agent is $[0, T] \cup \{\mathbb{T}\}$. We say that an agent is active at any time $t \in [0, T]$ if and only if he has not attacked by time t. We define $N_t = \int_i \mathbb{1}\{t_i \leq t\} \mathrm{d}i$ as the mass of attack until time t (inclusive), which is non-decreasing in time t.

Timing We consider a continuous-time model. We use the notation t^- (and t^+) to capture the instance before (and after) t but excluding t. Between time t and time $t + \Delta t$ (where $\Delta t \to 0$), we allow for a sequence of events to occur in a particular order. First, the agents receive information (if there is any). Next, the agents decide whether to attack if they have not attacked already. Then, the regime outcome (whether the regime changes if it has not already) is determined. The details of how the regime outcome is determined will be discussed shortly. Following that, the agents who attack collect their payoffs. Finally, the shock may arrive if it has not arrived yet. There is no time discounting between these events. Figure 1 depicts the order of events.

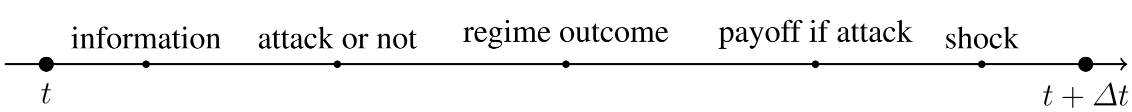

Figure 1: Timing

First, note that when the agents attack, they collect their payoff instantaneously in the same period. Second, note that we use the conventions that the shock arrives after the regime outcome is determined at any t. We use this convention to allow the agents a chance to attack and collect their payoffs at $t_i = 0$ even before any shock can arrive. Otherwise, our result will be even stronger. We discuss alternative timing conventions and some alternative cases in which the payoffs from attacking are not collected instantaneously in Section 3.1.

Regime Change There is a differentiable function $R(\theta, N_t)$ governing the regime's preparedness to face the shock at any time t. A stronger fundamental makes the regime better prepared (i.e., $\frac{\partial R}{\partial \theta} > 0$). On the other hand, more attacks weaken the regime's preparedness (i.e., $\frac{\partial R}{\partial N_t} < 0$). As we incorporate a timing dimension in the standard regime change game, the time when the regime changes, denoted by t_c, plays an important part in our analysis. Given θ and $(N_t)_{t \in [0,T]}$, if $R(\theta, N_T) \geq 0$, the regime never changes and survives in the end. This is denoted by $t_c = \infty$. Otherwise, if $R(\theta, N_T) < 0$, the regime does not survive in the end, it changes at the first instance after the shock arrives ($t > t_s$) when $R(\theta, N_t) < 0$.

Formally, a regime survives until time t (inclusive) either if (1) the shock has not arrived by the time the regime outcome is determined for time t, i.e., $t \leq t_s$,[10] or (2) the shock has arrived ($t > t_s$) but the fundamental is strong enough to withstand the attack so far, i.e., $R(\theta, N_t) \geq 0$. Assuming that the regime changes in the end ($R(\theta, N_T) < 0$), we define

$$t_{nc} := \sup\{\{t : 0 \leq t \leq t_s\} \cup \{t : t_s < t \leq T \;\&\; R(\theta, N_t) \geq 0\}\}$$

as the last date beyond which the regime cannot survive. By definition, $t_{nc} \geq t_s$. The equality holds when the regime changes the next instance of the shock arrival, i.e., $R(\theta, N_{t_s^+}) < 0$. Notice that this makes the second set in the above expression empty. Finally, we define t_c, the time at which the regime changes as

$$t_c = \begin{cases} t_{nc}^+ \text{ if } R(\theta, N_T) < 0 \\ \infty \text{ if } R(\theta, N_T) \geq 0. \end{cases} \tag{1}$$

By definition, either the regime never changes ($t_c = \infty$), or it changes at some date in the time interval $[0^+, T^+]$.

Payoff The agents are ex ante identical. They have a stationary discount rate $\beta > 0$ and are expected utility maximizers. If an agent chooses to attack at time $t_i \in [0, T]$, then he obtains

$$\pi(t_i, t_c) = e^{-\beta t_i}[\mathbb{1}\{t_i < t_c\}\overline{U} + \mathbb{1}\{t_i \geq t_c\}\underline{U}]. \tag{2}$$

Based on the above specification, the payoff from attacking depends on the agent's position in the queue of attacking agents. If he attacks early enough before most of the other attacking agents and thus before the time of the regime change ($t_i < t_c$), he obtains a high payoff $\overline{U}$; otherwise, he is too late to attack ($t_i \geq t_c$) and obtains a low payoff $\underline{U}$.[11]

[10]Recall that, by convention (See Figure 1), the regime outcome at time t is determined before the shock arrives. Therefore, if the shock arrives at time t, it does not influence the regime outcome at t but at t^+.

[11]In Example 1, where the shock arrives at the deterministic date T, the regime can either change at $t_c = T^+$ or never change $t_c = \infty$. Therefore, $t_i < t_c$ holds true for any $t_i \in [0, T]$. Accordingly, the payoff from attacking boils down to $\pi(t_i, t_c) = e^{-\beta t_i}\overline{U}$.

Assumption 1 (First-mover Advantage) *The payoff from attacking before the regime changes is higher than attacking after the regime changes:* $\overline{U} > \underline{U}$.

Note that an agent's action, by itself, does not affect t_c. Therefore, under this payoff specification, if an agent decides to attack, he is better off by attacking as early as possible — that is, attack at $t_i = 0$. Therefore, without any disclosure from the principal, this game boils down to the canonical binary action static regime change game. Notice that if the principal can achieve more than what she can achieve under static persuasion, then it is because of her ability to disclose information dynamically and not because the dynamic game is different from the static game.[12]

On the other hand, if an agent does not attack at all ($t_i = \mathbb{T}$), then he obtains

$$\pi(\mathbb{T}, t_c) = e^{-\beta T}[\mathbb{1}\{\mathbb{T} < t_c\}\overline{V} + \mathbb{1}\{\mathbb{T} \geq t_c\}\underline{V}].$$

For the agent who chooses not to attack, his payoff is realized at time T^+ depending on whether or not the regime changes in the end. If the regime does not change ($t_c = \infty > \mathbb{T}$), he gets a high payoff $\overline{V}$, and if the regime changes ($t_c \leq T^+ \leq \mathbb{T}$), he gets a lower payoff. We assume $\underline{V} \geq 0$.

Assumption 2 (Complementarity) *The payoffs satisfy the following inequalities:*

$$e^{-\beta T}\overline{V} > \overline{U}, \text{ and } \underline{U} > \underline{V}.$$

To understand the above assumption, first, suppose sufficiently many other agents choose to attack and, accordingly, the regime does not survive in the end — that is, $t_c \neq \infty$. Strategic complementarity dictates that conditional on a regime change, the payoff from attacking at any $t_i \in [0, T]$ is strictly higher than not attacking $t_i = \mathbb{T}$; that is,

$$\pi(t_i, t_c \neq \infty) - \pi(\mathbb{T}, t_c \neq \infty) = e^{-\beta t_i}[\mathbb{1}\{t_i < t_c\}\overline{U} + \mathbb{1}\{t_i \geq t_c\}\underline{U}] - e^{-\beta T}\underline{V} > 0.$$

The above inequality holding true for any $t_i \in [0, T]$ necessitates $\underline{U} > \underline{V}$. Second, suppose that insufficiently many agents attack and, accordingly, the regime survives in the end — that is, $t_c = \infty$. Strategic complementarity dictates that conditional on the regime surviving, the payoff from attacking at any $t_i \in [0, T]$ is strictly lower than that from not attacking $t_i = \mathbb{T}$; that is,

$$\pi(t_i, t_c = \infty) - \pi(\mathbb{T}, t_c = \infty) = e^{-\beta t_i}\overline{U} - e^{-\beta T}\overline{V} < 0$$

[12] Although this setup is natural for our examples, this may not be true in some other dynamic setup. For instance, consider riding a stock bubble, where an agent wants to wait until the last minute before the regime changes (see Abreu and Brunnermeier (2003)). Notice that in their setup, even without any disclosure, agents still have incentive to wait and therefore the game is very different from a canonical regime change game. We think this is a very interesting but substantially different setup. We leave the information design question in such a setup for future research.

The above inequality holding true for any $t_i \in [0, T]$ necessitates $\overline{U} < e^{-\beta T}\overline{V}$.

In Section 3.1, we consider some variants of this payoff structure. For instance, we consider a setup where there is no first-mover advantage or payoff depends on the final regime outcome, or it takes time to collect the payoff from attacking. We show that the results are robust under such variants.

Dominance Regions We assume that there exist $\underline{\theta}$ and $\overline{\theta} \in \Theta$ such that $R(\underline{\theta}, 0) = R(\overline{\theta}, 1) = 0$. Since the preparedness function R is strictly decreasing in the accumulated attack N_t and $N_t \in [0, 1]$ for any t, when $\theta \in \Theta^L \equiv \Theta \cap (-\infty, \underline{\theta})$, regardless of agents' strategy, the regime cannot survive, and it changes right after the shock arrives (i.e., $t_c = t_s^+$). When $\theta \in \Theta^U \equiv \Theta \cap [\bar{\theta}, +\infty)$, the regime will always survive (i.e., $t_c = \infty$) regardless of the attacks. We refer to Θ^U (or Θ^L) as the upper (or lower) dominance region, where not attacking $t_i = \mathbb{T}$ (or attacking right away $t_i = 0$) is the dominant strategy. Throughout the paper, we assume the existence of the dominance regions (i.e., $\Theta^L, \Theta^U \neq \emptyset$). Recall that the shock will definitely arrive by the end of the time window ($G(T) = 1$). However, it is possible that the fundamental is so strong ($\theta \in \Theta^U$) that the shock does not make the regime vulnerable to the attacks at all.

Exogenous Information The agents receive some noisy signals $s_i \in \mathbb{S}$ before the game begins. We assume that, given any underlying fundamental θ, the signal profile $s(\theta) \in \mathbb{S}^{[0,1]}$ is drawn from a distribution $F(s|\theta)$ with associated density $f(s|\theta)$. The signals may not be conditionally independent. For instance, if some agents share a common information source, or if they have some communication among themselves, then conditional on θ, the signals could be correlated. We allow for any arbitrary correlation, ranging from conditionally independent signals to perfectly correlated signals (homogeneous beliefs). The exogenous information structure is common knowledge.

We assume that the information-generating process F satisfies the following property.

Assumption 3 *(Doubt) There exists $\epsilon > 0$, such that any agent i with noisy signal $s_i \in \mathbb{S}$ believes that*

$$\mathbb{P}(\theta \in \Theta^U | s_i) = \frac{\int_{\theta \in \Theta^U} f_i(s_i|\theta) \mathrm{d}\Psi(\theta)}{\int_{\theta \in \Theta} f_i(s_i|\theta) \mathrm{d}\Psi(\theta)} > \epsilon, \ \mathbb{P}(\theta \in \Theta^L | s_i) = \frac{\int_{\theta \in \Theta^L} f_i(s_i|\theta) \mathrm{d}\Psi(\theta)}{\int_{\theta \in \Theta} f_i(s_i|\theta) \mathrm{d}\Psi(\theta)} > \epsilon,$$

where $f_i(s_i|\theta) = marg_{s_{-i}} f(s|\theta)$.

It holds that, regardless of the noisy signal an agent receives, he always assigns at least 2ϵ probability that θ is in the dominance regions for some $\epsilon > 0$. Recall that a regime with fundamental $\theta \in \Theta^U$ ($\theta \in \Theta^L$) survives (changes) regardless of the attacks. Hence, this assumption is equivalent to saying that, regardless of his signal and other agents' actions, an agent always has some doubt about whether his action is the right choice.[13]

[13] In particular, if the marginal distribution has full support and a bounded density — that is, $f_i(s_i|\theta) \in [\underline{f}, \overline{f}]$ for all $\theta \in \Theta$,

In Section 3.4, we discuss why the doubt assumption is more than necessary for our main result. For simplicity of exposition, we assume that the agents do not observe any new information over time. Later we relax this assumption. We show that the main insight can easily be extended when the agents observe the regime change (if it happens) (see Section 3.2) or infrequently receive additional noisy signals regarding the fundamental θ (see Section 3.3).

Principal As is standard in literature, we assume that the principal is concerned about the final outcome of the regime and prefers the survival of the regime. The principal's payoff is simply $\mathbb{1}\{R(\theta, N_T) \geq 0\}$. To go back to our examples, a regulator who values financial stability wants the bank to survive, IMF or world bank wants to avoid the sovereign debt crisis.[14] The principal does not know the private signals that the agents have received. However, unlike the agents, at any time t, the principal can observe the endogenous history of attacks until then; that is, $(N_s)_{s<t}$. In addition, through all necessary due diligence, the principal can figure out the fundamental of the regime (θ) and, thus, can assess the preparedness of the regime to face the shock even if the shock has not yet arrived.[15]

Disclosure Policy The principal *commits* to a dynamic public information disclosure policy denoted by $\Gamma = (d, \mathcal{S})$, where $\mathcal{S}$ is the message space, and $d : [0,T] \times \Theta \times [0,1] \to \Delta(\mathcal{S})$ is a mapping that specifies a public message depending on the date $\tau \in [0,T]$, the fundamental $\theta \in \Theta$, and the history of attack until then N_{τ^-}. Note that, at $\tau = 0$, any message can only be about the fundamental; for any $\tau > 0$, the message can vary depending on how many agents wait for the message. It is important to note that this disclosure rule allows the principal to send messages that vary with the history of attack N_{τ^-} (or, equivalently, the fraction of agents choosing to wait $1 - N_{\tau^-}$). This property is referred to as *endogenous disclosure*. Moreover, since the fundamental θ is realized before the game begins, the principal can also design information based on θ and disclose that information even before the shock actually realizes at t_s. We refer to this property as *forecasting*.[16]

Solution Concept and Adversarial Selection We consider the strategic form representation of the dynamic game. Given the disclosure policy Γ, a strategy of an agent of any type is to make a contingency

and $s_i \in \mathbb{S}$, where $0 < \underline{f} \leq \overline{f} < +\infty$ — then Assumption 3 holds true for any $\epsilon \in (0, \min\{\frac{(1-\Psi(\overline{\theta}))\underline{f}}{\overline{f}}, \frac{\Psi(\underline{\theta})\underline{f}}{\overline{f}}\})$.

[14]Our main result is robust if the principal wants to reduce the size attack when such attacks are unwarranted, or wants to maximize welfare. We discuss more nuanced preferences and the associated challenges in Section 4.4.

[15]This is consistent with the fact that bank supervisors and international organizations rely on historical data and econometric models to forecast the regime's preparedness to face some future shock. In Section 4.3, we discuss small errors in prediction.

[16]We can also allow for disclosure conditional on whether the shock has arrived, the whole history of attack (i.e., $(N_s)_{s \leq \tau^-}$), or private disclosure. However, as will be clear soon, these alternatives do not add any value in our setup. To focus our attention on the information design problem, we assume away any cost of acquiring information (or due diligence) and disclosing information. In real-world applications, these costs can be negligible compared with the loss brought about by a banking crisis or a sovereign debt crisis.

plan for when to attack ($t_i \in [0, T] \cup \mathbb{T}$) at the initial history $\emptyset$ and at every history based on the information disclosed by the principal. We use iterated elimination of strictly dominated strategies in the strategic form game as our *solution concept*. A strategy is rationalizable if it can survive the iterated elimination of strictly dominated strategies. Given any disclosure policy Γ, let $\mathcal{R}(\Gamma)$ be the set of rationalizable strategy profiles.[17,18]

The principal does not expect the agents to play the rationalizable strategy profile that is advantageous to her. Rather, she anticipates, state by state, the worst-case scenario that is consistent with some rationalizable strategy profile. We refer to this as *adversarial selection*. Define

$$\Theta^C(\Gamma) := \{\theta \in \Theta | R(\theta, N_T(x)) < 0 \text{ for some } x \in \mathcal{R}(\Gamma)\}.$$

The superscript C stands for regime change. In words, if $\theta \notin \Theta^C(\Gamma)$, then the regime will survive, regardless of whatever rationalizable strategies the agents play; otherwise, if $\theta \in \Theta^C(\Gamma)$, the regime will not survive under some rationalizable strategy profile $x \in \mathcal{R}(\Gamma)$.

Panic and Robust Information Design A distinctive feature of the regime change game is that when $\theta \in \Theta^L$, it is inevitable that the regime will change, regardless of what the agents do. Hence, no disclosure policy Γ can ensure the survival of such a regime (i.e., $\Theta^L \subseteq \Theta^C(\Gamma)$ for any Γ). However, a regime could also change even when it is not warranted ($\theta \notin \Theta^L$), as agents may choose to attack the regime if they believe that many others will do the same. We refer to this as *panic-based attacks*. Let us define $\Theta^P(\Gamma) := \Theta^C(\Gamma) \setminus \Theta^L$ for any policy Γ as the set of fundamentals in which the regime change may happen under some rationalizable strategy profile, and if it happens, it is caused by panic-based attacks.

Given the principal's payoff specification and under adversarial selection, she chooses Γ to maximize $\mathbb{E}[\mathbb{1}\{\theta \notin \Theta^C(\Gamma)\}]$. Since the regime will change regardless of Γ for any $\theta \in \Theta^L$, the principal's objective is equivalent to

$$\min_{\Gamma} \mathbb{P}(\theta \in \Theta^P(\Gamma)),$$

where $\mathbb{P}(\theta \in \Theta^P(\Gamma)) = \int_{\Theta^C(\Gamma)} \mathrm{d}\Psi(\theta)$. This is referred to as *adversarial (or robust) information design*. If $\Theta^P(\Gamma) = \emptyset$, then we can conclude that policy Γ eliminates panic in a robust sense.

[17] Note that the set of rationalizable strategy $\mathcal{R}(\Gamma)$ could be different based on the underlying exogenous information structure F. However, since F is exogenous and commonly known, for convenience, we suppress F from $\mathcal{R}(\cdot)$.

[18] Note that rationalizability requires common knowledge of rationality. However, as will be seen in Section 2, our main result only requires that the agents are rational and that they believe others are rational. Any higher order rationality is not required for our theory. Additionally, our results will hold true for more restrictive solution concepts such as Perfect Bayesian Equilibrium (PBE) (see the Online Appendix).

2 Main Result

In this section, we construct a simple dynamic information disclosure policy. This policy induces (even in the worst case) the agents to perfectly coordinate their actions and never attack a regime when it is not warranted ($\theta \notin \Theta^L$), thereby eliminating panic in the robust sense.

2.1 Static Benchmarks

To fix ideas, let us start with time 0 disclosure policies. Recall that the principal can send the message at $\tau = 0$, before the agents have their first opportunity to attack. Depending on the message from the principal and their own signals, agents may or may not attack. However, since a delayed attack is costly, if an agent decides to attack, he will do so immediately. Thus, the game boils down to the canonical static regime change game.

Full Disclosure Suppose that the fundamental θ is fully revealed at time 0. A possible equilibrium outcome is that all the agents perfectly coordinate on not attacking when $\theta \notin \Theta^L$. However, given the strategic complementarity, this is not the only rationalizable strategy profile. In another possible equilibrium, all agents attack at $t = 0$ whenever $\theta \notin \Theta^U$. Therefore, for a principal who anticipates the adversarial outcome, full disclosure is the worst policy because it maximizes the chance of regime change.

No Disclosure Suppose the principal does not provide any information. Then, depending on the distribution of the signals that the agents receive, there can be multiple equilibria. For instance, suppose that the agents share a homogeneous belief about θ. If they believe that the probability that $\theta \in \Theta^L$ is sufficiently small, there exists an equilibrium in which they do not attack. However, if they believe that the probability that $\theta \in \Theta^U$ is sufficiently small, then there also exists another equilibrium in which they all panic and attack. Using global game perturbation, Morris and Shin (2003) show that if the agents have sufficiently heterogeneous beliefs, there is a unique rationalizable strategy: attack if and only if the private signal is below a threshold. This means that unless the fundamental is sufficiently strong, there are panic-based attacks and the regime changes (see Section 3.4 for details).

Partial Disclosure There exist many time-0 partial disclosure policies. Consider the following policy Γ^0: at time 0, the principal publicly discloses whether or not the regime change is warranted ($\theta \in \Theta^L$, or equivalently, $R(\theta, 0) < 0$). There is an equilibrium in which the agents attack if and only if the principal sends the message that the regime change is warranted. However, there are other equilibria in which panic-based attacks are possible (see Angeletos, Hellwig and Pavan (2007)). To ensure that the agents never attack in any equilibrium, the positive news needs to be stronger than $\theta \notin \Theta^L$. Under conditionally independent Gaussian signals, Goldstein and Huang (2016) find the optimal binary (pass/fail) monotone public disclosure policy. The authors show that if the principal gives a pass grade iff $R(\theta, 0) > k$ for a sufficiently large k, then, under the adversarial selection, the agents will not attack when they see the regime gets a pass grade, and attack when they see the regime gets a fail grade. Notice that this disclosure policy perfectly coordinate the agents actions. However, when θ is such that $R(\theta, 0) \in [0, k]$, the regime gets a fail grade, and all the agents coordinate on attacking. Therefore, the regime changes although the the fundamental does not warrant a change (since $\theta \geq \underline{\theta}$). Inostroza and Pavan (2020) generalizes this result by showing that under conditionally independent super-modular signal, the optimal binary disclosure is a pass/fail disclosure that perfectly coordinates the agents actions.[19] Notice that under such optimal static disclosure $\Theta^P \neq \emptyset$, that is, under adversarial selection, panic cannot be averted using static persuasion.

The above discussion underscores two important aspects of the problem. First, to understand how to reduce the chance of panic, we cannot simply select the principal's preferred equilibrium. Many policies can eliminate panic if the agents play the principal's preferred equilibrium (such as full disclosure, or Γ^0 policy). However, these policies cannot do so in the robust sense. Second, panic cannot be eliminated simply by providing information early (at time 0, before the agents have any chance to attack). Below, we allow for a fairly flexible information environment in which, ex ante, the agents receive some arbitrarily correlated noisy signals (as long as there is some doubt). We show that the principal can exploit the endogenous waiting and construct a simple dynamic disclosure policy that eliminates panic in the robust sense.

2.2 A Simple Dynamic Disclosure Policy: Disaster Alert

We consider a simple partial disclosure policy. The principal sends a binary signal $\mathcal{S} = \{0, 1\}$ at a fixed date $\tau \in (0, T)$. We refer to this policy as Γ^τ.[20] The public signal $d^\tau \in \{0, 1\}$ is generated based on the underlying fundamental θ and the history of attacks N_{τ^-} as follows:

$$d^\tau(\theta, N_{\tau^-}) = \begin{cases} 1 \text{ if } R(\theta, N_{\tau^-}) < 0; \\ 0 \text{ otherwise.} \end{cases}$$

We call d^τ a *disaster alert*: the alert is triggered if $d^\tau = 1$ and is not triggered if $d^\tau = 0$.[21] When the alert is triggered ($d^\tau = 1$), it implies that regardless of the agents' actions after the disclosure, the regime cannot survive (i.e., $t_c < \infty$). To see this, first consider the case in which, ex post, the shock arrives before the time of disclosure (i.e., $t_s < \tau$). Since $R(\theta, N_{\tau^-}) < 0$, the regime has already changed (i.e.,

[19] The authors also show that under some conditions, such optimal pass/fail disclosure could be non-monotonic.

[20] Since this is a one-time disclosure at a specific date τ, we simplify notation by using τ as a superscript (rather than as an argument in the disclosure policy).

[21] In the spirit of Bayesian persuasion, this can be thought of as the principal sending a recommendation at time τ to the agents to "attack" when the disaster alert is triggered and to "not attack" otherwise.

$t_c \le \tau$). Next, consider the case in which, ex post, the shock has not yet arrived by the time of disclosure (i.e., $t_s \ge \tau$). However, since $R(\theta, N_{\tau^-}) < 0$, the regime will change as soon as the shock arrives (i.e., $t_c = t_s^+$).

On the other hand, if the alert is not triggered ($d^\tau = 0$), the agents learn that, without further attacks, the regime will survive regardless of when the shock arrives (i.e., $R(\theta, N_T = N_{\tau^-}) \ge 0$). However, the survival of the regime is not guaranteed following $d^\tau = 0$. If some agents choose to attack after no alert, the regime may still change.

In this spirit, the proposed disaster alert policy resembles a real-world EWS, which aims to generate an accurate forecast of a future crisis by all historical information. Notice that the proposed policy is simple: it is binary and public. However, it is an *endogenous disclosure.* The message at date τ depends on the fraction of agents who choose to wait for the disclosure ($1 - N_{\tau^-}$). If some agents do not wait for the disaster alert and attack before time τ, then the attack may trigger the alert ($R(\theta, N_{\tau^-}) < 0$). Moreover, it *forecasts*. The alert is triggered ($R(\theta, N_{\tau^-}) < 0$) even when the regime has not yet changed ($t_s \ge \tau$), but it is evident that it will change.

2.3 Optimality of Disaster Alert

The proposed policy can be interpreted simply as an assurance from the principal. That is, "Do not panic; just wait until time τ, and at that time, I will send you an alert if a regime change is inevitable." The following theorem shows that if the principal asks the agents to wait for only a little while (small enough τ), then, under the unique rationalizable strategy, the agents will always do so and follow the principal's advice regardless of their own information.

Theorem 1 *Under $G(0) = 0$, there exists $\hat{\tau} > 0$ such that under the disclosure policy Γ^τ with $\tau \in (0, \hat{\tau})$, panic is eliminated in the robust sense; that is, $\Theta^P(\Gamma^\tau) = \emptyset$.*

Theorem 1 claims that when the principal sets the disaster alert in a timely manner, she achieves her most preferred outcome as the unique rationalizable outcome. We prove this theorem using three lemmas that we will discuss next.

Option Value of Waiting

Consider the game under the disclosure policy Γ^τ where $\tau \in (0, T)$. Nature moves first and selects (θ, s), where s is the profile of signals for each agent. Each agent $i \in [0, 1]$ learns his own type s_i. An agent i of type s_i makes a contingency plan for when to attack at three possible histories: the initial history ($\emptyset$), the one after the alert is not triggered ($d^\tau = 0$), and the one after the alert is triggered ($d^\tau = 1$). Let $t_\emptyset$, t_0, and

t_1 be the time of attack at these three histories, respectively. Therefore, the strategy space for each agent is

$$\mathcal{X} := \{(t_\emptyset, t_0, t_1) \in [0, T] \cup \{\mathbb{T}\} \times ([\tau, T] \cup \{\mathbb{T}\})^2\}.$$

Consider an agent i of type s_i who plays a strategy $x_i = (t_\emptyset, t_0, t_1)$, while others play strategy x_{-i}. Let $u_i(x_i, x_{-i}; s_i)$ be his expected payoff. If $t_\emptyset \in [0, \tau)$, then

$$u_i(x_i, x_{-i}; s_i) = \mathbb{E}[\pi(t_\emptyset, t_c)|x_{-i}, s_i].$$

Note that if $t_\emptyset \in [0, \tau)$, then the agent is no longer active at time τ. Therefore, the path of play and, hence, the payoff are unaffected by the specification of t_0 and t_1. On the other hand, if $t_\emptyset \in [\tau, T] \cup \{\mathbb{T}\}$, then

$$\begin{aligned} u_i(x_i, x_{-i}; s_i) =& \mathbb{P}(d^\tau = 0|x_{-i}, s_i)\mathbb{E}[\pi(t_0, t_c)|x_{-i}, s_i, d^\tau = 0] \\ &+ \mathbb{P}(d^\tau = 1|x_{-i}, s_i)\mathbb{E}[\pi(t_1, t_c)|x_{-i}, s_i, d^\tau = 1]. \end{aligned}$$

As the agent is still active at time τ, the path of play and, hence, the payoff of the agents are determined only by t_0 and t_1 but not $t_\emptyset$.

Lemma 1 *(Option value) For any agent i of any type s_i, the only rationalizable strategies are*

- *strategy $\mathcal{A}$: $(0, t_0, t_1)$, where $t_0, t_1 \in [\tau, T] \cup \{\mathbb{T}\}$; and*
- *strategy $\mathcal{W}$: $(t_\emptyset, \mathbb{T}, \tau)$, where $t_\emptyset \in [\tau, T] \cup \{\mathbb{T}\}$.*

As we have already mentioned, for any agent i, given any strategy profile played by others, for any $t_0, t_1 \in [\tau, T] \cup \{\mathbb{T}\}$, the strategy $(0, t_0, t_1)$ leads to the same path of play and, hence, the same payoff for this agent. We refer to all such strategies as $\mathcal{A}$ (attack immediately). Similarly, for any $t_\emptyset \in [\tau, T] \cup \{\mathbb{T}\}$, strategy $(t_\emptyset, \mathbb{T}, \tau)$ leads to the same path of play and, hence, the same payoff for agent i. We refer to all such strategies as $\mathcal{W}$ (wait and follow).

The intuition behind Lemma 1 is simple. First, consider an agent who makes a plan not to wait for the alert or, equivalently, to attack before the disclosure date (i.e., $t_\emptyset \in [0, \tau)$). Since waiting is costly, in the absence of any new information, he will attack as early as possible to avoid an unnecessary delay cost. Therefore, any strategy that involves $t_\emptyset \in (0, \tau)$ is strictly dominated by attacking immediately at $t_\emptyset = 0$ (strategy $\mathcal{A}$).

Next, suppose that an agent makes a contingency plan whereby he waits for the alert (i.e., $t_\emptyset \geq \tau$). Conditional on $d^\tau = 1$, as it predicts a regime change regardless of what others do thereafter, this agent will attack and will do so immediately at $t_1 = \tau$ to save some waiting cost. As such, any strategy $(t_\emptyset, t_0, t_1)$ that involves $t_\emptyset \geq \tau$ and $t_1 > \tau$ is dominated by $(t_\emptyset, t_0, t_1 = \tau)$ for the same $t_\emptyset$ and t_0. It follows from the doubt assumption that the agent assigns a positive probability that $d^\tau = 1$ regardless of his type. This makes the dominance strict.

However, given that an agent makes a contingency plan whereby he waits and attacks immediately after $d^\tau = 1$ (i.e., $t_\emptyset \geq \tau$ and $t_1 = \tau$), if he also attacks after $d^\tau = 0$ (i.e., $t_0 \in [\tau, T]$), then the information d^τ has no value. This agent should have attacked at time 0 and saved the cost of waiting for the alert. That is, strategy $\mathcal{A}$ strictly dominates any strategy involving waiting and attacking regardless of d^τ (i.e., $(t_\emptyset, t_0, t_1 = \tau)$ with $t_\emptyset \geq \tau$ and $t_0 \in [\tau, T]$).[22] Thus, if a rational agent chooses to wait for the disaster alert, it must be that he will follow it — that is, attack if and only if the alert is triggered (strategy $\mathcal{W}$).[23]

Different from disclosing information at $t = 0$, under Γ^τ, the agents need to wait to see the message from the principal. To understand the role of future disclosure (i.e., $\tau > 0$), let us contrast this with the Γ^0 disclosure policy (see Section 2.1). Recall that under the policy Γ^0, there exists an equilibrium in which an agent may attack after learning $d^0 = 0$ (i.e., $R(\theta, 0) \geq 0$). Note that under policy Γ^τ with $\tau > 0$, an agent may still plan to attack at the information set $d^\tau = 0$ (i.e., $t_0 \in [\tau, T]$). However, based on Lemma 1, this case can occur only under strategy $\mathcal{A}$; that is, this agent will attack at time 0 and will no longer be active at the time of disclosure. Thus, when the agents play some rationalizable strategies, on path, attacks cannot happen after $d^\tau = 0$.

Predictability of Disaster Alert

By Lemma 1, when an agent plays some rationalizable strategy, if he waits for the disclosure, he will follow the disaster alert d^τ (i.e., attack if and only if the disaster alert is triggered). Accordingly, the public signal d^τ perfectly coordinates the actions of these agents from time τ onward. Recall that the agents who remain active at the disclosure time τ are the ones who play the "wait and follow" strategy ($\mathcal{W}$). Conditional on no alert ($d^\tau = 0$), they will never attack (i.e., $t_0 = \mathbb{T}$). Therefore, the regime always survives following $d^\tau = 0$ since $R(\theta, N_T) = R(\theta, N_{\tau^-}) \geq 0$. Consequently, as long as agents play rationalizable strategies (either $\mathcal{A}$ or $\mathcal{W}$), the regime changes if and only if $d^\tau = 1$. In other words, the binary public signal d^τ predicts the ultimate regime status.

Following Lemma 1, we can also infer the time of regime change t_c given that $d^\tau = 1$. By definition of d^τ, the alert is triggered ($d^\tau = 1$) if and only if $R(\theta, N_{\tau^-}) < 0$. By Lemma 1, the agents who attack

[22] It is worth noting that the result in Lemma 1 only involves one round of elimination of dominated strategies, and it does not even require the agents to hold the belief that others will never play dominated strategies.

[23] A similar option value of the waiting argument appears in the context of social learning, in which an agent can learn from others' actions, but such actions do not affect his payoff (see Chamley and Gale (1994) and Gul and Lundholm (1995)). The intuition is simple: Consider two agents deciding whether to attack at time 1 or time 2. If an agent waits to see whether the other agent attacks, it must hold that he will take different actions, conditional on whether the other agent attacks at time 1. Otherwise, there is no positive option value of waiting. The result also has an intuitive connection to the coordination with an outside option example in the forward-induction literature, introduced by Kohlberg and Mertens (1986). The outside option appears naturally in our dynamic setting, as the agents can choose to attack immediately rather than waiting for the future disclosure. Forward induction is formalized using iterated weak dominance (Ben-Porath and Dekel, 1992). However, note that given the doubt assumption, we can use the standard iterated elimination of strictly dominated strategies (IESDS).

before the disclosure time τ are the ones who play strategy $\mathcal{A}$ and attack at time 0. Therefore, $N_{\tau^-} = N_0$, implying that $R(\theta, N_0) < 0$. By definition of t_c, this means that the regime changes as soon as the shock arrives (i.e., $t_c = t_s^+$). The following lemma summarizes this.

Lemma 2 *(Perfect Predictability) Under policy Γ^τ, (1) if $d^\tau = 0$, the regime survives — that is, $t_c = \infty$; and (2) if $d^\tau = 1$, the regime changes as soon as the shock arrives — that is, $t_c = t_s^+$.*

Timely Alert and Negligible Waiting Cost

Consider an agent who receives a signal s_i. He faces some fundamental uncertainty regarding θ: he has some doubt about whether his action is going to be the right choice (regardless of what others do). Moreover, he still faces strategic uncertainty about whether or not other agents will wait for the disclosure. However, as long as all agents who wait for the disclosure will follow the disaster alert (Lemma 1), the disaster alert can perfectly predict the regime outcome (Lemma 2). Therefore, if he waits for the alert, his private information s_i becomes irrelevant. The disaster alert can help him avoid the mistake of attacking a regime that will survive in the end. Nevertheless, an agent who waits for disclosure may end up attacking the regime at a later date. Such a delay in attacking is costly, especially so if, by waiting, the agent misses the opportunity to attack the regime before it changes (i.e., $t_1 = \tau \geq t_c$). Thus, an agent will be reluctant to play strategy $\mathcal{W}$ if he believes that the alert is likely to be triggered.

Lemmas 1 and 2 hold true regardless of the time of disclosure τ as long as $\tau > 0$. Next, we explore the timing of disclosure and discuss the role of a "timely" disclosure.

Lemma 3 *(Timely Alert) There exists $\hat{\tau} > 0$ such that under the disclosure policy Γ^τ with any $\tau \in (0, \hat{\tau})$, the unique rationalizable strategy for any agent $i \in [0, 1]$ and any signal $s_i \in \mathbb{S}$ is $\mathcal{W}$; that is, wait for the disclosure and then follow the alert.*

Lemma 3 establishes the dominance of the "wait and follow" strategy ($\mathcal{W}$) over attacking immediately ($\mathcal{A}$) under a timely disaster alert. By convention, if the agent attacks immediately ($\mathcal{A}$), he obtains $\overline{U}$ regardless of t_c. That is, regardless of the fundamental, others' strategies, and the shock arrival time, he gets the highest possible payoff from attacking if he attacks immediately. In comparison, waiting for the alert ($\mathcal{W}$) is profitable if the alert is not triggered ($d^\tau = 0$), but it is costly if the alert is triggered ($d^\tau = 1$).

The Benefit of Waiting If no alert is triggered ($d^\tau = 0$), the agent expects the regime to survive in the end (Lemma 2). In this case, the "wait and follow" strategy ($\mathcal{W}$) prevents agents from making the mistake of attacking a regime that survives. That defines the benefit of waiting. Conditional on $d^\tau = 0$, the expected benefit from playing $\mathcal{W}$, as compared to $\mathcal{A}$, is

$$\mathbb{B}(\Gamma^\tau, x_{-i}, s_i) := \mathbb{E}\left[\pi(\mathbb{T}, t_c) - \pi(0, t_c) \middle| x_{-i}, s_i, d^\tau = 0\right].$$

Recall that the agent who plays $\mathcal{W}$ chooses $t_0 = \mathbb{T}$ when $d^\tau = 0$. It follows from Lemma 2 that as long as others play a rationalizable strategy (i.e., $x_j \in \{\mathcal{A}, \mathcal{W}\}$), conditional on $d^\tau = 0$, the regime never changes ($t_c = \infty$). Therefore,

$$\mathbb{B}(\varGamma^\tau, x_{-i}, s_i) = \pi(\mathbb{T}, \infty) - \pi(0, \infty) = e^{-\beta T}\overline{V} - \overline{U}. \tag{3}$$

This expected benefit is independent of the disclosure time τ, noisy information s_i, and others' rationalizable strategies x_{-i}, and it is strictly positive (by Assumption 2).

The Cost of Waiting On the other hand, conditional on $d^\tau = 1$, the regime changes as soon as the shock arrives (i.e., $t_c = t_s^+$) (Lemma 2). Recall that, given $d^\tau = 1$, an agent who plays strategy $\mathcal{W}$ attacks at $t_1 = \tau$. Therefore, the expected payoff from the "wait and follow" strategy ($\mathcal{W}$) is[24]

$$\begin{aligned}\mathbb{E}(\pi(\tau, t_c)|x_{-i}, s_i, d^\tau = 1) =& e^{-\beta\tau}\int_{t_s=0}^{T} \left(\mathbb{1}\{t_s \geq \tau\}\overline{U} + \mathbb{1}\{t_s < \tau\}\underline{U}\right) \mathrm{d}G(t_s) \\ =& e^{-\beta\tau}\left((1 - G(\tau^-))\overline{U} + G(\tau^-)\underline{U}\right). \end{aligned} \tag{4}$$

Accordingly, the expected cost of waiting is

$$\begin{aligned}\mathbb{C}(\varGamma^\tau, x_{-i}, s_i) :=& \mathbb{E}\left[\pi(0, t_c) - \pi(\tau, t_c)\middle| x_{-i}, s_i, d^\tau = 1\right] \\ =& (1 - G(\tau^-))(1 - e^{-\beta\tau})\overline{U} + G(\tau^-)(\overline{U} - e^{-\beta\tau}\underline{U}). \end{aligned} \tag{5}$$

Similar to the benefit, the expected cost $\mathbb{C}$ is independent of the signal s_i since $d^\tau = 1$ perfectly predicts that the regime will change as soon as the shock arrives. Unlike the benefit, the cost depends on the disclosure time τ. To see this, note that, with probability $G(\tau^-)$, the shock arrives and, consequently, the regime changes before the time of disclosure (i.e., $t_s < \tau$ and $t_c = t_s^+ \leq \tau$). In this case, waiting induces a reduction in the payoff from $\overline{U}$ to $e^{-\beta\tau}\underline{U}$. However, with probability $(1 - G(\tau^-))$, the shock does not arrive and, consequently, the regime changes after time τ (i.e., $t_s \geq \tau$ and $t_c = t_s^+ > \tau$). In this case, the cost induced by waiting is simply the loss of time value of $\overline{U}$ (i.e., $(1 - e^{-\beta\tau})\overline{U}$).

An early disclosure limits the waiting cost since $\mathbb{C}(\varGamma^\tau, x_{-i}, s_i)$ falls as τ decreases. The loss of time value $(1 - e^{-\beta\tau})\overline{U}$ decreases continuously to 0 when the disclosure time τ decreases to 0. The assumption that $G(0) = 0$ and the fact that $G(\cdot)$ is right continuous (since it is a cumulative distribution function) ensure that the probability of a significant payoff drop $G(\tau^-)$ is negligibly small when τ is set sufficiently close to 0. Therefore, although the decrease in payoff $\overline{U} - e^{-\beta\tau}\underline{U}$ is bounded away from 0, when τ decreases to 0, the expected cost is $\mathbb{C}(\varGamma^\tau, x_{-i}, s_i) \to 0$. This holds true for any agent i, regardless of his

[24]Notice that if $t_s \geq \tau$, then, following our timing convention (see Figure 1), the regime does not change when the agent attacks at time τ.

signal s_i and others' rationalizable strategy x_{-i}.

The Dominance of $\mathcal{W}$ over $\mathcal{A}$ It follows from Assumption 3 that regardless of the signal s_i and other agents' actions, there is at least ϵ chance that the disaster alert will not be triggered, and that, accordingly, attacking early will be a mistake. This enables us to get a lower bound for the net expected benefit from playing $\mathcal{W}$ as compared to $\mathcal{A}$ for any given s_i; that is,

$$\begin{aligned}\mathbb{D}(\Gamma^\tau, x_{-i}, s_i) &:= \mathbb{P}(d^\tau = 0|x_{-i}, s_i)\mathbb{B}(\Gamma^\tau, x_{-i}, s_i) - \mathbb{P}(d^\tau = 1|x_{-i}, s_i)\mathbb{C}(\Gamma^\tau, x_{-i}, s_i) \\ &\geq \epsilon \cdot \mathbb{B}(\Gamma^\tau, x_{-i}, s_i) - (1-\epsilon) \cdot \mathbb{C}(\Gamma^\tau, x_{-i}, s_i). \end{aligned} \tag{6}$$

Recall that the benefit $\mathbb{B}(\Gamma^\tau, x_{-i}, s_i)$ and the cost $\mathbb{C}(\Gamma^\tau, x_{-i}, s_i)$ are independent of x_{-i} and s_i. The benefit is strictly positive and is independent of τ, whereas the cost is increasing in τ and converges to 0 as τ decreases to 0. Therefore, there exists $\hat{\tau} > 0$ such that the net benefit of waiting $\mathbb{D}(\Gamma^\tau, x_{-i}, s_i)$ is strictly positive if the time of disclosure is set at any $\tau \in (0, \hat{\tau})$. As such, a timely disaster alert ensures that the expected benefit outweighs the expected cost, irrespective of the signal s_i and other agents' strategies (as long as they play rationalizable strategies $\mathcal{W}$ and $\mathcal{A}$). In other words, under a timely disaster alert, for $s_i \in \mathbb{S}$, strategy $\mathcal{W}$ strictly dominates strategy $\mathcal{A}$.

It is worth noting that the strict dominance of $\mathcal{W}$ involves only two rounds of iterated elimination of strictly dominated strategies, whereas higher orders of rationality are not required.

Theorem 1 follows immediately from the above lemmas. Given that the disaster alert is set in a timely manner, under the unique rationalizable strategy profile, all agents wait for the disclosure and follow the alert afterward. Since all the agents wait for the alert, any regime that can survive without attack ($\theta \notin \Theta^L$) will not trigger the alert ($d^\tau = 0$). Because the agents will not attack when the alert is not triggered, any regime with $\theta \notin \Theta^L$ will never change. Therefore, any timely disaster alert policy Γ^τ with $\tau \in (0, \hat{\tau})$ eliminates panic in the robust sense (i.e., $\Theta^P(\Gamma^\tau) = \emptyset$).

3 Extensions

In this section, we discuss the robustness of the timely disaster alert policy by extending the dynamic regime change model to consider alternative scenarios of the exogenous information structure, timing setup, and payoff specification.

3.1 Time Gap and Negligible Waiting Cost

Since the shock arrives in the future, and the regime becomes vulnerable only after the shock arrives, it creates a time gap between when the agents can start attacking (time 0) and when the regime changes (t_c).

When the shock arrives at a deterministic date T (as in Example 1), if the regime changes, it changes only at $t_c = T^+$. Thus, the regime cannot change while the agent waits for the disclosure at time τ. When the shock arrives at a stochastic date $t_s \in [0, T]$ (as in Example 2), the regime can change at any date $t_c \in [0^+, T^+]$. The probability that the regime can change before any given disclosure date τ is, at most, $G(\tau^-)$. Under the assumption $G(0) = 0$, the probability that the regime will change before the time of disclosure can be made arbitrarily small when the disclosure time τ is sufficiently close to zero.

This time gap is essential for our main result. Recall that the cost of waiting for the disaster alert is (see (5))

$$\mathbb{C}(\Gamma^\tau, x_{-i}, s_i) = (1 - G(\tau^-))(1 - e^{-\beta\tau})\overline{U} + \underbrace{G(\tau^-)(\overline{U} - e^{-\beta\tau}\underline{U})}_{\mathbb{C}_2(.)}. \tag{7}$$

The second term in this cost, denoted by $\mathbb{C}_2(.)$, is a result of the missed opportunity of attacking the regime before it changes. Suppose that there is a significant probability that the shock arrives before the disclosure date. For instance, consider the extreme case with $G(0) = 1$. This means that regardless of τ when the alert is triggered ($d^\tau = 1$), the regime has already changed. Hence, there is no time gap. In this case, the first term in (7), which captures the cost of waiting purely due to time discounting, will be negligibly small when the disclosure time τ is sufficiently close to 0. However, the second term is bounded away from 0; that is, independent of τ,

$$\mathbb{C}_2(\Gamma^\tau, x_{-i}, s_i) \geq G(0)(\overline{U} - \underline{U}) > 0.$$

As a result, the disclosure policy may fail to persuade the agents to wait for the alert.

However, in some specific environments, even with no time gap, a timely disaster alert can still be effective. Below, we describe a few of these environments by changing some features of our dynamic setup. We argue that in these alternative environments, our main result in Theorem 1 will hold even when the shock arrives before time 0 with certainty (i.e., $G(0) = 1$).

Alternative Timing Convention

According to our timing convention (see Figure 1), the regime cannot change at time 0. This allows the agents to attack at time zero and collect $\overline{U}$ even before any shock arrives. Consider an alternative setting in which between the time interval t and $t + \Delta t$, the shock arrives before the regime outcome is determined (either before or after the agents decide to attack). Then, if the shock arrives at time 0, the regime can change before any agents can attack the regime. In this case, the agents who attack at time 0 get only $\underline{U}$ (instead of $\overline{U}$). Therefore, the probability that an agent misses the opportunity to attack a regime before it changes while waiting for the alert is $G(\tau^-) - G(0)$ instead of $G(\tau^-)$. Therefore, $\mathbb{C}_2(.)$ changes to

$$\mathbb{C}_2^A(\Gamma^\tau, x_{-i}, s_i) = [G(\tau^-) - G(0)](\overline{U} - e^{-\beta\tau}\underline{U}).$$

Since $G(.)$ is a cumulative density function (CDF), it is right continuous, and hence, the cost $\mathbb{C}_2^A(.)$ can be made arbitrarily small when the disclosure time τ is sufficiently close to 0. Therefore, regardless of $G(\cdot)$, Theorem 1 will hold true.

No First-mover Advantage

Suppose there is no first-mover advantage (i.e., $\overline{U} = \underline{U} = U$). This means that the payoff from attacking a regime is independent of whether the attack happens before or after the regime changes. Accordingly, $\mathbb{C}_2(.)$ becomes

$$\mathbb{C}_2^{NF}(\Gamma^\tau, x_{-i}, s_i) = G(\tau^-)(1 - e^{-\beta\tau})U,$$

and the total cost of waiting, $\mathbb{C}^{NF}(\Gamma^\tau, s_i, x_{-i}) = (1 - e^{-\beta\tau})U$, is independent of $G(\cdot)$ and decreases to 0 continuously when the disclosure time τ decreases to 0. This makes Theorem 1 even stronger.[25]

Payoff depends only on final outcome

In our benchmark model, we assume that the payoff from attacking is collected instantaneously, which depends only on the regime status at the time of attack. Suppose, instead, that the payoff from attacking at t_i is determined by the final regime outcome: $\pi^{NI}(t_i, t_c) = e^{-\beta t_i}\left[\mathbb{1}\{t_c = \infty\}\overline{U} + \mathbb{1}\{t_c < \infty\}\underline{U}\right]$. Note that, under this payoff specification, there is no first-mover advantage. Although the payoff from attacking is higher when the regime survives ($t_c = \infty$), as compared to that conditional on regime change ($t_c < \infty$), attacking at a later date can never cause a discrete drop from $\overline{U}$ to $\underline{U}$ in the payoff. Conditional on the the alert being triggered ($d^\tau = 1$ and, accordingly, $t_c < \infty$), an agent gets $\underline{U}$ if he attacks at $t_i = 0$ and $e^{-\beta\tau}\underline{U}$ if he waits for the disaster alert. Hence, similar to that in the case of no first-mover advantage, the total cost of waiting is

$$\mathbb{C}^{NI}(\Gamma^\tau, s_i, x_{-i}) = (1 - e^{-\beta\tau})\underline{U}.$$

Obviously, regardless of $G(0)$, this cost of waiting decreases continuous to 0 when the disclosure time τ decreases to 0. Therefore, when the payoffs from attacking are not collected instantaneously but dependent only on the final regime outcome, Theorem 1 holds true regardless of $G(\cdot)$.[26]

[25] In practice, some policies are designed to remove the first-mover advantage. For instance, the bankruptcy code "avoidable preference" prevents creditors from being treated more favorably than others. For details, see Sections 547 and 550 in Chapter 11 of the U.S. Bankruptcy Code. It is implemented by calling back debt repayments made shortly before bankruptcy. The proceeds are then shared among all creditors in bankruptcy court. Under this bankruptcy code, there will be no first-mover advantage, and the payoff that the attacking agents get may be subject to the final status of the regime, thereby slightly complicating the payoff specifications. Nevertheless, given that this type of policy helps to remove the first-mover advantage, the same arguments can be used to demonstrate that such policy makes the timely disaster alert policy more effective at preventing panic.

[26] If one moves beyond the financial market applications (for instance, political regime change), it is plausible that an agent gets a lower payoff from attacking when the regime survives at the end ($t_c = \infty$) than when it changes ($t_c < \infty$). Thus, the condition $\overline{U} > \underline{U}$ (Assumption 1) is violated. Nonetheless, when the payoff of attacking is determined only by the final

Slow-moving Capital

Another reason that the payoffs from attacking are not paid instantaneously is that moving investments away from a country or withdrawing funds from financial institutions often takes time. This could be a result of some exogenous constraints. For example, there are constraints on cross-board capital flows, and withdrawing from money market mutual funds or hedge funds might be subject to redemption gates. Or delays can happen simply because moving capital away, which involves selling off properties and firing employees, is time consuming.

Suppose, as in our benchmark setup, the agents choose when to attack and an attack is irreversible. However, it takes $l > 0$ time to collect the entire payoffs from an attack. An agent who starts attacking at time t_i receives a high flow payoff $\overline{u}$ at any instant $t \in [t_i, t_i + l]$ if the regime has not changed yet ($t < t_c$) and a low flow payoff $\underline{u} \in (0, \overline{u})$ if the regime has already changed ($t \geq t_c$). We assume that any agent who has started attacking at t_i finishes $B\left(\frac{t-t_i}{l}\right)$ part of his attack by time t, where $B(\cdot)$ has support $[0, 1]$ and admits a density $b(\cdot)$. This generalizes our benchmark model, which becomes a special case with $l \to 0$. We assume the first-mover advantage (Assumption 1) and complementarity (Assumption 2) with regard to the flow payoffs.

Proposition 1 *If $l > 0$ and $b(\cdot)$ is Lipschitz continuous, then regardless of $G(\cdot)$, there exists $\hat{\tau}_l \in (0, l]$ such that under disaster alert Γ^τ with $\tau \in (0, \hat{\tau}_l)$, $\Theta^P(\Gamma^\tau) = \emptyset$.*

Proposition 1 demonstrates that, when the market moves slowly, the waiting cost for a timely disclosure can be made negligible regardless of $G(\cdot)$. To see this argument, consider the special case with uniform distribution (i.e., $B(\cdot) = \mathcal{U}[0, 1]$). Note that, conditional on $d^\tau = 1$, an agent gets the same flow payoff in the interval $[\tau, l]$ regardless of whether he plays $\mathcal{A}$ or $\mathcal{W}$ and, thus, the cost of waiting $\mathbb{C}(\Gamma^\tau, x_{-i}, s_i)$ is no higher than the flow payoff he loses in the interval $[0, \tau)$ — that is, $\frac{\overline{u}\tau}{l}$. Therefore, as τ decreases to 0, the cost of waiting for the timely disaster alert decreases continuously to 0. Accordingly, following the same insight from Theorem 1, we can show that a timely disaster alert eliminates panic in the robust sense.

From the above discussions, it should be clear that the assumptions on the payoff specification and timing we adopted in the benchmark model (that is, instantaneous payment, first-mover advantage, and the timing convention) are most conducive to panic and, thus, represent the most challenging case for the principal. They are not necessary for our result to hold true, whereas relaxing these assumptions strengthens our result.

---

regime outcome, what is essential to our theory is the complementarity in payoffs (see Assumption 2) and costly delay, whereas Assumption 1 plays no role in our analysis.

3.2 Observable Regime Change

In many applications (e.g., bank runs), the agents might naturally observe when the regime changes. Consider an exogenous information structure in which regime change is observable. If the agents have seen no regime change, then the timely disaster alert policy serves purely as a warning of impending regime change, and, thus, it will become more effective, meaning that the time cutoff $\hat{\tau}$ can be greater.

To see this, note that when the agents can observe the regime change, under a disaster alert policy Γ^{τ}, the only rationalizable strategies are: (1) attack right away ($\mathcal{A}$) and (2) attack as soon as the regime has changed, or attack at time τ if the regime has not changed by time τ but the alert is triggered ($d^{\tau} = 1$), and do not attack otherwise. Let us call this second strategy $\mathcal{W}_o$. The difference between strategies $\mathcal{W}_o$ and $\mathcal{W}$ is that, with $\mathcal{W}_o$, the agent, during the waiting period, can attack at an earlier date if the regime change occurs before the disclosure time τ. Thus, compared with the case in which the regime change cannot be observed, the cost of waiting for the disaster alert decreases. Formally,

$$\mathbb{C}_2^O(\Gamma^{\tau}, s_i, x_{-i}) = G(\tau^-)\overline{U} - \int_0^{\tau^-} e^{-\beta t}\underline{U}\mathrm{d}G(t) < G(\tau^-)\left(\overline{U} - e^{-\beta\tau}\underline{U}\right).$$

As such, it is even easier to dissuade the agents from playing $\mathcal{A}$. Under the same timely disaster alert policy Γ^{τ} with $\tau \in (0, \hat{\tau})$, all the agents play strategy $\mathcal{W}_o$, and, thus, panic is eliminated even when the regime change is observable.

3.3 Arrival of Additional Information

In our benchmark setup, we assume that the principal has full control of the information after the game begins, and the agents do not get any additional information over time. In practice, this may not always be the case, especially if the time horizon is long. For example, if the time horizon is a month, agents may receive weekly updates about the fundamental θ from other sources.

Suppose that an agent i receives a vector of noisy signals $\boldsymbol{s}_i := (s_i^0, s_i^1 \ldots, s_i^K)$ at specified dates $\{t_0, t_1, \ldots t_K\}$, where $t_0 = 0$ and $t_K < T$. As before, the noise can be arbitrarily correlated, and regardless of $\boldsymbol{s}_i$, an agent i always has some doubt. We assume that the information arrives infrequently; that is,

$$\overline{\tau} := \min_{j=1}^{K+1}\{t_j - t_{j-1}\} > 0,$$

where $t_{K+1} = T$. This implies that whenever the agents receive some information, there is at least a time window of $\overline{\tau}$ before new information arrives or the game ends. $\overline{\tau} > 0$ gives the principal the scope to set a timely disaster alert shortly after each arrival of new information and before the next information arrives or the game ends. We construct an extended Γ^{τ} policy as follows: a timely disaster alert every time after the new information arrives. That is, the k-th disaster alert is triggered — that is, $d^{t_k+\tau} = 1$, if and only if

$R(\theta, N_{t_k+\tau^-}) < 0$, where $k = 0, 1, \ldots K$.

Proposition 2 *If $G(\cdot)$ is atomless, and exogenous information arrives infrequently over time, there exists $\tilde{\tau} > 0$ such that, under the extended disclosure policy Γ^τ with $\tau \in (0, \min\{\tilde{\tau}, \bar{\tau}\})$, the unique rationalizable strategy is to wait for all the alerts and then to follow them. Under this strategy, panic is eliminated in the robust sense; that is, $\Theta^P(\Gamma^\tau) = \emptyset$.*

When the agents get additional information over time, one disaster alert d^τ is not enough to be persuasive. An agent may act on his new information arriving later than the disaster alert and attack at that time. As Proposition 2 states, however, as long as the principal can set a timely disaster alert every time after the arrival of new information, the agents will never act on their own signals. Instead, they always choose to wait for the next disclosure and to follow the disaster alert. We prove this by extending the iterated elimination argument from Theorem 1. We show that the strategy of waiting for all alerts and refraining from attacking unless any alert is triggered strictly dominates any strategy involving waiting for the alerts before time t_k ($k = 0, 1, \ldots, K$), then attacking at time t_k but not waiting for the later alert(s). The formal proof is relegated to the Appendix. Following this result, we can see that even when agents receive additional information about fundamental θ, if the disaster alert is continuously in place, panic will be eliminated.

3.4 Doubt

Throughout this paper, we have maintained the assumption that regardless of their private signals, the agents always have some doubt. Recall that waiting is profitable only when no alert is triggered ($d^\tau = 0$) and the doubt assumption guarantees that $d^\tau = 0$ arises with a positive probability bounded away from 0, regardless of what others do.

Suppose that the agents share a homogeneous belief about the fundamental. If the agents do not doubt and believe that the regime changes with certainty if all others attack, then they may not wait for the disaster alert. Thus, panic may occur. In this sense, doubt is necessary for eliminating panic.

However, under heterogeneous beliefs, agents who do not have doubt may still believe that many others have doubt, and, thus, those agents will wait for the alert (Lemma 3). That makes the endogenous signal $d^\tau = 0$ more likely to arise, thereby increasing the benefit from waiting (infection argument). Thus, a timely disaster alert can be effective even when the doubt assumption is relaxed to some extent.

To illustrate this, we consider a specialized environment commonly used in the global game literature. We assume that the regime change function $R(\theta, N) = \theta - N$. Nature draws θ from the improper prior $\mathcal{U}[-\infty, \infty]$, and agents receive a conditionally independent noisy signal $s_i = \theta + \sigma\varepsilon_i$, where $\varepsilon_i \sim^{iid} N(0, 1)$ and $\sigma > 0$ is a scale parameter. Note that, for any given ϵ, an agent who receives signal $s_i >$

$1+\sigma\Phi^{-1}(\epsilon)$ assigns a probability less than ϵ to $\theta \in \Theta^U$, which violates the doubt assumption.[27]

Proposition 3 *In the above specialized environment,*

1. *Under no disclosure, the unique rationalizable strategy is to attack at $t=0$ for $s_i < \hat{s} = \hat{p}+\sigma\Phi^{-1}(\hat{p})$ and to never attack for $s_i \geq \hat{s}$. Accordingly, the regime survives if and only if $\theta \geq \hat{\theta} = \hat{p}$, where*
$$\hat{p} := \frac{\overline{U} - e^{-\beta T}\underline{V}}{e^{-\beta T}\overline{V} - e^{-\beta T}\underline{V}}.$$

2. *Under disaster alert policy Γ^τ with $\tau \in (0,T)$, the unique rationalizable strategy is to play $\mathcal{A}$ for $s_i < s^*(\tau) = p^*(\tau) + \sigma\Phi^{-1}(p^*(\tau))$ and to play $\mathcal{W}$ for $s_i \geq s^*(\tau)$. Accordingly, the regime survives if and only if $\theta \geq \theta^*(\tau) = p^*(\tau)$, where*
$$p^*(\tau) := \frac{\overline{U} - e^{-\beta\tau}\left(G(\tau^-)\underline{U} + (1-G(\tau^-)\overline{U}\right)}{e^{-\beta T}\overline{V} - e^{-\beta\tau}\left(G(\tau^-)\underline{U} + (1-G(\tau^-)\overline{U}\right)} \in (0,1).$$

3. *Under no disclosure, $\Theta^P = [0,\hat{\theta})$ and, under policy Γ^τ, $\Theta^P(\Gamma^\tau) = [0,\theta^*(\tau))$. The disaster alert policy Γ^τ reduces the probability of panic, that is, $\theta^*(\tau) < \hat{\theta}$ for any $\tau \in (0,T)$.*

4. *For any $\tau > 0$, $\theta^*(\tau)$ decreases as τ decreases; that is, an earlier disaster alert is more effective. Additionally, for any $\zeta > 0$ (however small), there exists $\hat{\tau}(\zeta)$ such that, under the policy Γ^τ with $\tau \in (0,\hat{\tau}(\zeta))$, $\theta^*(\tau) < \zeta$.*

Recall that, under no disclosure, the game boils down to a static regime change game, where the agents choose between "attack" and "not attack." Morris and Shin (2003) use the infection argument and show that for each type s_i, a unique strategy survives the iterated elimination of strictly dominated strategies (IESDS) — that is, attack for $s_i < \hat{s}$ and not attack for $s_i \geq \hat{s}$. This implies that the regime survives if and only if the fundamental is beyond some cutoff $\hat{\theta}$. Given the above specialized environment, this cutoff is $\hat{\theta} = \hat{p}$ (see the Appendix for the formal argument). Notice that $\hat{p}$ is the probability of survival that makes an agent indifferent between attacking and not attacking.

Under a disaster alert policy, the difference is that, instead of "not attack," the agents play strategy $\mathcal{W}$ — that is, they wait but attack if the disaster alert is triggered (See Lemma 1). Clearly, an agent gets a higher payoff from $\mathcal{W}$ than from "not attack" as $\mathcal{W}$ provides the agents with an opportunity to attack later when the regime change is expected to happen. Therefore, when the waiting option is available, agents are less likely to attack the regime at time 0. Accordingly, regardless of disclosure time τ, after IESDS, the

[27] In this specialized environment, $\Theta^L = (-\infty,0)$ and $\Theta^U = [1,\infty)$, and for any signal s_i, $\mathbb{P}(\theta \in \Theta^L|s_i) > 0$ and $\mathbb{P}(\theta \in \Theta^U|s_i) > 0$ (although not bounded away from 0 for all s_i). This is sufficient to guarantee that the only rationalizable strategies are $\mathcal{A}$ and $\mathcal{W}$ (see the proof of Lemma 1).

cutoff $\theta^*(\tau)$ is lower than the fundamental cutoff $\hat{\theta}$ under the static game. As τ falls, the expected payoff from $\mathcal{W}$ further increases, and, accordingly, $\theta^*(\tau)$ falls. Proposition 3 shows that the fundamental cutoff $\theta^*(\tau) \to 0$ and the panic set $\Theta^p(\Gamma^\tau)$ converges to an empty set when τ decreases to 0.

4 Discussion: Deconstructing Disaster Alert

We saw that an EWS, when designed appropriately, can be remarkably effective under a reasonably general setup. This section spells out the essential features of our disaster alert policy and the properties of the underlying environment that account for the effectiveness. We also elaborate on the differences with the existing literature that studies how to manipulate information to foster coordination.

4.1 Endogenous Timing and Future Binary Disclosure

The disaster alert Γ^τ sends a binary public signal $d^\tau \in \{0, 1\}$ at a future date $\tau > 0$, and when $d^\tau = 1$, a regime change is inevitable (the regime will surely change if it has not already). This feature is essential for Lemma 1 (the option value argument): an agent who waits for the alert will not attack after the alert is not triggered. It is necessary for this option value argument that the agents have the option to wait for the alert.[28]

Consider an alternative setup where the agents do not have the option to wait. For instance, in a simultaneous move game, the agents can only choose between whether to attack or not attack. As we have mentioned before, a static binary disclosure that reveals whether the regime will change without further attack cannot eliminate panic in the robust sense. If the agents believe that others will ignore their private information and coordinate based on this publicly disclosed information, then there are no panic-based attacks. However, they may not believe that others will do so, which creates other equilibria with panic-based attacks.

Basak and Zhou (2020) consider a dynamic regime change game where agents move sequentially but in an exogenous order. At a specified date, an agent decides whether or not to attack. However, the agent does not have the option to wait. This is motivated by the example where creditors of a staggered debt portfolio make rollover decisions sequentially. Similar to the static setting, when agents move in an exogenous order, a one-time binary disclosure regarding whether the regime can survive without further attack may not dissuade the agents from attacking. The authors propose frequent disclosure instead of a

[28] Note the inherent asymmetry in our setup: Only one of the actions — namely, attack — is irreversible, and the principal wants to dissuade the agents from taking this irreversible action. The option value argument does not apply if the principal wants to persuade the agents to take the irreversible action (for instance, a dynamic coordination game of investment, as in Dasgupta (2007)). One may conjecture that if the attack were also a reversible action, then the result would be stronger because the agents who do not attack would be reassured by the fact that agents who had attacked could change their minds later. However, they may adopt the following strategy ($\mathcal{W}^c$): attack right away and reverse the action only if the alert is not triggered. The adoption of this strategy could trigger the alert although the regime change was not warranted, and, thus, panic is not eliminated.

onetime disclosure. They build an inductive argument that shows that each disclosure is persuasive if the subsequent ones are persuasive for agents moving later.[29]

Basak and Zhou (2020) demonstrate the value of frequent disclosure under exogenous timing and explain the (informational) benefit from a debt profile with diversified maturity dates. Unlike Basak and Zhou (2020), this paper speaks to the effectiveness of dynamic disclosure (such as supervisory bank stress tests and debt sustainability analysis) in very different applications where agents can choose the timing of their attack. The main difference between the underlying mechanisms in these two papers is as follows. When an agent learns that the regime can survive without further attack, he may still attack because he faces strategic uncertainty regarding what others will do. When the agents move in exogenous order, the principal must assure an agent that she will also persuade those moving later; this necessitates frequent disclosure. In contrast, when agents can choose the timing of the attack, as Lemma 2 shows, a one-time alert perfectly predicts the regime outcome. That is, there is no strategic uncertainty regarding all those agents who have waited. Therefore, frequent disclosure is not required, but the timing of disclosure matters as it controls the cost of waiting.

4.2 Endogenous Disclosure

The option value argument shows that the agents who wait for the alert will follow. However, it does not mean the agents will wait for the alert. The alert must provide valuable information so as to persuade the agents to wait for it. In Lemma 2, we show that regardless of whether an agent believes that others will wait or attack, a disaster alert is always valuable in the sense that it can perfectly predict the regime outcome. An essential feature for this perfect predictability of the disaster alert is endogenous disclosure; that is, the principal sends different messages based on the number of agents who endogenously choose to wait for the alert.

Note that when the agents move simultaneously, any disclosure policy (see, for instance, Goldstein and Huang (2016), Inostroza and Pavan (2020), Morris, Oyama and Takahashi (2020), and Li, Song and Zhao (2019)) is solely based on the fundamental and, thus, is not an endogenous disclosure. However, in a dynamic setting with multiple agents, the message disclosed can be based on historical endogenous actions.

To see why endogenous disclosure is an essential property, consider an alternative policy $\tilde{\Gamma}^{\tau}$, which is the same as Γ^{τ} except that the alert is triggered if and only if $R(\theta, 0) < 0$ (or, equivalently, $\theta \in \Theta^{L}$).

[29] Basak and Zhou (2020) consider a more restrictive environment in which the agents are endowed with conditionally independent signals (as in the canonical global games) and are restricted to playing monotone strategies. We do not impose any such restrictions in the present paper.

Formally,

$$\tilde{d}^{\tau}(\theta, N_{\tau^-}) = \begin{cases} 1 \text{ if } R(\theta, 0) < 0; \\ 0 \text{ otherwise.} \end{cases}$$

Under this policy, the principal sends a binary future signal, and the message $\tilde{d}^{\tau} = 1$ predicts the regime change regardless of the agents' choices. Therefore, whoever waits for the future binary disclosure $\tilde{d}^{\tau}$ would follow it (Lemma 1 holds true). However, the message $\tilde{d}^{\tau} = 0$, which means "the regime will survive if no one attacks it," cannot predict the survival of the regime. If some agents do not wait for the future disclosure but choose to attack (choose $\mathcal{A}$ over $\mathcal{W}$), then following $\tilde{d}^{\tau} = 0$, the regime can still change even without any further attack. As such, the binary signal $\tilde{d}^{\tau}$ cannot predict the regime outcome (Lemma 2 fails), and, therefore, agents may not find the alert $\tilde{d}^{\tau}$ worth waiting for.[30]

The crucial difference between the above policy $\tilde{\Gamma}^{\tau}$ and a disaster alert Γ^{τ} is that $d^{\tau} = 0$ means that "the regime will survive if there is no attack *after the disclosure time* τ." Unlike $\tilde{d}^{\tau} = 0$, it predicts no regime change as long as the agents who wait for the disclosure will follow it, regardless of the size of attack N_{τ^-} before the disclosure.[31] This makes the disaster alert d^{τ} more valuable than $\tilde{d}^{\tau}$ under the alternative disclosure policy $\tilde{\Gamma}^{\tau}$.

Interestingly, under a timely disaster alert Γ^{τ}, as the agents play the unique rationalizable strategy $\mathcal{W}$, on path, at the disclosure time τ, $d^{\tau} = 0$ shares the same meaning regarding the fundamental state θ as $\tilde{d}^{\tau}$ since no one attacks before time τ (i.e., $N_{\tau^-} = 0$). However, as the above discussion shows, the history-dependent design has a significant impact on whether agents would wait for future disclosure. Since agents may not choose to wait for $\tilde{d}^{\tau}$ at $t = 0$, the policy $\tilde{\Gamma}^{\tau}$ cannot be robust in eliminating panic.

4.3 Forecasting

However, the history-dependent design, by itself, is not enough to guarantee the perfect predictability of the disaster alert d^{τ}. Under our disaster alert, even if the shock has not arrived by the time of disclosure, if it is evident that the regime will not survive when the shock arrives, the principal triggers the alert (sends message $d^{\tau} = 1$). We call this feature of disaster alert *forecasting*. This forecasting property is also

[30] To see this formally, suppose that an agent believes that the regime will survive with probability $\mathbb{P}(t_c = \infty | x_{-i}, s_i, \tilde{d}^{\tau} = 0) = p$. Then, the benefit of waiting is

$$\mathbb{B}(\tilde{\Gamma}^{\tau}, x_{-i}, s_i) := \mathbb{E}\left[\pi(\mathbb{T}, t_c) - \pi(0, t_c) \middle| x_{-i}, s_i, \tilde{d}^{\tau} = 0\right] = e^{-\beta T}[(1-p)\underline{V} + p\overline{V}] - \overline{U}.$$

If x_{-i} is such that most of the other agents choose $\mathcal{A}$ over $\mathcal{W}$, then p is sufficiently small. Accordingly, the above benefit may not be positive. This means that there can be an equilibrium in which the agents panic and attack rather than wait for the disclosure $\tilde{d}^{\tau}$. In contrast, under policy Γ^{τ}, $P(t_c = \infty | x_{-i}, s_i, d^{\tau} = 0) = 1$ (Lemma 2). Thus, regardless of whether others choose $\mathcal{A}$ or $\mathcal{W}$, waiting for d^{τ} always has a strictly positive benefit.

[31] Formally, $d^{\tau} = 0$ (i.e., $R(\theta, N_{\tau^-}) \geq 0$) implies $\tilde{d}^{\tau} = 0$ (i.e., $R(\theta, 0) \geq 0$) since $R(\theta, N)$ is decreasing in N. However, it is possible to have $d^{\tau} = 1$ and $\tilde{d}^{\tau} = 0$. Consider $\theta \notin \Theta^L \cup \Theta^U$. Then, $\tilde{d}^{\tau} = 0$ since $R(\theta, 0) \geq 0$. In that case, if sufficiently many agents attack before time τ such that $R(\theta, N_{\tau^-}) < 0$, then $d^{\tau} = 1$.

essential for perfect predictability (Lemma 2).

To see this, consider this alternative policy $\hat{\Gamma}^{\tau}$, which is the same as Γ^{τ} except that the alert is triggered if and only if the regime change has occurred before time τ; that is, $t_c < \tau$. Formally,[32]

$$\hat{d}^{\tau}(\theta, N_{\tau^-}, \mathbb{1}\{t_s < \tau\}) = \begin{cases} 1 \text{ if } R(\theta, N_{\tau^-}) < 0 \ \& \ \mathbb{1}\{t_s < \tau\} = 1; \\ 0 \text{ otherwise.} \end{cases}$$

Unlike under Γ^{τ} policy, when the regime is doomed to change but has not yet changed ($R(\theta, N_{\tau^-}) < 0$ and $\mathbb{1}\{t_s < \tau\} = 0$), the principal does not trigger the alert. Note that, since $\hat{d}^{\tau}$ is a future binary signal with $\hat{d}^{\tau} = 1$ predicting the regime change, Lemma 1 holds true under $\hat{\Gamma}^{\tau}$. However, the agents could see "no alert" ($\hat{d}^{\tau} = 0$) just because the shock has not arrived yet (i.e., $t_s \geq \tau$), but the regime will change once the shock arrives later. Thus, $\hat{d}^{\tau} = 0$ does not guarantee no future regime change even without any further attack (Lemma 2 fails). Therefore, forecasting is an essential feature of our disaster alert.

Notice that the disclosure policy $\hat{\Gamma}^{\tau}$ only reveals whether or not the regime has already changed. The literature often considers this policy as naturally available information rather than a deliberate information design policy. For instance, in Angeletos, Hellwig and Pavan (2007) and Basak and Zhou (2020), agents may move at different dates, and when they move, they know whether or not the regime has already changed. The above-mentioned studies assume that the shock has already arrived. However, if the shock arrives at a future date (as in our current setup), then as we have argued, this alternative disclosure policy $\hat{\Gamma}^{\tau}$ fails Lemma 2. Consequently, the agents may not want to wait for such an alert. This is true even if the principal uses frequent disclosure. Thus, the frequent disclosure as proposed in Basak and Zhou (2020) fails to eliminate panic in our present setup.

Forecasting Errors

In practice, forecasting requires due diligence.[33] However, it may be difficult, if not impossible, to learn about θ before the shock arrives, making forecasting infeasible or at least erroneous. A false negative happens when the alert is triggered ($d^{\tau} = 1$) but $R(\theta, N_{\tau^-}) \geq 0$; a false positive happens when the alert is not triggered ($d^{\tau} = 0$) but $R(\theta, N_{\tau^-}) < 0$. A forecasting error makes the timely disaster alert less

[32] Notice that the principal's message is also conditional on whether the shock has already arrived. Abusing notation, we incorporate the indicator $\mathbb{1}\{t_s < \tau\}$ into the argument of d^{τ}.

[33] For forward-looking bank stress tests, bank supervisors collect data from the individual banks and, based on their models, investigate how the banks' balance sheets perform under some projected future stress scenarios. For debt sustainability analysis, IMF and the World Bank use historical data regarding a country's debt burdens and its macroeconomic conditions; perform a forward-looking analysis of the country's fiscal standing under different scenarios with plausible future shocks; conduct stress tests to evaluate country-specific risks stemming from future potential shocks (e.g., natural disasters, volatile commodity prices); and rely on quantitative models to identify the vulnerabilities in its public debts. For the details of the DSA program run by the IMF, see International Monetary Fund, "The Debt Sustainability Framework for Low-Income Countries," July 13, 2018, https://www.imf.org/external/pubs/ft/dsa/lic.htm. For econometric methods adopted in the development of an EWS for sovereign debt crises, see Dawood, Horsewood and Strobel (2017).

valuable. A false negative can be disastrous since the agents will all attack following the alert, although regime change is not warranted (i.e., $\theta \notin \Theta^L$). Therefore, the principal should apply very stringent rules in generating the warning — sending the message of "no warning" unless the crisis is certain. However, this action could generate false positives more often. Although a false positive does not entail the same risk, if the probability of a false positive is high, the agents may ignore the disclosure and act based on their own signals. Accordingly, panic could ensue. The following proposition shows that a timely disaster alert is effective as long as there is no false negative and the probability of a false positive is sufficiently small.

Proposition 4 *Suppose that there is no false negative and that a false positive happens with probability, at most, $\eta > 0$. If $\eta < \eta_0 \equiv \frac{\overline{V}-e^{\beta T}\overline{U}}{\overline{V}-\underline{V}}$, then there exists $\hat{\tau}(\eta) > 0$ such that under Γ^τ with $\tau \in (0, \hat{\tau}(\eta))$, $\Theta^P(\Gamma^\tau) = \emptyset$.*

The intuition behind this result is simple. Since there is no false negative, the agents will attack immediately after learning $d^\tau = 1$. Therefore, the option value argument (Lemma 1) still holds true. Because it is possible to have a false positive, $d^\tau = 0$ cannot be a perfect predictor of the survival of the regime, thereby reducing the benefit of waiting. However, if the principal can set the disaster alert sufficiently early to save the cost of waiting, then agents will still wait for the future disclosure provided that the incidence of a false positive is small enough.

4.4 Principal's preference and commitment

Our principal has a simple preference: she only cares about the ultimate regime outcome and prefers that the regime survives.[34] Under such preference, ex ante commitment is not necessary for our optimal disclosure policy. To see this, at the disclosure time τ, if $R(\theta, N_{\tau^-}) < 0$, then the regime change is inevitable regardless of the agents' actions after time τ. Therefore, the principal has no incentive to misreport or delay the disclosure ex post. Moreover, from an ex ante perspective, policy Γ^τ is able to completely eliminate the panic and obtain the most preferred outcome for the principal. Therefore, she has no incentive to deviate to any other disclosure rules (e.g., send other messages before time τ). However, if the principal has more nuanced preferences, she may need commitment power to implement an EWS or may not want an EWS in the first place.

Recall that under the proposed optimal disclosure policy, when $\theta \in \Theta^L$, regardless of what the agents do, the regime changes as soon as the shock arrives; and when $\theta \notin \Theta^L$, all agents coordinate on not attacking and the regime never changes. It is easy to see that the timely disaster alert policy remains optimal if the principal derives a positive flow utility for every moment that the regime survives, or if the principal not only cares about the survival of the regime but also gets a higher payoff when the size

[34] This preference specification is consistent with that of international organizations trying to reduce the panic associated with sovereign debt defaults (Example 1) or of bank supervisors and regulators trying to stop runs on healthy banks (Example 2).

of the attack on a fundamentally sound regime ($\theta \notin \Theta^L$) is smaller. Moreover, if the principal wants to maximize the welfare of the agents, then she should schedule the disclosure time τ as early as possible (but still after time 0) because the agents bear an unnecessary cost while waiting for the alert.[35] However, if the principal's objective is in conflict with that of the agents — if, for example, she wants to minimize the size of the attack or delay the attack as much as possible even when regime change is inevitable — then the timely disaster alert is not a desirable policy. This is because, under this policy, when the regime change is inevitable, all the agents attack and do so immediately after the alert is triggered. For such problems, the regulator should use private messages to generate miscoordination among agents.[36]

5 Conclusion

We extend the canonical regime change game to a dynamic setup in which privately informed agents choose when to attack (if at all) and the first mover gets an advantage. We study an adversarial information design in this dynamic game under a fairly flexible exogenous information structure. We construct a dynamic disclosure policy that publicly sends a binary message (alert or no alert) at a future date but in a timely manner. We show that, under this simple policy, the unique rationalizable strategy is to wait for and then follow the alert, regardless of agents' private information. This timely disaster alert is proved optimal as, under the unique rationalizable strategy profile, it perfectly coordinates agents on taking the right course of action and achieves the principal's most preferred outcome. The main contribution of the paper is to show that, unlike the case under static persuasion, when a timing dimension is added to the coordination problem, the principal can eliminate the undesirable outcomes that can arise from agents' failure to coordinate on the right course of action.

Relating our optimal policy to the reality, we show that the proposed policy resembles an EWS, such as debt sustainability analysis or forward-looking bank stress tests. These information disclosure policies caution agents about the impending crisis. This paper provides a rationale for adopting such a policy by uncovering the underlying mechanism that makes an EWS effective at preventing panic. Our theory demonstrates the optimality of a binary pass/fail design and emphasizes that "endogenous disclosure," "forecasting," and "timely disclosure" are necessary features of an effective EWS.[37]

[35] The principal can also supplement the disclosure policy Γ^τ with an additional public disaster alert at time 0, which is purely fundamental based and reveals whether or not $\theta \in \Theta^L$. With this additional disclosure, if $\theta \in \Theta^L$, agents attack at time 0; otherwise, they wait for the disclosure at time τ and never attack after observing $d^\tau = 0$. However, strict dominance does not work if the agents know $\theta \notin \Theta^L$. Nevertheless, we can use PBE as our solution concept and show that, under this policy, all agents will wait and follow d^τ if $\theta \notin \Theta^L$ and attack at time 0 if $\theta \in \Theta^L$ (see the Online Appendix).

[36] Optimal private disclosure in such environment is an open question and beyond the scope of this paper. See Ely (2017) for an interesting example with two agents and a bank that wants to maximize the time when the last depositor runs.

[37] Anderson (2016) provides some anecdotal evidence demonstrating the benefit of promptly conducting bank stress tests. The author shows that forward-looking stress tests were quite effective in the U.S. but not as effective in Europe. The author argues that this contrasting experience can be attributed to the fact that the stress test attempt in the U.S. was "timely," whereas it was perhaps "too late" in Europe.

References

Abreu, Dilip, and Markus K Brunnermeier. 2003. "Bubbles and crashes." *Econometrica*, 71(1): 173–204.

Anderson, R. W. 2016. "Stress testing and macroprudential regulation: A transatlantic assessment." *CEPR Press*.

Angeletos, George-Marios, and Chen Lian. 2017. "Incomplete Information in Macroeconomics: Accommodating Frictions in Coordination." *Handbook of Macroeconomics*, 2: 1065–1240.

Angeletos, George-Marios, Christian Hellwig, and Alessandro Pavan. 2007. "Dynamic Global Games of Regime Change: Learning, Multiplicity, and the Timing of Attacks." *Econometrica*, 75(3): 711–756.

Au, Pak Hung. 2015. "Dynamic information disclosure." *The RAND Journal of Economics*, 46(4): 791–823.

Ball, Ian. 2019. "Dynamic Information Provision: Rewarding the Past and Guiding the Future." *Available at SSRN 3103127*.

Basak, Deepal, and Zhen Zhou. 2020. "Diffusing Coordination Risk." *American Economic Review*, 110.1: 271–297.

Ben-Porath, Elchanan, and Eddie Dekel. 1992. "Signaling Future Actions and Potential for Sacrifice." *Journal of Economic Theory*, 57: 36–51.

Bergemann, Dirk, and Stephen Morris. 2016. "Information Design, Bayesian Persuasion, and Bayes Correlated Equilibrium." *The American Economic Review*, 106(5): 586–591.

Bergemann, Dirk, and Stephen Morris. 2019. "Information design: A unified perspective." *Journal of Economic Literature*, 57(1): 44–95.

Chamley, Christophe. 1999. "Coordinating Regime Switches." *Quarterly Journal of Economics*, 869–905.

Chamley, Christophe. 2003. "Dynamic speculative attacks." *American Economic Review*, 93(3): 603–621.

Chamley, Christophe, and Douglas Gale. 1994. "Information revelation and strategic delay in a model of investment." *Econometrica*, 62(5): 1065–1085.

Chassang, Sylvain. 2010. "Fear of Miscoordination and the Robustness of Cooperation in Dynamic Global Games with Exit." *Econometrica*, 78(3): 973–1006.

Dasgupta, Amil. 2007. "Coordination and Delay in Global Games." *Journal of Economic Theory*, 134(1): 195–225.

Dasgupta, Amil, Jakub Steiner, and Colin Stewart. 2012. "Dynamic Coordination with Individual Learning." *Games and Economic Behavior*, 74(1): 83–101.

Dawood, Mary, Nicholas Horsewood, and Frank Strobel. 2017. "Predicting sovereign debt crises: an early warning system approach." *Journal of Financial Stability*, 28: 16–28.

Ely, Jeffrey C. 2017. "Beeps." *The American Economic Review*, 107(1): 31–53.

Ely, Jeffrey C, and Martin Szydlowski. 2020. "Moving the goalposts." *Journal of Political Economy*, 128(2): 468–506.

Ely, Jeffrey, George Georgiadis, Sina Moghadas Khorasani, and Luis Rayo. 2021. "Optimal feedback in contests." *Working Paper*.

Goldstein, Itay, and Chong Huang. 2016. "Bayesian Persuasion in Coordination Games." *American Economic Review*, 106(5): 592–596.

Goldstein, Itay, and Yaron Leitner. 2018. "Stress tests and information disclosure." *Journal of Economic Theory*, 177: 34–69.

Gul, Faruk, and Russell Lundholm. 1995. "Endogenous timing and the clustering of agents' decisions." *Journal of Political Economy*, 103(5): 1039–1066.

Inostroza, Nicolas. 2019. "Persuading multiple audiences: An information design approach to banking regulation." *Available at SSRN 3450981*.

Inostroza, Nicolas, and Alessandro Pavan. 2020. "Persuasion in Global Games with Application to Stress Testing." *Working Paper*.

Kamenica, Emir. 2019. "Bayesian persuasion and information design." *Annual Review of Economics*, 11: 249–272.

Kamenica, Emir, and Matthew Gentzkow. 2011. "Bayesian Persuasion." *The American Economic Review*, 101(6): 2590–2615.

Kohlberg, Elon, and Jean-Francois Mertens. 1986. "On the Strategic Stability of Equilibria." *Econometrica*, 54(5): 1003–1037.

Li, Fei, Yangbo Song, and Mofei Zhao. 2019. "Global Manipulation by Local Obfuscation: Information Design in Coordination Games." *Available at SSRN 3471491*.

Li, Xuelin, Martin Szydlowski, and Fangyuan Yu. 2021. "Hype Cycles: Dynamic Information Design with Two Audiences." *Available at SSRN 3923908*.

Mathevet, Laurent, Jacopo Perego, and Ina Taneva. 2020. "On information design in games." *Journal of Political Economy*, 128(4): 1370–1404.

Morris, Stephen, and Hyun Song Shin. 2003. "Global Games: Theory and Applications." In *Advances in Economics and Econometrics (Proceeding of the Eighth World Congress of the Econometric Society)*. , ed. Dewatripont, Hansen and Turnovsky. Cambridge University Press.

Morris, Stephen, Daisuke Oyama, and Satoru Takahashi. 2020. "Implementation via information design in binary-action supermodular games." *Available at SSRN 3697335*.

Orlov, Dmitry, Andrzej Skrzypacz, and Pavel Zryumov. 2020. "Persuading the Principal To Wait." *Journal of Political Economy*, 128(7): 2542–2578.

Orlov, Dmitry, Pavel Zryumov, and Andrzej Skrzypacz. 2018. "Design of macro-prudential stress tests." *Working Paper*.

Sakovics, Jozsef, and Jakub Steiner. 2012. "Who Matters in Coordination Problems?" *The American Economic Review*, 102(7): 3439–3461.

Smolin, Alex. 2021. "Dynamic Evaluation Design." *American Economic Journal: Microeconomics*, 13(4): 300–331.

Appendix

Proof of Lemma 1. We prove that the undominated strategy for any type s_i is $\mathcal{A}$ and $\mathcal{W}$ in three steps.

Step 1. $(0, t_0, t_1) \succ (t_\emptyset, t_0, t_1)$ for all $t_\emptyset \in (0, \tau)$, $t_0, t_1 \in [\tau, T] \cup \{\mathbb{T}\}$.

Consider any agent i with any signal s_i, and take any strategy profile x_{-i} as given. For any $t_\emptyset \in [0, \tau)$,

$$u_i((t_\emptyset, t_0, t_1), x_{-i}) = e^{-\beta t_\emptyset} \left[\mathbb{P}(t_\emptyset < t_c | x_{-i}, s_i) \overline{U} + (1 - \mathbb{P}(t_\emptyset < t_c | x_{-i}, s_i)) \underline{U} \right].$$

For any given s_i and x_{-i}, the distribution of $t_c(s_i, x_{-i})$ is fixed and, therefore, $\mathbb{P}(t_\emptyset < t_c | x_{-i}, s_i)$ (weakly) decreases with $t_\emptyset$. Because $\overline{U} > \underline{U} > 0$, and $e^{-\beta t_\emptyset}$ strictly decreases with $t_\emptyset$, the expected payoff $\mathbb{E}[u_i((t_\emptyset, t_0, t_1), x_{-i}) | s_i]$ strictly decreases with $t_\emptyset$ for $t_\emptyset \in [0, \tau)$, thereby proving the strict dominance of $(0, t_0, t_1)$.

Step 2. $(t_\emptyset, t_0, \tau) \succ (t_\emptyset, t_0, t_1)$ for all $t_\emptyset, t_0 \in [\tau, T] \cup \{\mathbb{T}\}$, and $t_1 \in (\tau, T] \cup \{\mathbb{T}\}$.

Fix any agent i, consider any signal s_i and take any strategy profile x_{-i} as given, for any $t_\emptyset \in [\tau. T] \cup \{\mathbb{T}\}$ and $t_0, t_1 \in [\tau, T] \cup \{\mathbb{T}\}$,

$$\begin{aligned} u_i((t_\emptyset, t_0, t_1), x_{-i}) \; &= \mathbb{P}(R(\theta, N_{\tau^-}(x_{-i})) \geq 0 | s_i) \mathbb{E}[\pi_i(t_0, t_c) | d^\tau = 0, x_{-i}, s_i] \\ &+ \quad \mathbb{P}(R(\theta, N_{\tau^-}(x_{-i})) < 0 | s_i) \mathbb{E}[\pi_i(t_1, t_c) | d^\tau = 1, x_{-i}, s_i]. \end{aligned} \tag{A.1}$$

Note that the first term on the RHS is independent of t_1. Let us consider the expected payoff conditional on $d^\tau = 1$ in the second term on the RHS. For $t_1 \in [\tau, T]$,

$$\mathbb{E}[\pi_i(t_1, t_c) | d^\tau = 1, x_{-i}, s_i] = e^{-\beta t_1} \left[\underline{U} + \mathbb{P}(t_1 < t_c | d^\tau = 1, x_{-i}, s_i)(\overline{U} - \underline{U}) \right]. \tag{A.2}$$

Observe that $\mathbb{P}(t_1 < t_c | d^\tau = 1, x_{-i}, s_i)$ weakly decreases with t_1. Since $\overline{U} > \underline{U} > 0$ and $\beta > 0$, we have $\mathbb{E}[\pi_i(\tau, t_c) | d^\tau = 1, x_{-i}, s_i] > \mathbb{E}[\pi_i(t_1, t_c) | d^\tau = 1, x_{-i}, s_i]$ for any $t_1 \in (\tau, T]$.

Next, consider $t_1 = \mathbb{T}$. By definition of (1), $t_c \leq T^+ \leq \mathbb{T}$ given that $R(\theta, N_{\tau^-}) < 0$. This implies that, regardless of x_{-i} and s_i, $\mathbb{E}[\pi_i(\mathbb{T}, t_c) | d^\tau = 1, x_{-i}, s_i] = e^{-\beta T} \underline{V}$. Since $\underline{U} > \underline{V}$ and $\beta > 0$, we have (see (A.2))

$$\mathbb{E}[\pi_i(\tau, t_c) | d^\tau = 1, x_{-i}, s_i] \geq e^{-\beta \tau} \underline{U} > e^{-\beta T} \underline{U} > e^{-\beta T} \underline{V} = \mathbb{E}[\pi_i(\mathbb{T}, t_c) | d^\tau = 1, x_{-i}, s_i].$$

Therefore, the following inequality holds for any $t_1 \neq \tau$ regardless of s_i and x_{-i},

$$\mathbb{E}[\pi_i(\tau, t_c) | d^\tau = 1, x_{-i}, s_i] > \mathbb{E}[\pi_i(t_1, t_c) | d^\tau = 1, x_{-i}, s_i].$$

Under Assumption 3, for any x_{-i}, s_i, the probability

$$\mathbb{P}(R(\theta, N_{\tau^-}(x_{-i})) < 0 | s_i) \geq \mathbb{P}(R(\theta, 0) < 0 | s_i) = \mathbb{P}(\theta \in \Theta^L | s_i) > 0.$$

This implies that $\mathbb{E}[u_i((t_\emptyset, t_0, \tau), x_{-i})|s_i] > \mathbb{E}[u_i((t_\emptyset, t_0, t_1), x_{-i})|s_i]$ for all $t_\emptyset \in [\tau, T] \cup \{\mathbb{T}\}$, $t_0 \in [\tau, T] \cup \{\mathbb{T}\}$, $t_1 \in (\tau, T] \cup \{\mathbb{T}\}$.

Step 3. $(0, t_0', t_1') \succ (t_\emptyset, t_0, \tau)$ for any $t_0', t_1' \in [\tau, T] \cup \{\mathbb{T}\}$ and $t_\emptyset \in [\tau, T] \cup \{\mathbb{T}\}, t_0 \in [\tau, T]$.

First, let us compare the expected payoffs from these two strategies for any fixed s_i and x_{-i}. Under $(t_\emptyset = 0, t_0, t_1)$, we have $u_i((t_\emptyset = 0, t_0', t_1'), x_{-i}) = \overline{U}$. Then, observe that, since $\overline{U} > \underline{U} > 0$, the expected payoff from $(t_\emptyset, t_0, \tau)$ with $t_\emptyset \geq \tau$ and $t_0 \in [\tau, T]$ satisfies

$$u_i((t_\emptyset, t_0, \tau), x_{-i}) \leq e^{-\beta t_0}\mathbb{P}(d^\tau = 0|s_i, x_{-i})\overline{U} + e^{-\beta\tau}\mathbb{P}(d^\tau = 1|s_i, x_{-i})\overline{U} \leq e^{-\beta\tau}\overline{U}.$$

Given that $\beta > 0$ and $\overline{U} > 0$, $\mathbb{E}[u_i((t_\emptyset = 0, t_0', t_1'), x_{-i})|s_i] > \mathbb{E}[u_i((t_\emptyset, t_0, \tau), x_{-i})|s_i]$.[38]

Therefore, for agent with any type s_i, for any strategy satisfying $t_\emptyset \in [0, \tau)$, the only undominated strategy is $\mathcal{A} = (0, t_0, t_1)$ where $t_0, t_1 \in [\tau, T] \cup \{\mathbb{T}\}$; and for any strategy satisfying $t_\emptyset \in [\tau, T] \cup \{\mathbb{T}\}$, the only undominated strategy is $\mathcal{W} = (t_\emptyset, \mathbb{T}, \tau)$, where $t_\emptyset \in [\tau, T] \cup \{\mathbb{T}\}$.

It is worth noting that there is only one round of elimination, as we do not require agents to hold the belief that other agents are not playing dominated strategies. Further, regarding the the belief about fundamental θ, this elimination only requires $\mathbb{P}(\theta \in \Theta^L|s_i) > 0$ for any s_i in Step 2, which obviously holds true under Assumption 3. ■

Proof of Lemma 2. Given that the only rationalizable strategies are $\mathcal{A}$ and $\mathcal{W}$ for any type s_i, conditional on $d^\tau = 0$, $N_T = N_{\tau^-}$ as $t_0 = \mathbb{T}$ under strategy $\mathcal{W}$. Therefore, by definition, $t_c = \infty$ as $R(\theta, N_T) = R(\theta, N_{\tau^-}) \geq 0$. Moreover, under either strategy $\mathcal{A}$ or $\mathcal{W}$, $N_{\tau^-} = N_0$ since $t_\emptyset \notin (0, \tau)$. Therefore, conditional on $d^\tau = 1$, $R(\theta, N_0) = R(\theta, N_{\tau^-}) < 0$ which implies that $t_c = t_s^+$ (see the definition of t_c in (1)). ■

Proof of Lemma 3. The proof follows the discussion after Lemma 3 directly. To complete the proof, here, we set out to find $\hat{\tau}$ explicitly. First, from (5), the expected cost $\mathbb{C}(\Gamma^\tau, x_{-i}, s_i)$ satisfies the following inequality:

$$\mathbb{C}(\Gamma^\tau, x_{-i}, s_i) = G(\tau^-)e^{-\beta\tau}(\overline{U} - \underline{U}) + (1 - e^{-\beta\tau})\overline{U} < G(\tau^-)(\overline{U} - \underline{U}) + \beta\tau\overline{U}.$$

From (6), the expected payoff difference is at least

$$\mathbb{D}(\Gamma^\tau, x_{-i}, s_i) > -(1 - \epsilon)\left[G(\tau^-)(\overline{U} - \underline{U}) + \beta\tau\overline{U}\right] + \epsilon(e^{-\beta T}\overline{V} - \overline{U}).$$

[38] Note that this step does not require the agent to hold the belief that $\mathbb{P}(d^\tau = 0|s_i, x_{-i}) > 0$. Even when the agent believes that this probability is zero, he would attack immediately if anticipating that $d^\tau = 1$ will arise with certainty.

Note that the payoff differences are independent of s_i and x_{-i}. Let us define

$$\hat{\tau} := \min\left\{\frac{\epsilon}{2(1-\epsilon)}\frac{e^{-\beta T}\overline{V}-\overline{U}}{\beta\overline{U}}, G^{-1}\left(\frac{\epsilon}{2(1-\epsilon)}\frac{e^{-\beta T}\bar{V}-\overline{U}}{\overline{U}-\underline{U}}\right)\right\}.$$

As $G(0)=0$, $\hat{\tau}>0$ for any $\epsilon>0$. It is easy to check that, for any given $\tau\in(0,\hat{\tau})$, $\mathbb{D}(\Gamma^\tau, s_i)>0$. Thus, under Γ^τ with $\tau\in(0,\hat{\tau})$, $\mathcal{A}$ is strictly dominated by $\mathcal{W}$. ■

Proof of Theorem 1. By Lemma 3, under policy Γ^τ with $\tau\in(0,\hat{\tau})$, for any agent $i\in[0,1]$ with any type $s_i\in\mathbb{S}$, the only rationalizable strategy is $\mathcal{W}$. Therefore, $N_{\tau^-}=0$. Consider any $\theta\notin\Theta^L$. By the definition of Θ^L, $R(\theta, N_{\tau^-}=0)\geq 0$ and, thus, $d^\tau=0$. By Lemma 2, conditional on $d^\tau=0$, $t_c=\infty$. Hence, $\Theta^C(\Gamma^\tau)=\Theta^L$ and $\Theta^P(\Gamma^\tau)=\emptyset$. ■

Proof of Proposition 1. Since it takes an agent $l>0$ time to finish the attack, at any date t, there is a difference between the pledged attack N_t and the materialized attack M_t, where $M_t := \int_{x\leq t} B\left(\frac{t-x}{l}\right)\mathrm{d}N_x$. By definition, $M_t\leq N_t$ and $M_{T+l}=N_T$. The regime changes at the first instance when $R(\theta, M_t)<0$, and if $R(\theta, M_{T+l})\geq 0$, then the regime never changes. Accordingly, we redefine t_c as follows.

$$t_c := \begin{cases}\min\{t\in[t_s^+,\infty)|R(\theta,M_t)<0\} \text{ if } R(\theta,M_{T+l})<0\\ \infty \text{ if } R(\theta,M_{T+l})\geq 0.\end{cases} \tag{A.3}$$

The agent, who attacks at time t, gets a flow payoff from t until the attack is finished, $t+l$. Therefore, the payoff from starting an attack at time $t\in[0,T]$ can be written as

$$\pi(t,t_c)=e^{-\beta t}\int_0^l e^{-\beta z}\left[\mathbb{1}\{t+z\geq t_c\}\underline{u}+\mathbb{1}\{t+z<t_c\}\bar{u}\right]\mathrm{d}B(\frac{z}{l}). \tag{A.4}$$

To make the payoffs comparable, we write the payoff from not attacking (i.e., $t=\mathbb{T}$) as

$$\pi(\mathbb{T},t_c)=e^{-\beta T}\int_0^l e^{-\beta z}\left[\mathbb{1}\{t_c<\infty\}\underline{v}+\mathbb{1}\{t_c=\infty\}\overline{v}\right]\mathrm{d}B(\frac{z}{l}).$$

Complementarity (Assumption 2) dictates that the flow payoffs satisfy $e^{-\beta T}\overline{v}>\overline{u}>\underline{u}>\underline{v}\geq 0$.[39]

By definition, under the disaster alert policy Γ^τ with $\tau\in(0,l)$, the alert is triggered ($d^\tau=1$) if θ and the pledged attack N_{τ^-} are such that $R(\theta, N_{\tau^-})<0$. As in our benchmark setup, when $d^\tau=1$, the regime changes ($t_c<\infty$) regardless of what the agents do afterwards. Moreover, for any $t'>t$ where $t,t'\in[0,T]$, and any t_c, $\pi(t',t_c)<\pi(t,t_c)$. Thus, delayed attack is always costly. Therefore, Lemma 1 and 2 hold true. However, note that, unlike in the benchmark model, we do not have $t_c=t_s^+$. Although

[39] To connect this augmented model to our benchmark setup, one can interpret the instantaneous payoffs $\overline{U}$ and $\underline{U}$ in our benchmark setup as $\overline{Y}=\overline{y}\cdot\left(\int_0^l e^{-\beta z}\mathrm{d}B(\frac{z}{l})\right)$ and $\underline{Y}=\underline{y}\cdot\left(\int_0^l e^{-\beta z}\mathrm{d}B(\frac{z}{l})\right)$ where Y stands for U,V and y stands for u and v. It is worth noting that when the time lag l decreases to 0, all flow payoffs (in lowercase letters) converge to the original payoffs (in uppercase letters), and the materialized proportion of the attack is no different from the pledged attack —i.e., $M_t\to N_t$. This brings us back to the benchmark model, where attacks are instantaneous ($l=0$).

the agents may start attacking at time 0, the regime may not change as soon as the shock arrives because all the attacks may not have materialized by then.

It follows from Lemma 2 that conditional on $d^\tau = 0$, $t_c = \infty$. Therefore, the benefit of playing $\mathcal{W}$ over $\mathcal{A}$ is

$$\mathbb{B}^S(\mathit{\Gamma}^\tau, x_{-i}, s_i) := \mathbb{E}[\pi(\mathbb{T}, t_c) - \pi(0, t_c)|d^\tau = 0, x_{-i}, s_i] = \left(e^{-\beta T}\bar{v} - \bar{u}\right)\int_0^l e^{-\beta z}\mathrm{d}B(\frac{z}{l}) \tag{A.5}$$

Note that $\mathbb{B}^S(\mathit{\Gamma}^\tau, x_{-i}, s_i) > 0$ and it is independent of τ, whether others choose $\mathcal{A}$ or $\mathcal{W}$, and s_i.

Consider an agent who starts to attack at t. Conditional on $d^\tau = 1$, $t + z \geq t_c$ iff $t_s < t + z$ and $R(\theta, M_{t+z}) < 0$. Since t_s is an independent random variable, we have

$$\mathbb{P}(t + z \geq t_c|d^\tau = 1, x_{-i}, s_i) = \mathbb{P}(R(\theta, M_{t+z}) < 0|x_{-i}, s_i, d^\tau = 1)\mathbb{P}(t_s < t + z).$$

For convenience, we write $\mathcal{P}(t + z) := \mathbb{P}(R(\theta, M_{t+z}) < 0|s_i, x_{-i}, d^\tau = 1)$. It is understood that this probability depends on s_i, x_{-i}, but we suppress them to reduce the burden of notation. Therefore, the expected payoff from start to attack at time 0 ($\mathcal{A}$) conditional on $d^\tau = 1$, x_{-i} and s_i is

$$\mathbb{E}\left[\pi(0, t_c)|d^\tau = 1, x_{-i}, s_i\right] = \int_0^l e^{-\beta z}\left[\mathcal{P}(z)G(z^-)\underline{u} + (1 - \mathcal{P}(z)G(z^-))\overline{u}\right]\mathrm{d}B\left(\frac{z}{l}\right), \tag{A.6}$$

and the expected payoff from start attacking at τ conditional on $d^\tau = 1$, x_{-i} and s_i is

$$\begin{aligned}\mathbb{E}\left[\pi(\tau, t_c)|d^\tau = 1, x_{-i}, s_i\right] &= \int_0^l e^{-\beta(z+\tau)}\left[\mathcal{P}(\tau + z)G(z + \tau^-)\underline{u} + (1 - \mathcal{P}(\tau + z)G(z + \tau^-))\overline{u}\right]\mathrm{d}B\left(\frac{z}{l}\right)\\ &= \int_\tau^{l+\tau} e^{-\beta z}\left[\mathcal{P}(z)G(z^-)\underline{u} + (1 - \mathcal{P}(z)G(z^-))\overline{u}\right]\mathrm{d}B\left(\frac{z-\tau}{l}\right)\\ &\geq \int_\tau^l e^{-\beta z}\left[\mathcal{P}(z)G(z^-)\underline{u} + (1 - \mathcal{P}(z)G(z^-))\overline{u}\right]\frac{1}{l}b\left(\frac{z-\tau}{l}\right)\mathrm{d}z. \end{aligned}\tag{A.7}$$

We substitute $\tau + z$ with z to get the second equality, and the last inequality comes from the fact that $\overline{u} > \underline{u} \geq 0$. It follows from (A.6) and (A.7) that conditional on $d^\tau = 1$, x_{-i}, and s_i, the cost of playing $\mathcal{W}$ rather than $\mathcal{A}$, denoted by $\mathbb{C}^S(\mathit{\Gamma}^\tau, x_{-i}, s_i)$, is

$$\begin{aligned}\mathbb{E}\left[\pi(0, t_c) - \pi(\tau, t_c)\middle|d^\tau = 1, x_{-i}, s_i\right] &\leq \int_0^\tau e^{-\beta z}\left[\mathcal{P}(z)G(z^-)\underline{u} + (1 - \mathcal{P}(z)G(z^-))\overline{u}\right]\frac{1}{l}b\left(\frac{z}{l}\right)\mathrm{d}z\\ &+ \int_\tau^l e^{-\beta z}\left[\mathcal{P}(z)G(z^-)\underline{u} + (1 - \mathcal{P}(z)G(z^-))\overline{u}\right]\frac{1}{l}\left(b\left(\frac{z}{l}\right) - b\left(\frac{z-\tau}{l}\right)\right)\mathrm{d}z.\end{aligned}$$

Notice that $e^{-\beta z}\left[\mathcal{P}(z)G(z^-)\underline{u} + (1 - \mathcal{P}(z)G(z^-))\overline{u}\right] \leq \bar{u}$. Since $b(\cdot)$ is Lipschitz continuous, there exists $\kappa \in (0, \infty)$ such that $b\left(\frac{z}{l}\right) - b\left(\frac{z-\tau}{l}\right) \leq |b\left(\frac{z}{l}\right) - b\left(\frac{z-\tau}{l}\right)| \leq \kappa\frac{\tau}{l}$. Moreover, since b is continuous and defined on a closed domain $[0, 1]$, it is bounded; that is, there exist $\bar{b} > 0$ such that $b\left(\frac{z}{l}\right) \leq \bar{b}$. Therefore, for any s_i and $\tau \in (0, l)$,

$$\mathbb{C}^S(\mathit{\Gamma}^\tau, x_{-i}, s_i) \leq \overline{u}\left(\int_\tau^l \frac{1}{l}\kappa\frac{\tau}{l}\mathrm{d}z + \frac{\tau}{l}\bar{b}\right) < \frac{\overline{u}(\bar{b}+\kappa)}{l}\tau. \tag{A.8}$$

The first inequality follows immediately from the above discussions, and the second inequality relies on $\tau > 0$. Note that the expected cost $\mathbb{C}^S(\mathit{\Gamma}^\tau, x_{-i}, s_i)$, by definition, is positive. Its upper bound $\frac{\overline{u}(\bar{b}+\kappa)}{l}\tau$ is independent of s_i, x_{-i} and G, and it is strictly increasing in τ with $\lim_{\tau\to 0}\frac{\overline{u}(\bar{b}+\kappa)}{l}\tau = 0$. Following (A.5), (A.8) and Assumption 3, the expected payoff difference between $\mathcal{W}$ and $\mathcal{A}$ is

$$\begin{aligned}\mathbb{D}^S(\mathit{\Gamma}^\tau, x_{-i}, s_i) =& \mathbb{P}(d^\tau = 0|s_i)\mathbb{B}^S(\mathit{\Gamma}^\tau, x_{-i}, s_i) - \mathbb{P}(d^\tau = 1|x_{-i}, s_i)\mathbb{C}^S(\mathit{\Gamma}^\tau, x_{-i}, s_i)\\ >& \epsilon(e^{-\beta T}\bar{v} - \bar{u})\int_0^l e^{-\beta z}\mathrm{dB}\left(\frac{z}{l}\right) - (1-\epsilon)\frac{\overline{u}(\bar{b}+\kappa)}{l}\tau.\end{aligned}$$

Let us define $\hat{\tau}_l := \min\{\frac{\epsilon}{1-\epsilon}\cdot\left(\frac{e^{-\beta T}\overline{v}-\overline{u}}{\overline{u}(\bar{b}+\kappa)}\right)\int_0^l e^{-\beta z}\mathrm{d}B\left(\frac{z}{l}\right)\cdot l, l\}$. Clearly, for any $\epsilon > 0$, $\hat{\tau}_l > 0$. Therefore, under $\mathit{\Gamma}^\tau$ with any $\tau \in (0, \hat{\tau}_l)$, $\mathbb{D}(\mathit{\Gamma}^\tau, x_{-i}, s_i) > 0$ regardless of s_i and other agents' choices between $\mathcal{A}$ and $\mathcal{W}$, thereby proving the dominance of $\mathcal{W}$ over $\mathcal{A}$. ■

Proof of Proposition 2. Under the extended $\mathit{\Gamma}^\tau$ policy, where $\tau < \overline{\tau}$, each agent chooses a contingency plan $(t^k, t_0^k, t_1^k)_{k=0}^K$, where t^k, t_0^k and t_1^k denote the time of attack after receiving the private signal at t_k, seeing $d^{t_k+\tau} = 0$ and $d^{t_k+\tau} = 1$. It follows from the same argument as in Lemma 1 that (1) since waiting is costly, an agent either attacks at t_k, or he waits at least until the next information arrives at $t_k + \tau$; and (2) if he waits, then it must be that he plans to use the information. Since the regime will surely change following $d^{t_k+\tau} = 1$, he must plan to attack at $t_k + \tau$ after learning this. This implies, after $d^{t_k+\tau} = 0$, he must wait at least until the next information arrives (at time t_{k+1}). This implies that the only rationalizable strategies are $(\mathcal{W})_{k=0}^{K+1}$, in which $\mathcal{W}^k$ ($k = 1, 2, \ldots, K$) means (1) $\forall l < k$, $t^l \geq t_l + \tau, t_0^l \geq t_{l+1}, t_1^l = t_l + \tau$; (2) $t^k = t_k$; and (3) $\forall l > k$, $t^l \geq t_l$, and $t_0^l, t_1^l \geq t_l + \tau$. In addition, $\mathcal{W}^0$ means $t^0 = t_0 = 0$, and $t_0^l, t_1^l \geq t_l + \tau$ for all $l \geq 0$; and W^{K+1} means $t_0^K = \mathbb{T}$, and $t^l \geq t_l + \tau$, $t_0^l \geq t_{l+1}$ and $t_1^l = t_l + \tau$ for all $l \leq K$.

The $\mathcal{W}^k$ strategy means that the agent waits for all the first k alerts and not attack unless an alert is triggered, but then attack at t_k after receiving the new private signal, and not wait for the next alert. Note that $\mathcal{W}^0 = \mathcal{A}$ strategy means attack at time 0 and not wait for any alert; and $\mathcal{W}^{K+1}$ strategy means wait for all the alerts and never attack unless an alert is triggered.

Next, we prove that $\mathcal{W}^{K+1}$ is the unique rationalizable strategy for any type s_i. We prove this by iterated elimination: we first establish the strict dominance of $\mathcal{W}^{K+1}$ over $\mathcal{W}^K$; and then we show that, for any $k \in \{0, 1, \ldots, K-1\}$, if $\mathcal{W}^{K+1}$ strictly dominates any $\mathcal{W}^{k'}$ for all $k' > k$, then $\mathcal{W}^{K+1}$ strictly dominates $\mathcal{W}^k$.

To prove that $\mathcal{W}^K$ is strictly dominated by $\mathcal{W}^{K+1}$, first note that both strategies follow the same path if any of the first K^{th} alerts is triggered, or $d^{t_{K-1}+\tau} = 1$. Therefore, when comparing these two strategies, we only need to consider the histories following $d^{t_{K-1}+\tau} = 0$. Conditional on $d^{t_K+\tau} = 0$, under all possible rationalizable strategy profile $x_{-i} = (x_j)_{j \neq i}$ in which $x_j \in \{\mathcal{W}^k\}_{k=0}^{K+1}$, $N_T = N_{t_K+\tau^-}$. Therefore, as in Lemma 2, $t_c = \infty$ following $d^{t_K+\tau} = 0$. Therefore, similar to (3), the benefit of $\mathcal{W}^{K+1}$ compared to $\mathcal{W}^K$, conditional on $d^{t_K+\tau} = 0$, is

$$\mathbb{B}_K(\mathit{\Gamma}^\tau, x_{-i}, \boldsymbol{s}_i) := \pi(\mathbb{T}, \mathbb{T}) - \pi(t_K, \mathbb{T}) = e^{-\beta T}\overline{V} - e^{-\beta t_K}\overline{U}.$$

On the other hand, if $d^{t_K+\tau} = 1$, under strategy $\mathcal{W}^{K+1}$, the agent attacks at time $t_K + \tau$, whereas he attacks at t_K under strategy $\mathcal{W}^K$. Therefore, the expected cost of $\mathcal{W}^{K+1}$ compared to $\mathcal{W}^K$ can be written as

$$\mathbb{C}_K(\mathit{\Gamma}^\tau, x_{-i}, \boldsymbol{s}_i) := \mathbb{E}\left[\pi(t_K, t_c) - \pi(t_K + \tau, t_c) \middle| d^{t_{K-1}+\tau} = 0, d^{t_K+\tau} = 1, \boldsymbol{s}_i\right].$$

Recall that, under any rationalizable strategy profile x_{-i}, conditional on $d^{t_{K-1}+\tau} = 0$, attacks between time $t_{K-1} + \tau$ and $t_K + \tau$ only occur at time t_K. Therefore, conditional on $d^{t_{K-1}+\tau} = 0$ and $d^{t_K+\tau} = 1$, $R(\theta, N_t) \geq 0$ for any $t < t_K$, and $R(\theta, N_t) < 0$ for any $t \geq t_K$. By the definition of t_c, $t_c = t_K$ if $t_s < t_K$, $t_c \in (t_K, t_K + \tau]$ if $t_s \in [t_K, t_K + \tau)$ and $t_c > t_K + \tau$ if $t_s \geq t_K + \tau$.

Therefore, the cost of waiting $\mathbb{C}_K(\mathit{\Gamma}^\tau, x_{-i}, \boldsymbol{s}_i)$ is

$$e^{-\beta t_K}\left[(1 - G(t_K^-))\overline{U} + G(t_K^-)\underline{U}\right] - e^{-\beta(t_K+\tau)}\left[(1 - G(t_K + \tau^-))\overline{U} + G(t_K + \tau^-)\underline{U}\right].$$

It is easy to see $\mathbb{C}_K(\mathit{\Gamma}^\tau, x_{-i}, \boldsymbol{s}_i)$ is strictly increasing in τ. Under the assumption that G is atomless, as $\tau \to 0$, $G(t_K + \tau^-) - G(t_K^-) \to 0$. Therefore, $\lim_{\tau \to 0} \mathbb{C}_K(\mathit{\Gamma}^\tau, \boldsymbol{s}_i) = 0$.

We can write the expected payoff difference from $\mathcal{W}^{K+1}$ as compared to $\mathcal{W}^K$ as

$$\begin{aligned}
\mathbb{D}_K(\mathit{\Gamma}^\tau, x_{-i}, \boldsymbol{s}_i) \;&:=\; \mathbb{P}(d^{t_{K-1}+\tau} = 1|\boldsymbol{s}_i) \cdot 0 + \mathbb{P}(d^{t_K+\tau} = d^{t_{K-1}+\tau} = 0|\boldsymbol{s}_i)\mathbb{B}_K(\mathit{\Gamma}^\tau, \boldsymbol{s}_i) \\
&\quad\; - \mathbb{P}(d^{t_K+\tau} = 1, d^{t_{K-1}+\tau} = 0|\boldsymbol{s}_i)\mathbb{C}_K(\mathit{\Gamma}^\tau, \boldsymbol{s}_i) \\
&> \;\epsilon \cdot \mathbb{B}_K(\mathit{\Gamma}^\tau, x_{-i}, \boldsymbol{s}_i) - (1 - \epsilon) \cdot \mathbb{C}_K(\mathit{\Gamma}^\tau, x_{-i}, \boldsymbol{s}_i).
\end{aligned}$$

The last inequality follows from the fact that, under the doubt assumption, for any $\boldsymbol{s}_i$, $\mathbb{P}(d^{t_K+\tau} = d^{t_{K-1}+\tau} = 0|\boldsymbol{s}_i) > \epsilon$, and, accordingly, $\mathbb{P}(d^{t_K+\tau} = 1, d^{t_{K-1}+\tau} = 0|\boldsymbol{s}_i) < 1 - \epsilon$. The expected benefit $\mathbb{B}_K$ is strictly positive and is independent of τ and $\boldsymbol{s}_i$, whereas the expected cost $\mathbb{C}_K(\mathit{\Gamma}^\tau, x_{-i}, \boldsymbol{s}_i) \to 0$ if τ decreases to 0. Therefore, as in Lemma 3, for any $\epsilon > 0$, we can find $\tilde{\tau}_K > 0$ such that for any $\tau \in (0, \min\{\tilde{\tau}_K, \bar{\tau}\})$, $D_K(\mathit{\Gamma}^\tau, x_{-i}, \boldsymbol{s}_i) > 0$ irrespective of $\boldsymbol{s}_i$. As this holds true for any x_{-i}, $\mathcal{W}^{K+1}$ strictly dominates $\mathcal{W}^K$.

Now, given that strategy $\mathcal{W}^K$ is strictly dominated, there is no attack at time t_K under any undominated strategies. Therefore, the second to last alert $d^{t_{K-1}+\tau}$ perfectly predicts the regime outcome, i.e., $d^{t_K+\tau} = d^{t_{K-1}+\tau} = \mathbb{1}(t_c < \infty)$. Next, we repeat the same argument and evaluate the expected benefit and cost of strategy $\mathcal{W}^{K+1}$ compared to strategy $\mathcal{W}^{K-1}$. There is no difference between these two strategies if $d^{t_{K-2}+\tau} = 0$. Conditional on $d^{t_{K-2}+\tau} = 0$, the benefit of $\mathcal{W}^{K+1}$ comes from the event when $d^{t_{K-1}+\tau} = 0$

(and accordingly $t_c = \infty$), whereas the cost comes from the event when $d^{t_{K-1}+\tau} = 1$ (and accordingly $t_c < \infty$). Conditional on $d^{t_{K-2}+\tau} = 0$ and $d^{t_{K-1}+\tau} = 1$, the agent attacks at time t_{K-1} under $\mathcal{W}^{K-1}$ and he attacks at time $t_{K-1} + \tau$ under $\mathcal{W}^{K+1}$. Following the same procedures as we did for the comparison between $\mathcal{W}^{K+1}$ and $\mathcal{W}^{K}$, we can show the existence of the cutoff $\tilde{\tau}_{K-1} > 0$ and establish the strict dominance of $\mathcal{W}^{K+1}$ over $\mathcal{W}^{K-1}$ under any policy Γ^τ with $\tau \in (0, \min\{\tilde{\tau}_{K-1}, \tilde{\tau}_K, \bar{\tau}\})$.

Following the same procedures, we can find $(K+1)$ cutoffs $\{\tilde{\tau}_k\}_{k=0}^{K}$, and establish the strict dominance of $\mathcal{W}^{K+1}$ over $\mathcal{W}^{k}$ (for all $k = 0, 1, \ldots, K$) under any disclosure policy Γ^τ with $\tau \in (0, \min\{\tilde{\tau}, \bar{\tau}\})$, where $\tilde{\tau} := \min_{k=0}^{K} \tilde{\tau}_k > 0$. This completes the proof. ■

Proof of Proposition 3.

No Disclosure If there is no disclosure, the game boils down to a binary action static regime change game similar to Morris and Shin (2003) (henceforth MS), where each agents decides whether to "attack" or "not attack" at time 0. If an agent attacks, then his payoff is $\overline{U}$; and if he does not attack, his payoff is $e^{-\beta T}\overline{V}$ if the regime survives at T (i.e., $t_c = \infty$) and $e^{-\beta T}\underline{V}$ if the regime changes (i.e., $t_c < \infty$). The regime survives if and only if $R(\theta, N_T = N_0) = \theta - N_0 \geq 0$. Thus, under no disclosure, an agent will not attack if and only if the regime survives with probability

$$\mathbb{P}(\theta \geq N_0 | s_i) \geq \frac{\overline{U} - e^{-\beta T}\underline{V}}{e^{-\beta T}\overline{V} - e^{-\beta T}\underline{V}} = \hat{p}.$$

Following the standard global game argument, IESDS yields a unique rationalizable strategy, under which, agents attack if and only if $s_i < \hat{s}$ for some cutoff $\hat{s}$, and accordingly, the regime changes if and only if $\theta < \hat{\theta}$. When others follow such a monotone strategy, the agent who receives higher signal is more optimistic about the regime surviving. The marginal agent with $s_i = \hat{s}$ must be indifferent between attacking and not attacking. This gives us $\mathbb{P}(\theta \geq \hat{\theta} | \hat{s}) = \hat{p}$. Given uniform prior and Gaussian noise, $\theta | s_i \sim N(s_i, \sigma^2)$. Therefore, the above indifference condition simplifies to $\Phi(\frac{\hat{s}-\hat{\theta}}{\sigma}) = \hat{p}$. Moreover, when other agents play such monotone strategy, there is less attack when θ is higher. The fundamental cutoff $\theta = \hat{\theta}$ is such that $N_0(\hat{\theta}) = \mathbb{P}\left(s < \hat{s} | \theta = \hat{\theta}\right) = \Phi(\frac{\hat{s}-\hat{\theta}}{\sigma}) = \hat{\theta}$. Therefore, we have $\hat{\theta} = \hat{p}$.

Policy Γ^τ Under a disaster alert Γ^τ, since $\mathbb{P}(\theta \in \Theta^L = (-\infty, 0) | s_i) > 0$ and $\mathbb{P}(\theta \in \Theta^U = [1, +\infty) | s_i) > 0$ for all s_i, following Lemma 1, the rationalizable strategies are $\mathcal{A}$ and $\mathcal{W}$.[40] Similar to the static game, if an agent plays $\mathcal{A}$ and attack at $t = 0$, he gets $\overline{U}$. However, if an agent plays the strategy $\mathcal{W}$, conditional on $d^\tau = 0$, he never attacks and gets $e^{-\beta T}\overline{V}$ (since the regime survives following $d^\tau = 0$), and, conditional on $d^\tau = 1$ (and consequently $t_c = t_s^+$), he attacks at time τ and gets the expected payoff

$$\mathbb{E}[\pi(\tau, t_c)] | x_{-i}, s_i, d^\tau = 1] = e^{-\beta \tau}\left(G(\tau^-)\underline{U} + (1 - G(\tau^-))\overline{U}\right).$$

[40]Note that this is the only condition on exogenous information that is required for the proof of Lemma 1 (See footnote 38).

Note that this expected payoff is strictly greater than $e^{-\beta T}\underline{V}$, the payoff from not attacking a failed regime in the static game.

Recall that under policy Γ^τ, the regime survives if and only $d^\tau = 0$ (Lemma 2), and, similar to the static game, $d^\tau = 0$ if and only if $R(\theta, N_0) = \theta - N_0 \geq 0$. Therefore, an agent plays $\mathcal{W}$ if and only if he believes that $d^\tau = 0$, and, accordingly, the regime survives (i.e., $\theta \geq N_0$) with probability

$$\mathbb{P}(\theta \geq N_0|s_i) \geq \frac{\overline{U} - e^{-\beta\tau}\left(G(\tau^-)\underline{U} + (1 - G(\tau^-)\overline{U}\right)}{e^{-\beta T}\overline{V} - e^{-\beta\tau}\left(G(\tau^-)\underline{U} + (1 - G(\tau^-)\overline{U}\right)} = p^*(\tau).$$

Note that $p^*(\tau) < \hat{p}$ for any $\tau \in (0, T)$ because

$$e^{-\beta T}\underline{V} < e^{-\beta\tau}\left(G(\tau^-)\underline{U} + (1 - G(\tau^-))\overline{U}\right) < \overline{U} < e^{-\beta T}\overline{V}.$$

In addition, it is easy to check that $p^*(\tau)$ increases with τ and $\lim_{\tau\downarrow 0} p^*(\tau) = 0$.

Following the standard global game argument (as in case of no disclosure policy), IESDS yields a unique rationalizable strategy, under which, agents takes $\mathcal{A}$ if and only if $s_i < s^*(\tau)$ for some cutoff $s^*(\tau)$, and accordingly, the regime changes if and only if $\theta < \theta^*(\tau)$. Therefore, the indifference condition for the marginal agent $s^*(\tau)$ is $\Phi\left(\frac{s^*(\tau)-\theta^*(\tau)}{\sigma}\right) = p^*(\tau)$, and at $\theta = \theta^*(\tau)$ the aggregate attack $N_0(\theta^*(\tau)) = \Phi\left(\frac{s^*(\tau)-\theta^*(\tau)}{\sigma}\right) = \theta^*(\tau)$. Thus, for any $\tau \in (0, T)$, we have

$$\theta^*(\tau) = p^*(\tau). \tag{A.9}$$

Therefore, $\Theta^P(\Gamma^\tau) = [0, p^*(\tau))$. Recall that, under no disclosure policy, $\Theta^P = [0, \hat{p})$. Since $p^*(\tau) < \hat{p}$, $\Theta^P(\Gamma^\tau) \subset \Theta^P$. Finally, recall that $p^*(\tau)$ strictly increases with τ and $\lim_{\tau\downarrow 0} p^*(\tau) = 0$. Therefore, for any $\zeta > 0$, we can always find $\hat{\tau}(\zeta) > 0$ such that for any $\tau < \hat{\tau}(\zeta)$, $\theta^*(\tau) = p^*(\tau) < \zeta$. ■

Proof of Proposition 4. First of all, the option value argument (Lemma 1) still holds with false-positive errors. That is because: (1) since $d^\tau = 1$ predicts $t_c < \infty$ regardless of what agents do after time τ (as there is no false alarm), agents will attack following $d^\tau = 1$; and (2) following the same argument as in Step 3 of the proof of Lemma 1, any strategy involving waiting ($t_\emptyset \geq \tau$) and then attacking after $d^\tau = 0$ ($t_0 \in [\tau, T]$) is strictly dominated by $\mathcal{A}$. Therefore, with false-positive errors, the only rationalizable strategies are $\mathcal{A}$ and $\mathcal{W}$.

Fix any rationalizable strategy profile x_{-i}, for convenience, let us denote $\mathbb{P}(d^\tau = 0, t_c < \infty|s_i, x_{-i})$ as $\mathcal{P}^e(s_i, x_{-i})$, $\mathbb{P}(d^\tau = 0, t_c = \infty|s_i, x_{-i})$ as $\mathcal{P}^0(s_i, x_{-i})$, and $\mathbb{P}(d^\tau = 1, t_c < \infty|s_i, x_{-i})$ as $\mathcal{P}^1(s_i, x_{-i})$. Therefore, we can write the expected payoff from $\mathcal{W}$ as

$$\mathcal{P}^1(s_i, x_{-i})e^{-\beta\tau}\left(G(\tau^-)\underline{U} + (1 - G(\tau^-))\overline{U}\right) + \mathcal{P}^0(s_i, x_{-i})e^{-\beta T}\overline{V} + \mathcal{P}^e(s_i, x_{-i})e^{-\beta T}\underline{V}. \tag{A.10}$$

As the payoff from $\mathcal{A}$ is $\overline{U}$, the expected payoff difference is

$$\begin{aligned}\mathbb{D}(\Gamma^\tau, x_{-i}, s_i) =&\mathcal{P}^1(s_i, x_{-i})\left(e^{-\beta\tau}\left[G(\tau^-)\underline{U} + (1-G(\tau^-))\overline{U}\right] - \overline{U}\right)\\ &+\mathcal{P}^0(s_i, x_{-i})\left(e^{-\beta T}\overline{V} - \overline{U}\right) + \mathcal{P}^e(s_i, x_{-i})\left(e^{-\beta T}\underline{V} - \overline{U}\right).\end{aligned}$$

Recall that the probability of a false-positive is, at most, η and this upper bound is independent of s_i and x_{-i} —i.e., $\mathbb{P}(t_c < \infty | d^\tau = 0, s_i) = \frac{\mathcal{P}^e(s_i, x_{-i})}{\mathcal{P}^e(s_i, x_{-i}) + \mathcal{P}^0(s_i, x_{-i})} \leq \eta$. Thus, $\mathcal{P}^e(s_i, x_{-i}) \leq \frac{\eta}{1-\eta}\mathcal{P}^0(s_i, x_{-i})$. In addition, as $\mathcal{P}^1(s_i, x_{-i}) + \mathcal{P}^0(s_i, x_{-i}) + \mathcal{P}^e(s_i, x_{-i}) = 1$, $\mathcal{P}^1(s_i, x_{-i}) \leq 1 - \mathcal{P}^0(s_i, x_{-i})$ for any s_i and x_{-i}. Since

$$e^{-\beta T}\underline{V} - \overline{U} < e^{-\beta\tau}\left[G(\tau^-)\underline{U} + (1-G(\tau^-))\overline{U}\right] - \overline{U} < 0 < e^{-\beta T}\overline{V} - \overline{U},$$

we have

$$\begin{aligned}\mathbb{D}(\Gamma^\tau, s_i) \geq &\left(1 - \mathcal{P}^0(s_i, x_{-i})\right)\left(e^{-\beta\tau}\left[G(\tau^-)\underline{U} + (1-G(\tau^-))\overline{U}\right] - \overline{U}\right)\\ &+\mathcal{P}^0(s_i, x_{-i})\left(e^{-\beta T}\left[\overline{V} + \frac{\eta}{1-\eta}\underline{V}\right] - \frac{1}{1-\eta}\overline{U}\right).\end{aligned}$$

Define $\eta_0 := \frac{\overline{V} - e^{\beta T}\overline{U}}{\overline{V} - \underline{V}}$. Under Assumption 2, η_0 is strictly positive, and for $\eta < \eta_0$, $e^{-\beta T}\left[\overline{V} + \frac{\eta}{1-\eta}\underline{V}\right] - \frac{1}{1-\eta}\overline{U} > 0$. Moreover, recall that, under Assumption 3, $\mathcal{P}^0(s_i, x_{-i}) \geq \mathbb{P}(\theta \in \Theta^U | s_i) > \epsilon$ for all $s_i \in \mathbb{S}$ and x_{-i}. Then, we can define $\hat{\tau}(\eta) > 0$ for any $\eta < \eta_0$ as follows.

$$\hat{\tau}(\eta) := \min\left\{\frac{\epsilon}{2(1-\epsilon)}\frac{e^{-\beta T}\left(\overline{V} + \frac{\eta}{1-\eta}\underline{V}\right) - \frac{\overline{U}}{1-\eta}}{\beta\overline{U}}, G^{-1}\left(\frac{\epsilon}{2(1-\epsilon)}\frac{e^{-\beta T}\left(\overline{V} + \frac{\eta}{1-\eta}\underline{V}\right) - \frac{\overline{U}}{1-\eta}}{\overline{U} - \underline{U}}\right)\right\}.$$

It is easy to check that, given $\eta < \eta_0$ and $\mathcal{P}^0(s_i, x_{-i}) > \epsilon$, $\mathbb{D}(\Gamma^\tau, s_i) > 0$ for any $\tau \in (0, \hat{\tau}(\eta))$. That is, $\mathcal{A}$ is strictly dominated by $\mathcal{W}$. Following the proofs of Lemma 3 and Theorem 1, $\Theta^P(\Gamma^\tau) = \emptyset$ for any $\tau \in (0, \hat{\tau}(\eta))$. ■