Panics and Early Warnings*

Deepal Basak[†] Zhen Zhou[‡]

Oct 2021

Abstract

We show how early warning about an impending regime change eliminates panic. Agents anticipate a future shock and decide when to attack. Waiting is costly, especially when others attack and cause a regime change while one waits. This may create panic. We model early warning as a dynamic disclosure policy and establish its optimality in a reasonably general setup. Under this policy, the unique rationalizable strategy is to ignore private information, wait for the warning and then follow it. We discuss the robustness and limitations of this policy by relating it to forward-looking bank stress tests and debt sustainability analysis.

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

Key Words: *Coordination, Dynamic Information Design, 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 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

When agents anticipate that an adverse shock may cause a crisis in the near future, they may panic and attack preemptively. This, in turn, may cause a crisis that is not warranted. An early warning system (EWS henceforth) sends a warning to the agents regarding this future crisis. In this paper, we formalize panic in a dynamic regime change game, where (1) a shock arrives at a future date (possibly stochastic), and once it arrives, it may make the regime vulnerable; and (2) the agents choose when to attack, if at all, and waiting is costly. We design a simple optimal dynamic information disclosure policy, which resembles an EWS. We show that under a reasonably general setup, this simple policy completely eliminates panic.

To fix ideas, consider the following examples. A country has a significant amount of outstanding sovereign debt that matures in, say, six months, and the country may default. In this case, we say that the shock arrives at this fixed maturity date. Foreign investors who are wary about the forthcoming crisis and resulting currency depreciation may panic and relocate their investments before the maturity date. These decisions worsen the fiscal condition of this country, which contributes to it not being able to repay its debt when the debt matures. Similarly, consider bank 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 only difference from the first example is that the date when the bank becomes vulnerable is stochastic rather than fixed. The creditors may panic and run on the bank before the shock actually arrives, which weakens the bank and contributes to its failure. The first-mover advantage — the creditor gets a higher payoff if he withdraws before others attack and the resulting the bank failure — could accelerate the run.

These examples are often modeled as a regime change game, in which the agents simultaneously decide whether to attack the regime, and the regime changes if the fundamental is not sufficiently strong to withstand the attack.[1] In our dynamic setting, because of discounting and first-mover advantage, delayed attack is always costly. In the absence of any new information, if an agent decides to attack, he will attack right away. Thus, without future information, our dynamic game boils down to this classical 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 (such as EWS).[2]

EWS is common in practice. For instance, the IMF and the World Bank regularly conduct forward-looking Debt Sustainability Analysis (DSA), aiming to provide the market with an early

[1] For instance, see Cole and Kehoe (2000) for sovereign debt crisis and Rochet and Vives (2004) and Goldstein and Pauzner (2005) for bank run.

[2] In the above examples, the gap arises 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, or there could be regulatory restrictions. Our insight applies to such environments, as well (See Proposition 3).

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 a bank would be able to sustain that adverse shock or not.[3]

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, regardless of what the agents do thereafter. However, if the alert is not triggered, it means the regime will survive provided that there is no further attack. We call this alert timely when $\tau \in (0, \hat{\tau})$ for some small but positive $\hat{\tau}$.

The disclosure policy is simple — binary and public. However, it is important to note two subtle properties. (1) It is an *endogenous disclosure*: if agents attack and do not wait for the alert, that may trigger the alert. (2) It *forecasts* an impending regime change: the alert gets triggered even though the regime has not changed yet, but it 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 private signals about the fundamental. The principal does not observe these signals. The signals need not be conditionally independent: we allow for any arbitrary correlation (some agents may share information). We assume only that regardless of their signals, and regardless of others' actions, an agent will always have some *doubt* about whether his action is the right action.

We show that under a timely disaster alert, the unique rationalizable strategy for any agent is to wait for and then follow the disaster alert, regardless of his signal. 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 thereafter. Therefore, panic is eliminated in the robust sense (even in the worst-case scenario).

To understand the logic behind our main result, note that an agent's incentive to wait for a future disclosure depends not only on what he will learn from that disclosure, but also on whether others would wait for that disclosure, and if they do, how they will respond.

First, consider an agent who has made a contingency plan that involves waiting for the disaster alert. We show that, regardless of others' strategy, this agent must follow the disaster alert — 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 makes a contingency plan to wait for the alert and then attack, regardless

[3]While this paper focuses on how to optimally design the bank stress test to avert panic, several recent studies look into other aspects of stress test as an information disclosure policy. For instance, Goldstein and Leitner (2018) studies how stress tests can be adopted to facilitate banks of heterogeneous quality to raise funds under adverse selection; Orlov, Zryumov and Skrzypacz (2018) studies how stress tests help to discover systemic risk and how it can be combined with capital regulations; Inostroza (2019) investigates a comprehensive disclosure policy including banks' liquidity position and asset quality.

of whether or not the alert is triggered. As such, there is no option value in waiting. Since waiting is costly, such contingency plans 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 as the *option value* argument (See Lemma 1).[4]

The disaster alert is so designed that when the alert is not triggered, regardless of how many agents choose to attack or wait, 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, there will be no further attack and the regime never changes. This demonstrates the *perfect predictability* of the disaster alert (See Lemma 2).

It follows from the perfect predictability of the alert that if an agent chooses to wait for it, his own noisy information becomes 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, if the regime changes while the agents wait for the alert, then the agent may miss the opportunity to attack a regime before it fails, 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 others attack or wait and follow, the benefit outweighs the cost of waiting. This shows the importance of a *timely disclosure* (See Lemma 3).

It follows from the three abovementioned Lemmas 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 perfectly coordinates the agents' actions and eliminates panic in the robust sense. For simplicity of exposition, we assume that the agents do not receive any other information over time. We show that the theoretical insight can be easily extended when the agents see when the regime changes, or infrequently receive additional exogenous information over time.

This result is in contrast to the literature on optimal disclosure in a static regime change game. Under conditionally independent signals, Goldstein and Huang (2016) finds the optimal monotone public disclosure policy. More recently, Inostroza and Pavan (2020) explores the benefit of non-

[4]It follows from the doubt assumption that the agents always assign positive probability to the event that the alert gets triggered and also to the event that the alert does not get triggered.

monotone public disclosures for agents endowed with private information; and Li, Song and Zhao (2019) finds the optimal private disclosure when the agents share a homogeneous belief. It is important to note that under all these optimal policies, a regime may change even though the fundamental does not warrant it. That is, panic cannot be eliminated in the robust sense.

The contrasting result comes from the fact that, unlike the static disclosure, the agents can endogenously choose to wait for the disaster alert or not. If they choose to wait, then they will follow the alert. However, this option value argument alone is not sufficient to persuade the agents to wait for the alert. It is essential that the disaster alert sends different messages depending on how many agents endogenously choose to wait for the disclosure.

To see this, consider an alternative EWS which at a future date τ discloses only whether or not the fundamental warrants a run. Unlike the disaster alert, this policy only discloses information about the exogenous fundamental, regardless of how many agents endogenously choose to attack before the disclosure. Notice that, under the timely disaster alert policy, on path, since all the agents wait for the alert, the agents learn the same information as under this alternative policy. However, that does not imply that the alternative policy can eliminate panic in a robust sense. Without the endogenous disclosure property, this alternative EWS cannot perfectly predict the regime outcome. This is because, the regime may still change due to past attacks even though agents who waited do not attack following no warning. Accordingly, agents may not wait for such a future disclosure and may panic. This shows why EWS should not only focus on the fundamental, but should also incorporate the historical market reactions.

The other essential property of the timely disaster alert policy that persuades the agents to wait for the disaster alert is that it is a forecast. In a companion paper, Basak and Zhou (2020) consider a different dynamic environment where agents move sequentially in an exogenous order (as under a staggered debt structure) and they learn whether the regime has already changed or not (no forecasting). The authors show that when the debt structure is sufficiently staggered — that is, the viability of the regime is disclosed sufficiently frequently — agents attack only when the regime is not viable. The underlying argument is inductive and it exploits the exogenous order — the agents follow the viability news because they believe that other agents moving later would also follow the subsequent disclosures. While under endogenous timing, the agents who choose to wait for the alert, will follow the alert regardless of what others do.

It is important to note that a frequent viability test does not eliminate panic when the timing of the attack is endogenous. Since frequent viability tests are not forecasts, the agents may not be warned about the impending regime change simply because shock has not arrived yet; and when the shock arrives, the regime can change even without further attack. Thus, perfect predictability fails. Moreover, since when the alert is triggered, the regime has already changed, it is too late to attack. This makes waiting for the alert significantly costly. Therefore, the agents may not find it worth waiting for such a future disclosure.[5]

We saw that a simple EWS can eliminate panic in a fairly general dynamic regime change setup. However, this is certainly not without limitation. For instance, we can find examples of the following situations: the principal may not only care about the regime outcome, but would prefer to delay the regime change even if it is inevitable; the principal makes mistakes while forecasting; or the shock has already arrived and thus the regime may change before the principal can disclose any information (so that the waiting cost is not negligible even under a timely disclosure). In Section 3, we discuss the limitations and robustness of the EWS by considering these alternative cases. This discussion yields some positive implications for the design of EWS, and demonstrates how supplementary policies can help to reinforce the effectiveness of EWS.

Literature This paper contributes to two stands of the literature: (1) dynamic coordination; and (2) 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), Dasgupta, Steiner and Stewart (2012)). The distinctive feature of our model is that we consider a principal who manipulates the agents' beliefs by optimally choosing how the agents learn about the past.

There is a large and growing literature on information design. See Bergemann and Morris (2019) and Kamenica (2019) for recent surveys. Most of the studies in this literature that consider multiple agents are restricted to simultaneous move games, and the ones that consider a dynamic setting, are restricted to a single agent.[6] In this paper, to capture the panic, we consider an adverse shock that arrives at a future date and a mass of privately informed agents who choose when to attack. This allows the information designer to send different messages, conditional on how many agents wait for the disclosure. As we can see, in the context of coordination, such an information design could be remarkably effective. Some recent studies have applied the optimal dynamic information design to multiple agents in other practical contexts. For instance, Che and Hörner (2018) finds the optimal "spam" recommendations to agents arriving early to facilitate social learning. Ely et al. (2021) shows that in a contest, to maximize effort from the agents, a principal should implement a review cycle — set a time limit for when she discloses past successes; if there is a success, then stop; otherwise, reset the clock. Li, Szydlowski and Yu (2021) finds that, in a dynamic entry game with two agents, the optimal disclosure to attract an adopter and deter a competitor features

[5] On the other hand, when timing of attack is exogenous, the option value argument fails. Therefore, a timely disaster alert may not eliminate panic in the robust sense when agents move in exogenous order.

[6] See Bergemann and Morris (2016) and Mathevet, Perego and Taneva (2020) for static disclosure with multiple agents, and Ely (2017), Orlov, Skrzypacz and Zryumov (2020), Ely and Szydlowski (2020) and Ball (2019) for dynamic disclosure with a single agent.

information reversals (good news followed by bad, and vice versa).

The rest of the paper is organized as follows. Section 1 describes the model and solution concept. Section 2 shows how a timely disaster alert eliminates panic. Section 3 discusses the limitations and robustness, and Section 4 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. Under the canonical regime change game, the agents simultaneously decide whether to attack the regime, and the regime survives if the fundamental is sound enough to withstand the attack. The agents prefer attacking if they believe that the regime will change, and vice versa. This can create panic; that is, a regime may change even though the fundamental does not warrant it.

However, as we have seen in our motivating examples, anticipating some future shock, agents may panic before the shock actually arrives. This often creates a gap between when the agents start panicking and when the regime changes. This gap could provide an opportunity for the principal to reduce such panics. To understand this 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 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 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 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, which is non-decreasing in time t.

Timing We consider a continuous time model. An agent who has not attacked by time t has an opportunity to attack again at $t + \Delta t$, where $\Delta t \to 0$. We use the notation t^- (and t^+) to capture the instance before (and after) t but excluding t. Between time t and time $t + \Delta t$, 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 is determined. The details of how the regime outcome is determined will be discussed shortly. Following that, the agents who attack collect their payoffs depending on the regime outcome. 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.

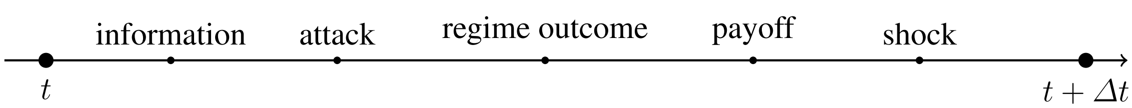

Figure 1: Timing

Note that we use the convention that the shock arrives after the regime outcome is determined at any t. This gives the agents a chance to attack and collect their payoffs at $t_i = 0$ even before any shock can arrive. If we use the convention that the shock arrives before they get the first opportunity to attack, then it is easier to dissuade the agents from attacking. See Footnote 16 for details.

Regime Change There is a differentiable function $R(\theta, N_t)$ governing the regime's preparedness to face the shock at any time t. Better 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., $\frac{\partial R}{\partial N_t} < 0$. At any date t after the shock arrives, the regime changes if and only if $R(\theta, N_t) < 0$. Let t_c be the time when the regime changes. By definition, $t_c \geq t_s^+$. Recall that N_t is non-decreasing in time, while θ is drawn at the beginning of the game and remains the same. Thus, t_c is the first instance after the shock arrives when $R(\theta, N_t) < 0$. If $R(\theta, N_T) \geq 0$, the regime never changes and survives in the end. This is denoted by $t_c = \infty$. Formally,

$$t_c := \begin{cases} \min\{t \in [t_s^+, \infty) | R(\theta, N_t) < 0\} \text{ if } R(\theta, N_T) < 0 \\ \infty \text{ if } R(\theta, N_T) \geq 0. \end{cases} \tag{1}$$

Note that if at $t \leq t_s$ (before the shock's arrival), $R(\theta, N_t) < 0$; then, regime does not change at time t, but it is inevitable that the regime will change as soon as the shock arrives. This creates a gap between when the agents start attacking and when the regime changes.

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

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

This payoff is influenced by other agents' choices $(t_j)_{j \neq i}$ through t_c. If an agents attacks before the regime changes ($t_i < t_c$), he gets a high payoff $\overline{U}$, and if he attacks after the regime changes ($t_i \geq t_s$), he gets a low payoff $\underline{U}$. Since, by convention, the regime cannot change at time 0, $\pi(0, t_c) = \overline{U}$. However, if an agent does not attack at all ($t_i = \mathbb{T}$), then he obtains

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

Note that $\mathbb{T} < t_c$ iff $t_c = \infty$. Therefore, if an agent never attacks, then at time T, he gets a high payoff $\overline{V}$ if the regime survives in the end ($t_c = \infty$), and a low payoff $\underline{V}$ otherwise ($t_c \in [0^+, T^+]$).

We extend the strategic complementarity assumption in a natural manner in this dynamic environment. We assume the following. First, if the regime changes at some date ($t_c < \infty$), an agent is better off if he attacks the regime at any date ($t_i \in [0, T]$) than if he never attacks ($t_i = \mathbb{T}$). This is captured through $\underline{U} > \underline{V}$. Second, given the regime changes, an agent is better off if he attacks before the regime changes ($t_i < t_c$) rather than after the regime changes. This is captured through $\overline{U} > \underline{U}$. Finally, if the regime never changes ($t_c = \infty$), an agent prefers to never attack the regime ($t_i = \mathbb{T}$) over attacking the regime at any date $t_i \in [0, T]$. This is captured through $e^{-\beta T}\overline{V} > \overline{U}$.

Assumption 1 *(Complementarity) The payoffs satisfy the following inequalities:*

$$e^{-\beta T}\overline{V} > \overline{U} > \underline{U} > \underline{V} \geq 0.$$

Notice that under this endogenous timing choice, the payoff specification features a natural "first-mover advantage." If others attack and cause a regime change, then an agent can get a higher payoff by attacking before others than attacking afterwards. This is arguably the main reason for panicking.[7]

[7] Note that since delay is costly ($\beta > 0$ and $\overline{U} > \underline{U} > 0$), if an agent knows that the regime will change, then he prefers attacking as early as possible. This is different from riding a stock bubble, where an agent wants to wait until the last minute before the regime changes (See Abreu and Brunnermeier (2003)). We think this is a very interesting but substantially different setup. We leave the information design question in such a setup for future research.

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 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).

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

Assumption 2 *(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 ϵ 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 action.[8]

In Section 3.3.2, 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

[8]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 $\theta \in \Theta$, and $s_i \in \mathbb{S}$, where $0 < \underline{f} \leq \overline{f} < +\infty$ — then Assumption 2 holds true for any $\epsilon \in (0, \min\{\frac{(1-\Psi(\overline{\theta}))\underline{f}}{\overline{f}}, \frac{\Psi(\underline{\theta})\underline{f}}{\overline{f}}\})$.

over time. We relax this assumption later. We show that the main insight can be easily extended when the agents observe the regime change (if it so happens) (See Section 2.5) or infrequently receive additional noisy signals regarding the fundamental θ (See Section 2.6).

Principal The principal is concerned about the final outcome of the regime and prefers the survival of the regime. The principal's payoff is simply $\mathbb{1}\{R(\theta, N_T) \geq 0\}$.[9] Similar to the agents, the principal does not observe exactly when the shock arrives. She also 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 before the shock actually arrives.[10]

Disclosure Policy The principal (she) *commits* to a dynamic public information disclosure policy denoted by $\mathit{\Gamma} = (q, \mathcal{S})$, where $\mathcal{S}$ is the message space, and $q : [0, T] \times \mathit{\Theta} \times [0, 1] \to \mathit{\Delta}(\mathcal{S})$ is a mapping that specifies a public message depending on the date $\tau \in [0, T]$, the fundamental $\theta \in \mathit{\Theta}$, and the history of attack until then N_{τ^-}. Note that, while 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 to hear the message.[11] 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*.

Solution Concept We consider the strategic form representation of the dynamic game. Given the disclosure policy $\mathit{\Gamma}$, a strategy of an agent of any type is to make a contingency 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

[9] More-nuanced preferences and the associated challenges will be discussed in Section 3.1.

[10] This is consistent with the fact that bank supervisors and international organizations rely on historical data and econometric models to forecast the preparedness to face some future shock. In Section 3.2, we discuss small errors in prediction.

[11] We can also allow for disclosure conditional of the whole history of attack (i.e., $(N_s)_{s \leq \tau^-}$), or private disclosure. However, as will be clear soon, this does 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.

of strictly dominated strategies. Our results will hold true for more-restrictive solution concepts such as PBE (See the Online Appendix).

Given the disclosure policy $\mathit{\Gamma}$, let $\mathcal{R}(\mathit{\Gamma})$ be the set of rationalizable strategy profiles. Define

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

Thus, if $\theta \notin \mathit{\Theta}^F(\mathit{\Gamma})$, then the regime will survive, regardless of whatever rationalizable strategies the agents play; otherwise, if $\theta \in \mathit{\Theta}^F(\mathit{\Gamma})$, the regime may not survive.

Robust/Adversarial Information Design 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. Her objective is

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

where $\mathbb{P}(\theta \in \mathit{\Theta}^F(\mathit{\Gamma})) = \int_{\mathit{\Theta}^F(\mathit{\Gamma})} \mathrm{d}\mathit{\Psi}(\theta)$. In words, she chooses policy $\mathit{\Gamma}$ to minimize the ex-ante chance that the regime may change under any possible rationalizable strategy profile $x \in \mathcal{R}(\mathit{\Gamma})$.

Panic A distinctive feature of the regime change game is that when $\theta \in \mathit{\Theta}^L$, it is inevitable that the regime will change, regardless of what the agents do. Hence, any disclosure policy $\mathit{\Gamma}$ cannot endure such a regime – i.e., $\mathit{\Theta}^L \subseteq \mathit{\Theta}^F(\mathit{\Gamma})$ for any $\mathit{\Gamma}$. However, a regime could also change even when it is not warranted ($\theta \notin \mathit{\Theta}^L$), 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 $\mathit{\Theta}^P(\mathit{\Gamma}) := \mathit{\Theta}^F(\mathit{\Gamma}) \setminus \mathit{\Theta}^L$ for any policy $\mathit{\Gamma}$ as the set of fundamentals in which the regime change may happen, and if it happens, it is caused by panic-based attacks. Therefore, the principal's objective is equivalent to minimizing the chance of panic $\mathbb{P}(\theta \in \mathit{\Theta}^P(\mathit{\Gamma}))$. If $\mathit{\Theta}^P(\mathit{\Gamma}) = \emptyset$, then we can conclude that policy $\mathit{\Gamma}$ eliminates panic in a robust sense.

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 ($\theta \notin \mathit{\Theta}^L$), 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 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 disclosed 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 that 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) shows that if the agents have sufficiently heterogeneous beliefs, there is a unique rationalizable strategy: attack if and only if the 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 3.3.2 for details).

Partial Disclosure There exist many time-0 partial disclosure polices. Consider the following policy Γ^0: at time 0, the principal publicly discloses whether the regime change is warranted ($\theta \in \Theta^L$, or equivalently, $R(\theta, 0) < 0$) or not. 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 signals, Goldstein and Huang (2016) finds the optimal monotone public disclosure policy. The authors show that if the principal discloses whether $R(\theta, 0) > k$ or not for a sufficiently large k, then the agents will never attack when $R(\theta, 0) > k$. Inostroza and Pavan (2020) investigates when monotone public disclosure is indeed optimal and the conditions under which non-monotone public disclosure can do better. When the agents have homogeneous beliefs, Li, Song and Zhao (2019) shows that under the optimal disclosure policy, the principal should locally exaggerate the state to some agents. Under all these optimal policies, a regime may change even though the fundamental does not warrant it ($\theta \notin \Theta^L$). Thus, under the static setting, although partial disclosure policies can improve the "worst possible" outcome, an unwarranted regime change cannot be avoided in the robust sense.

The above discussion points out 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 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, denoted as Γ^τ. The principal sends a binary signal $d^\tau \in \{0, 1\}$ at a fixed date $\tau \in (0, T)$. The public signal d^τ 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 $d^\tau = 0$.[12] 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., $t_c \leq \tau$. Next, consider the case in which, ex-post, the shock has not yet arrived by the time of disclosure – i.e., $t_s \geq \tau$. 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., $R(\theta, N_T = N_{\tau^-}) \geq 0$. However, the survival of the regime is not guaranteed following $d^\tau = 0$. If some agents choose to

[12] 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.

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 — 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 it 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 ($t_s \geq \tau$), but it is evident that it will change.

2.3 Don't Panic

The proposed policy can be interpreted simply as assurance from the principal — "Do not panic; just wait until time τ, and at that time, I will advise you to attack 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 the agents will always do so and follow the principal's advise regardless of their own information (even in the worst case). In other words, a timely disaster alert eliminates panic in the robust sense.

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

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 $\mathit{\Gamma}^\tau$ 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 the unnecessary delay cost. Therefore, any contingency plan 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 $t_0 \in [\tau, T]$. 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 $\mathcal{W}$).[13]

Different from disclosing information at $t = 0$, under $\mathit{\Gamma}^\tau$, 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 $\mathit{\Gamma}^0$ disclosure policy (See Section 2.1). Recall that under the policy $\mathit{\Gamma}^0$, 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 $\mathit{\Gamma}^\tau$ 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. Based on Lemma 1, 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 $d^\tau = 1$. Following Lemma 1, we can also infer the time of regime change t_c. By definition of d^τ, the alert is triggered ($d^\tau = 1$) if and only if $R(\theta, N_{\tau^-}) < 0$. Lemma 1 indicates that, if any agent attacks before the disclosure time τ, he plays strategy $\mathcal{A}$ and attacks 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.

[13] 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).

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

Lemma 2 demonstrates that if the agents believe that all others play rationalizable strategies ($\mathcal{A}$ or $\mathcal{W}$),[14] then they understand that d^τ perfectly predicts the ultimate regime outcome. The two properties of the disaster alert — endogenous disclosure and forecasting — are essential for this perfect predictability.

To see the role of endogenous disclosure, consider an alternative policy $\tilde{\Gamma}^\tau$: at time τ, the principal sends the message $\tilde{d}^\tau = 1$ if $R(\theta, 0) < 0$ (or, equivalently, $\theta \in \Theta^L$), and message $\tilde{d}^\tau = 0$, otherwise. Different from our policy Γ^τ, the design of this alert is based purely on the exogenous fundamental θ and it is independent of the endogenous attack N_{τ^-}. It is important to note that if $d^\tau = 0$, then $\tilde{d}^\tau = 0$ but not the other way around. Since $R(\theta, N)$ is decreasing in N, $d^\tau = 0$ ($R(\theta, N_{\tau^-}) \geq 0$) implies that $\tilde{d}^\tau = 0$ ($R(\theta, 0) \geq 0$). However, for any $\theta \notin \Theta^L \cup \Theta^U$, if there were attacks before time τ, then $\tilde{d}^\tau = 0$ but it is possible to have $d^\tau = 1$ (if N_{τ^-} is sufficiently large so that $R(\theta, N_{\tau^-}) < 0$). Note that option value argument (Lemma 1) holds true under $\tilde{\Gamma}^\tau$. However, the regime may still change even after $\tilde{d}^\tau = 0$ if some agents play strategy $\mathcal{A}$ and attack before the time of disclosure (since $R(\theta, 0) \geq 0$ cannot guarantee $R(\theta, N_T = N_{\tau^-}) \geq 0$). As such, $\tilde{d}^\tau = 0$ cannot predict the survival of the regime; that is, Lemma 2 fails under policy $\tilde{\Gamma}^\tau$.

Next, consider the forecasting property. Suppose that the principal cannot forecast at all. Then, the timely disaster alert policy boils down to the following: The principal sends the message $\hat{d}^\tau = 1$ if and only if the regime change has occurred before time τ; that is, $t_c < \tau$ or, equivalently, $R(\theta, N_{\tau^-}) < 0$ and $t_s < \tau$. We will refer to this policy as $\hat{\Gamma}^\tau$. Note that Lemma 1 holds true under $\hat{\Gamma}^\tau$. However, $\hat{d}^\tau$ could be 0 just because the shock has not arrived yet (i.e., $t_s \geq \tau$), but it will occur once the shock arrives at a future date. Thus, $\hat{d}^\tau = 0$ does not guarantee no future regime change, even without any further attack; that is, Lemma 2 fails under policy $\hat{\Gamma}^\tau$.

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. However, an

[14] 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.

agent who waits for disclosure may end up attacking the regime at a later date. Such a delay in attacking is costly, and is 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 $\mathcal{W}$ and wait if he believes that the alert is likely to be triggered.

Lemma 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. 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,

$$\mathbb{B}(\Gamma^\tau, x_{-i}, s_i) = \pi(\mathbb{T}, \infty) - \pi(0, \infty) = e^{-\beta T}\overline{V} - \overline{U}. \tag{2}$$

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 1).

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[15]

$$\begin{aligned}\mathbb{E}(\pi(\tau, t_c)|x_{-i}, s_i, d^\tau = 1) =& e^{-\beta\tau} \int_{t_s=0}^{T} \left(\mathbb{1}\{t_s \geq \tau\}\overline{U} + \mathbb{1}\{t_s < \tau\}\underline{U}\right) \mathrm{d}G(t_s) \\ =& e^{-\beta\tau} \left((1 - G(\tau^-))\overline{U} + G(\tau^-)\underline{U}\right). \end{aligned} \tag{3}$$

Accordingly, the expected cost of waiting is

$$\begin{aligned}\mathbb{C}(\Gamma^\tau, x_{-i}, s_i) :=& \mathbb{E}\left[\pi(0, t_c) - \pi(\tau, t_c)\middle| x_{-i}, s_i, d^\tau = 1\right] \\ =& (1 - G(\tau^-))(1 - e^{-\beta\tau})\overline{U} + G(\tau^-)(\overline{U} - e^{-\beta\tau}\underline{U}). \end{aligned} \tag{4}$$

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 $\mathbb{C}(\Gamma^\tau, x_{-i}, s_i)$ falls as τ decreases. Moreover, both the loss of time value $(1 - e^{-\beta\tau})\overline{U}$ and the probability of a significant payoff drop $G(\tau^-)$ can be made negligibly small when τ is set sufficiently close to 0. Therefore, for any agent i, regardless of his signal s_i and others' rationalizable strategy x_{-i}, when τ decreases to 0, the expected cost is $\mathbb{C}(\Gamma^\tau, x_{-i}, s_i) \to 0$.[16]

The Dominance of $\mathcal{W}$ over $\mathcal{A}$ It follows from Assumption 2 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, 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,

$$\begin{aligned}\mathbb{D}(\Gamma^\tau, x_{-i}, s_i) :=& \mathbb{P}(d^\tau = 0|x_{-i}, s_i)\mathbb{B}(\Gamma^\tau, x_{-i}, s_i) - \mathbb{P}(d^\tau = 1|x_{-i}, s_i)\mathbb{C}(\Gamma^\tau, x_{-i}, s_i) \\ \geq& \epsilon \cdot \mathbb{B}(\Gamma^\tau, x_{-i}, s_i) - (1 - \epsilon) \cdot \mathbb{C}(\Gamma^\tau, x_{-i}, s_i). \end{aligned} \tag{5}$$

[15] Notice that if $t_s \geq \tau$, then the regime does not change when the agent attacks at time τ.

[16] Since $G(0) = 0$ and $G(.)$ is right continuous (since a cumulative distribution function), as $\tau \to 0$, $G(\tau^-) \to 0$. If we consider an alternative setting in which the shock can arrive before the agents get the first opportunity to attack, then $\lim_{\tau \to 0} \mathbb{C}(\Gamma^\tau, s_i) = 0$ simply follows from right continuity of CDF G, and the $G(0) = 0$ assumption is not necessary. This is because the probability that an agent misses the opportunity to attack a regime before it changes while waiting for the alert is, at most, $G(\tau^-) - G(0)$ instead of $G(\tau^-)$.

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 $\mathcal{A}$.

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 afterwards. 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, the regime with $\theta \notin \Theta^L$ will never change. Therefore, any timely disaster alert policy Γ^τ with $\tau \in (0, \hat{\tau})$ eliminates the panic in the robust sense –i.e., $\Theta^P(\Gamma^\tau) = \emptyset$.

Remark 1 (Perfect Coordination) *When the disclosure is set timely, agents coordinate not only on following the disaster alert (Lemma 1), but also on waiting for it (Lemma 3). Recall that our flexible ex-ante information structure allows the agents to get different signals. In that case, they have heterogeneous beliefs about the fundamental. However, under the timely disaster alert, they perfectly coordinate their actions: when $\theta \in \Theta^L$, they all attack at time τ, and when $\theta \notin \Theta^L$, they never attack.*

Remark 2 (Principal's Preference and Commitment) *For our principal who concerns only the regime outcome, the ex-ante commitment is not necessary. To see this, consider that at disclosure time τ, if the principal observes that $R(\theta, N_{\tau^-}) < 0$, 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 the 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 signals before time τ). However, when the principal has more-nuanced preferences, she may need commitment power to implement EWS or may not want EWS in the first place. We discuss such preferences in Section 3.1.*

2.4 Two Essential Properties of the Disaster Alert

Endogenous Disclosure

As we mentioned before, unlike the static disclosure policies (Goldstein and Huang (2016), Inostroza and Pavan (2020), Li, Song and Zhao (2019)), a disaster alert sends different messages depending on how many agents endogenously choose to wait for this future disclosure. However, notice that under a timely disaster alert, since all the agents wait for the disclosure ($N_{\tau^-} = 0$), when $d^\tau = 0$, on path, the agents learn only that $R(\theta, 0) \geq 0$. Recall that policy $\tilde{\Gamma}^\tau$ also discloses the same information $R(\theta, 0) \geq 0$ or not at time τ (See the discussion after Lemma 2). However, as we argued earlier, since $\tilde{\Gamma}^\tau$ is not an endogenous disclosure policy, it cannot predict the regime outcome (Lemma 2 fails). To see this, note that if most agents do not wait for the disclosure and attack (choose $\mathcal{A}$ over $\mathcal{W}$), then even after $\tilde{d}^\tau = 0$, the regime can still change without any future attack. Therefore, it may not be profitable to wait for $\tilde{d}^\tau$ no matter how early the disclosure is scheduled. Unlike Γ^τ, the non-history-dependent policy $\tilde{\Gamma}^\tau$ cannot eliminate panic in the robust sense.[17]

Forecasting

When the principal is not capable of forecasting, the disaster alert boils down to the $\hat{\Gamma}^\tau$ policy, and the perfect predictability fails (See the discussion after Lemma 2). Moreover, the alert $\hat{d}^\tau = 1$ means that the regime has already changed, and, thus, it is too late for the agent who has waited for the disclosure to attack. He can get only $\underline{U}$ from attacking after $\hat{d}^\tau = 1$. This makes waiting substantially costly.[18]

This shows the contrast between the disclosure policy in this paper and that in the existing literature, such as Angeletos, Hellwig and Pavan (2007) and Basak and Zhou (2020). In these papers, the agents publicly learn about whether or not the regime has already changed. As the above discussion shows, disclosing the information about the regime change after the event may not dissuade the agents from attacking when they choose when to attack. However, Basak and Zhou (2020) shows that when the agents move sequentially in an exogenous order and frequently learn about whether or not the regime has changed yet (the alerts $\hat{d}^\tau$ are disclosed after small time intervals), there is a unique cutoff equilibrium where the agents do not attack a viable regime.[19]

[17] To see this formally, suppose that an agent believes that the regime will survive with probability $P(t_c = \infty | x_{-i}, s_i, \tilde{d}^\tau = 0) = p$. 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 $\mathcal{W}$, there is always a strictly positive benefit of waiting for d^τ.

[18] Formally, the cost of waiting is $\mathbb{C}(\hat{\Gamma}^\tau, x_{-i}, s_i) = \mathbb{E}\left(\pi(0, t_c) - \pi(\tau, t_c) | x_{-i}, s_i, \hat{d}^\tau = 1\right) = \overline{U} - e^{-\beta\tau}\underline{U}$. Since $\overline{U} > \underline{U}$, unlike the cost of waiting under Γ^τ (See $\mathbb{C}(\Gamma^\tau, x_{-i}, s_i)$ in equation (4)), this cost is bounded away from 0 regardless of the disclosure time τ.

[19] Basak and Zhou (2020) relies on a more restrictive environment in which the agents are endowed with dispersed private information (as in the canonical global games) and are restricted to playing monotone strategies. Note that we

2.5 Observable Regime Change

In many applications (e.g., bank runs), it could be natural that the agents can observe when the regime changes. Next, we 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.[20] 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.

2.6 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.

Formally, 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

do not impose such restrictions.

[20] More precisely, the second term in the expected cost of waiting $\mathbb{C}$ is $G(\tau-)\overline{U} - \int_0^{\tau-} e^{-\beta t} \underline{U} \mathrm{d}G(t)$, which is smaller than $G(\tau-) \left(\overline{U} - e^{-\beta\tau}\underline{U}\right)$ ((see (4)).

timely disaster alert every time after the new information arrives. That is, the k-th disaster alert is triggered — i.e., $d^{t_k+\tau} = 1$, if and only if $R(\theta, N_{t_k+\tau^-}) < 0$, where $k = 0, 1, \dots K$.

Theorem 2 *If $G(.)$ 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 Theorem 2 states, however, as long as the principal can set a timely disaster alert every time after the new information arrives, 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 never attack unless any alert is triggered strictly dominates any strategy involving waiting for the alerts before time t_k ($k = 0, 1, \dots, K$), then attacking at time t_k but not waiting for the next alert. 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 EWS is continuously in place, panic will be eliminated.

3 Discussion

Our proposed dynamic information disclosure policy resembles an EWS. We have seen that this simple policy can be remarkably effective in reducing the chance of panic. However, it is important to understand that the success of the EWS relies on a set of conditions that may not always hold true in practice. First, it may not always be in the principal's interest to implement an EWS. Second, it may not be feasible to implement such a policy. Finally, the strategic environment could be such that an early warning does not help. Below, we shed light on the limitations and possible improvements of EWS.

3.1 Conflict of Interest

We assume that the principal has a simple preference — she wants the regime to survive. This preference specification is consistent with that of international organizations trying to reduce the panic associated with sovereign debt defaults, or of bank supervisors and regulators trying to stop runs on healthy banks. However, for other applications, the principal's preference could be more nuanced. Below, we discuss when such a different preference supports an EWS and when it is in conflict with an EWS.

First, recall that under a timely disaster alert, when $\theta \in \Theta^L$, agents attack at time τ, and the regime changes instantaneously after the adverse shock arrives ($t_c = t_s^+$). On the other hand, if $\theta \notin \Theta^L$, no one attacks and the regime survives. Note that a timely disaster alert minimizes the size of the attack against any healthy regime. Therefore, it remains the optimal policy if the principal not only cares about the survival of the regime, but also gets a higher payoff when the size of the attack on a healthy regime is smaller. Moreover, since when $\theta \in \Theta^L$, the regime changes at $t_c = t_s^+$ regardless of what agents do. The same policy remains optimal if the principal derives a positive flow utility for every moment that the regime survives.

Second, consider a principal who wants to maximize the welfare of the agents. Note that when $\theta \in \Theta^L$, the agents bear an unnecessary cost while waiting for the alert. To avoid this cost, the principal should make the disclosure time τ as early as possible (but still after time 0).[21]

Third, consider a principal who wants to minimize the size of the attack or to delay the attack as much as possible even when regime change is inevitable. In this case, the principal's preference conflicts more with that of the agents. Ely (2017) builds an interesting example in which a bank (principal) knows that it will fail but wants to maximize the time when the last depositor runs. Since an early warning induces early attacks when regime change is inevitable, it is not a desirable policy for such a principal. Ely (2017) argues that in such a case, the principal should send private messages to generate mi-coordination among agents.

3.2 Error in Forecasting

The effectiveness of the EWS relies on the principal's ability to forecast the preparedness of the regime to face some future shock. Building an EWS requires due diligence. This is largely consistent with the forward-looking bank stress tests and the IMF's and World Bank's debt sustainability analysis in practice.[22] It is possible that such forecasts are not perfectly accurate.

[21]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). .

[22]In practice, 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. 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 IMF, see https://www.imf.org/external/pubs/ft/dsa/lic.htm. For econometric methods adopted

A false-negative happens when the alert is triggered ($d^\tau = 1$) but $R(\theta, N_{\tau^-}) \geq 0$; and 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 valuable. A false-negative will be disastrous since the agents will all attack, 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. This could generate false-positives more often. Although a false-positive does not entail the same risk, if there is a high probability of a false-positive, the agents may ignore the disclosure and act based on their own signals. Accordingly, there could be panic. The following proposition shows that timely persuasion is effective as long as there is no false-negative and the probability of false-positive is sufficiently small.

Proposition 1 *Suppose that there is no false-negative, and 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 as long as the incidence of false-positive is small enough.[23]

3.3 Essential Properties of the Strategic Environment

The dynamic regime change game we consider has some distinctive properties. Next, we discuss whether these properties are essential or can be relaxed.

3.3.1 Irreversible Attacks

There is an 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. 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

in the development of an EWS for sovereign debt crises, see Dawood, Horsewood and Strobel (2017).

[23] A false-positive may also arise when some agents enter the game at a later date after the disclosure, and attack even though the alert has not been triggered. The same argument applies to such environments.

away and reverse the action only if the alert is not triggered. This could trigger the alert, although the regime change was not warranted, and, thus, panic is not eliminated.[24]

3.3.2 Doubt

Throughout this paper we 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 homogeneous belief about the fundamental. If the agents do not doubt and believe that the regime changes for sure 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, they 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. 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 $\mathscr{U}[-\infty, \infty]$, and agents receive 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.[25] However, the following result shows that a timely disaster alert eliminates panic in the limit.

Proposition 2 *In the above specialized environment, given a disaster alert Γ^τ, under the unique rationalizable strategy, the regime survives iff $\theta \geq \theta^*(\tau)$; that is, $\Theta^P(\Gamma^\tau) = [0, \theta^*(\tau))$, where*

$$\theta^*(\tau) = \frac{\overline{U} - e^{-\beta\tau}\left(G(\tau^-)\underline{U} + (1 - G(\tau^-)\overline{U}\right)}{e^{-\beta T}\overline{V} - e^{-\beta\tau}\left(G(\tau^-)\underline{U} + (1 - G(\tau^-)\overline{U}\right)} \in (0, 1).$$

For any $\tau > 0$, $\theta^*(\tau)$ is lower than the threshold under no disclosure policy, and $\theta^*(\tau)$ declines as τ decreases. For any $\zeta > 0$ (however small), we can find $\hat{\tau}(\zeta)$ such that, under the policy Γ^τ with $\tau \in (0, \hat{\tau}(\zeta))$, $\theta^*(\tau) < \zeta$.

[24] Also, note that 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)).

[25] 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).

The formal proof is relegated to the Appendix. To understand the intuition, first consider the no disclosure policy. 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) shows that for each type s_i, a unique strategy survives the iterated elimination of dominated strategies (IESDS) — attack for $s_i < s^*$ and not attack for $s_i \geq s^*$. This implies that the regime survives iff the fundamental is beyond a cutoff. Here, the only 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. Thus, attacking right away becomes relatively less attractive than under the static regime change game. Therefore, agents are less likely to attack the regime at time 0 ($\mathcal{A}$) when the waiting option is available. Accordingly, regardless of disclosure time τ, after IESDS, the fundamental cutoff $\theta^*(\tau)$ is lower than the cutoff under the static game. As τ falls, the expected payoff from $\mathcal{W}$ further increases, and, accordingly, $\theta^*(\tau)$ falls. As shown in the above Proposition, 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.

3.3.3 Negligible Waiting Cost for a Timely Alert

We consider an economic environment in which the cost of waiting for a timely disclosure can be made negligible (See Lemma 3). It is not hard to think of situations where waiting for even a short time could be significantly costly. For instance, suppose there is a significant mass probability that the shock will arrive before any disclosure ($G(0) > 0$) and the regime can change instantaneously. Given the first-mover advantage ($\overline{U} > \underline{U}$), regardless of how timely the disaster alert, the agents will find it very costly to wait for it (See equation (4)). When this cost is significant, the agent may not wait for an EWS and panic may persist.

However, even when the agents are worried about the significant chance of the shock arriving while they wait for the disclosure ($G(0) > 0$), an EWS may work. This could happen when, for instance, other regulations in the market remove the first-mover advantage, or because the market moves slowly and the regime does not change instantaneously.

First-mover Advantage Waiting can induce a significant payoff drop if the agent, while waiting, misses the opportunity of attacking before the regime change. This captures "first-mover advantage," which makes waiting more costly and challenges the efficacy of EWS.

In reality, some existing regulatory policies and bankruptcy codes are designed precisely to remove this first-mover advantage. For example, money market fund managers are allowed to impose redemption fees on early redemption and transfer that to the staying fund investors.[26] The redemption fees can potentially cancel out the negative externality exerted by the earlier redemption on later withdrawal and eliminate the first-mover advantage. Another example is the bankruptcy code known as "avoidable preference," which prevents creditors from being treated more favorably than others.[27] It is implemented by calling back debt repayments made shortly before bankruptcy, and the proceeds are then shared among all creditors in bankruptcy court. Under this policy, withdrawing investments shortly before bankruptcy is not advantageous.

Under these regulations, there is no first-mover advantage and thus, $\overline{U} = \underline{U}$. This makes the waiting cost continuous in time (See equation (4)) and independent of the distribution of the shock arrival time ($G(.)$). Thus, regardless of how likely the shock arrives before the earliest possible disclosure ($G(0)$), the cost of waiting for a timely disclosure is negligible. Accordingly, the agents will wait for the alert rather than panicking.

Slow-moving Capital In our benchmark model, we assume that attacks are instantaneous. In practice, moving investments away from a country or withdrawing funds from financial institutions often takes time. This could be due to 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, it can happen simply because moving capital away, which involves selling off properties and firing employees, is time-consuming.

Next, we consider an environment in which attacks are not instantaneous, and show that the waiting cost for a timely disaster alert can be made negligible under slow-moving capital, regardless of the distribution of shock arrival time $G(.)$. As before, the agents choose when to attack and 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 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 $\underline{u} \in (0, \overline{u})$ if 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 $B(.)$ has support $[0, 1]$ and it admits a density $b(.)$, which is Lipschitz continuous. This generalizes our benchmark model, which becomes a special case with $l \to 0$. We assume the complementarity (Assumption 1) with regard to the flow payoffs.

Proposition 3 *If $l > 0$ and $b(.)$ is Lipschitz continuous, then regardless of $G(.)$, there exists $\hat{\tau}_l \in (0, l]$ such that under disaster alert Γ^τ with $\tau \in (0, \hat{\tau}_l)$, $\Theta^P(\Gamma^\tau) = \emptyset$.*

Note that the above Proposition allows the shock to arrive before any disclosure with significant

[26]For the detailed rules of imposing redemption fees, see the SEC's 2014 amendments to the rules that govern money market mutual funds (https://www.sec.gov/rules/final/2014/33-9616.pdf).

[27]For details, see Sections 547 and 550 in Chapter 11 of the U.S. Bankruptcy Codes.

probability (i.e., $G(0) > 0$). Recall that, in this case, waiting can be very costly if the regime changes instantaneously. However, when the market moves slowly, the waiting cost for a timely disclosure can be made negligible. To see this argument, consider the special case $B(.) = \mathscr{U}[0,1]$. 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 $\mathcal{W}$. Therefore, conditional on $d^\tau = 1$, the cost of waiting $\mathbb{C}(\mathit{\Gamma}^\tau, x_{-i}, s_i)$ is no higher than the flow payoff he loses in the interval $[0, \tau)$ — that is, $\frac{\overline{u}\tau}{l}$. Therefore, as $\tau \to 0$, waiting for the timely disaster alert is almost costless. Accordingly, following the same insight from Theorem 1, we can show that a timely disaster alert eliminates panic in the robust sense.

4 Conclusion

We model panic by extending the classical regime change game in a dynamic setup, where (1) a shock arrives at a stochastic date, and (2) the agents choose when to attack. We study the robust dynamic information design problem. The agents are privately informed, and we allow for a reasonably general exogenous information structure. The principal wants to dissuade the agents from attacking. We construct a simple policy — a timely disaster alert. Under this simple policy, the agents' unique rationalizable strategy is to wait for and then follow the alert, regardless of their signals. This eliminates panic in a robust sense.

Our proposed policy resembles an EWS such as a debt sustainability analysis or forward-looking bank stress test. In practice, these policies are adopted to caution agents about an impending crisis. This paper provides a rationale for adopting such an EWS and emphasizes the importance of "endogenous disclosure" and "forecasting." The effectiveness of the EWS hinges not only on being able to persuade the agents to follow the warning, but also on being able to persuade them to wait for the warning. From that perspective, "timely disclosure" is also critical. Anderson (2016) provides some anecdotal evidence demonstrating the benefit of conducting bank stress tests in a timely manner. The author shows that forward-looking stress tests were quite effective in the U.S. but not so much in Europe. The author argues that this contrasting experience can be attributed to the fact that the U.S.'s stress test attempt was "timely," while it was, perhaps, "too late" in Europe.[28] In addition, we discuss the crucial features of the environment that make early warning an effective information disclosure policy. We also point out when the EWS can be unreliable and how other supplementary policies and regulations can facilitate its efficacy.

[28] The first bank stress test in the U.S. was conducted in a timely manner. The plan for stress testing was announced on February 10, 2009. The white paper describing the procedures employed in SCAP was released on April 24, 2009 and the results of the SCAP were disclosed on May 7, 2009. On the other hand, although signs of instability in the financial system became apparent around the same time as in the U.S., the first Europe-wide stress test were conducted in October 2009 and the bank-level results were to remain confidential.

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.

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.

Che, Yeon-Koo, and Johannes Hörner. 2018. "Recommender systems as mechanisms for social learning." *The Quarterly Journal of Economics*, 133(2): 871–925.

Cole, Harold L, and Timothy J Kehoe. 2000. "Self-fulfilling debt crises." *The Review of Economic Studies*, 67(1): 91–116.

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 Ady Pauzner. 2005. "Demand–deposit contracts and the probability of bank runs." *Journal of Finance*, 60(3): 1293–1327.

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.

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.

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*.

Rochet, Jean-Charles, and Xavier Vives. 2004. "Coordination failures and the lender of last resort: was Bagehot right after all?" *Journal of the European Economic Association*, 2(6): 1116–1147.

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 $t_c(d^\tau, x_{-i})$ 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 2, 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 $\overline{U} > 0$, $\mathbb{E}[u_i((t_\emptyset = 0, t_0', t_1'), x_{-i})|s_i] > \mathbb{E}[u_i((t_\emptyset, t_0, \tau), x_{-i})|s_i]$.[29]

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 2. ■

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 (4), 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 (5), the expected payoff difference is at least

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

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, $\Theta^F(\Gamma^\tau) = \Theta^L$ and $\Theta^P(\Gamma^\tau) = \emptyset$. ■

[29] 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.

Proof of Theorem 2. Under the extended Γ^τ 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 (2), the benefit of $\mathcal{W}^{K+1}$ compared to $\mathcal{W}^K$, conditional on $d^{t_K+\tau} = 0$, is

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

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

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

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 $\mathit{\Gamma}^\tau$ 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 $\mathit{\Gamma}^\tau$ 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 1. 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

$$\mathcal{P}^1(s_i, x_{-i}) e^{-\beta\tau} \left(G(\tau^-)\underline{U} + (1 - G(\tau^-))\overline{U}\right) + \mathcal{P}^0(s_i, x_{-i}) e^{-\beta T}\overline{V} + \mathcal{P}^e(s_i, x_{-i}) e^{-\beta T}\underline{V}. \quad \text{(A.3)}$$

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 1, 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 2, $\mathcal{P}^0(s_i, x_{-i}) \geq \mathbb{P}(\theta \in \Theta^U | s_i) > \epsilon$ 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 $\tau \in (0, \hat{\tau}(\eta))$. ■

Proof of Proposition 2. Under a disaster alert Γ^τ, since $\mathbb{P}(\theta \in \Theta^L | s_i) > 0$ for all s_i, following Lemma 1, the rationalizable strategies are $\mathcal{A}$ and $\mathcal{W}$.[30] Thus, the game boils down to a binary action regime change game similar to Morris and Shin (2003) (henceforth MS). Recall

[30]Note that this is the only condition on ex-ante information that is required for the proof of Lemma 1 (See footnote 29).

that under no disclosure policy, the game is exactly as the static game in MS, where each agents decides whether to "attack" or "not attack."

Similar to attacking in the static game, if an agent plays $\mathcal{A}$, he gets $\overline{U}$. If N_0 people plays $\mathcal{A}$, then the regime survives if $R(\theta, N_0) = \theta - N_0 \geq 0$, and it changes otherwise. Therefore, under no disclosure policy, if an agent does not attack, he gets $e^{-\beta T}\overline{V}$ if the regime ultimately survives, and $e^{-\beta T}\underline{V}$ otherwise. Thus, under no disclosure, an agent will not attack if and only if the regime survives with probability

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

Compared to the "not attack" strategy, an agent who plays $\mathcal{W}$, gets the warning about the impending regime change $d^\tau = 1$ (i.e., $\theta < N_0$) and can attack after learning so. As in Lemma 2, when $d^\tau = 0$, the regime survives ($t_c = \infty$) and when $d^\tau = 1$, the regime changes as soon as the shock arrives. Thus, conditional on $d^\tau = 0$, and the agent who chooses $\mathcal{W}$ gets $e^{-\beta T}\overline{V}$. However, conditional on $d^\tau = 1$, the agent who plays $\mathcal{W}$ gets the expected payoff

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

It is clear that this payoff is strictly decreasing in τ, and it is greater than the payoff from "not attack" a failed regime in the static regime change game, $e^{-\beta T}\underline{V}$. Therefore, an agent plays $\mathcal{W}$ if and only if he believes that $d^\tau = 0$, and, accordingly, the regime survives (i.e., $\theta \geq N_T = N_0$) with probability

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

It is easy to see that $p^*(\tau) < p^*$ and $p^*(\tau)$ decreases as τ decreases.

Following the standard global game argument, 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 iff $\theta < \theta^*(\tau)$. When others follow such a monotone strategy, the agent who receives higher signal is more optimistic about the regime surviving. The marginal agent $s^*(\tau)$ must be indifferent between $\mathcal{A}$ and $\mathcal{W}$. This gives us $\mathbb{P}(\theta \geq \theta^*(\tau) | s^*(\tau)) = p^*(\tau)$. Given uniform prior and Gaussian noise, $\theta | s_i \sim N(s_i, \sigma^2)$. Therefore, the above indifference condition simplifies to

$$\Phi\left(\frac{s^*(\tau) - \theta^*(\tau)}{\sigma}\right) = p^*(\tau). \tag{A.4}$$

Moreover, when other agents play such monotone strategy, there is less attack when θ is higher. The fundamental cutoff $\theta = \theta^*(\tau)$ is such that $N_0(\theta^*(\tau)) = \mathbb{P}\left(s < s^*(\tau) | \theta = \theta^*(\tau)\right) = \theta^*(\tau)$.

This simplifies to

$$\Phi\left(\frac{s^*(\tau)-\theta^*(\tau)}{\sigma}\right)=\theta^*(\tau). \tag{A.5}$$

Combining the above two equations (A.4) and (A.5), we get that $\theta^*(\tau) = p^*(\tau)$. Therefore, $\Theta^P(\Gamma^\tau)=[0,p^*(\tau))$. Recall that, under no disclosure policy, $\Theta^P=[0,p^*)$. Since $p^*(\tau)<p^*$, for any τ, a disaster alert Γ^τ reduces the chance of panic compared to no disclosure.

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 3. 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 $M_t := \int_{x\le t} B\left(\frac{t-x}{l}\right)\mathrm{d}N_x$. By definition, $M_t\le N_t$ and $M_{T+l}=N_T$. The regime changes at the first instance when $R(\theta,M_t)<0$, and if $R(\theta,M_{T+l})\ge 0$, then the regime never changes. Accordingly, we redefine t_c as follows.

$$t_c := \begin{cases}\min\{t\in[t_s^+,\infty)|R(\theta,M_t)<0\} \text{ if } R(\theta,M_{T+l})<0\\ \infty \text{ if } R(\theta,M_{T+l})\ge 0.\end{cases} \tag{A.6}$$

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

$$\pi(t,t_c)=e^{-\beta t}\int_0^l e^{-\beta z}\left[\mathbb{1}\{t+z\ge t_c\}\underline{u}+\mathbb{1}\{t+z<t_c\}\bar{u}\right]\mathrm{d}B(\frac{z}{l}). \tag{A.7}$$

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 1) dictates that the flow payoffs satisfy $e^{-\beta T}\overline{v}>\overline{u}>\underline{u}>\underline{v}\ge 0$.[31]

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 the agents may start attacking at time 0, the regime may not change as

[31] 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 ($l=0$).

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

$$\mathbb{B}^S(\Gamma^\tau, x_{-i}, s_i) := \mathbb{E}[\pi(\mathbb{T}, t_c) - \pi(0, t_c)|d^\tau = 0, x_{-i}, s_i] = \left(e^{-\beta T}\bar{v} - \bar{u}\right) \int_0^l e^{-\beta z} \mathrm{d}B(\frac{z}{l}) \quad \text{(A.8)}$$

Note that $\mathbb{B}^S(\Gamma^\tau, x_{-i}, s_i) > 0$ 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

$$\mathbb{E}\left[\pi(0, t_c)|d^\tau = 1, x_{-i}, s_i\right] = \int_0^l e^{-\beta z} \left[\mathcal{P}(z)G(z^-)\underline{u} + (1 - \mathcal{P}(z)G(z^-))\overline{u}\right] \mathrm{d}B\left(\frac{z}{l}\right), \quad \text{(A.9)}$$

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

$$\begin{aligned}\mathbb{E}\left[\pi(\tau, t_c)|d^\tau = 1, x_{-i}, s_i\right] &= \int_0^l e^{-\beta(z+\tau)} \left[\mathcal{P}(\tau + z)G(z + \tau^-)\underline{u} + (1 - \mathcal{P}(\tau + z)G(z + \tau^-))\overline{u}\right] \mathrm{d}B\left(\frac{z}{l}\right) \\ &= \int_\tau^{l+\tau} e^{-\beta z} \left[\mathcal{P}(z)G(z^-)\underline{u} + (1 - \mathcal{P}(z)G(z^-))\overline{u}\right] \mathrm{d}B\left(\frac{z - \tau}{l}\right) \\ &\geq \int_\tau^l e^{-\beta z} \left[\mathcal{P}(z)G(z^-)\underline{u} + (1 - \mathcal{P}(z)G(z^-))\overline{u}\right] \frac{1}{l} b\left(\frac{z - \tau}{l}\right) \mathrm{d}z. \quad \text{(A.10)}\end{aligned}$$

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.9) and (A.10) that conditional on $d^\tau = 1$, x_{-i}, and s_i, the cost of playing $\mathcal{W}$ rather than $\mathcal{A}$, denoted by $\mathbb{C}^S(\Gamma^\tau, x_{-i}, s_i)$, 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 $b(.)$ 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 $\overline{b} > 0$ such that

$b\left(\frac{z}{l}\right) \leq \bar{b}$. Therefore, for any s_i and $\tau \in (0, l)$,

$$\mathbb{C}^S(\varGamma^\tau, x_{-i}, s_i) \leq \overline{u}\left(\int_\tau^l \frac{1}{l}\kappa\frac{\tau}{l}\mathrm{d}z + \frac{\tau}{l}\bar{b}\right) < \frac{\overline{u}(\bar{b}+\kappa)}{l}\tau. \tag{A.11}$$

The first inequality follows immediately from the above discussions, and the second inequality relies on $\tau > 0$. Note that the expected cost $\mathbb{C}^S(\varGamma^\tau, x_{-i}, s_i)$, 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.8), (A.11) and Assumption 2, the expected payoff difference between $\mathcal{W}$ and $\mathcal{A}$ is

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

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 $\varGamma^\tau$ with any $\tau \in (0, \hat{\tau}_l)$, $\mathbb{D}(\varGamma^\tau, x_{-i}, s_i) > 0$ regardless of s_i and other agents' choices between $\mathcal{A}$ and $\mathcal{W}$, thereby proving the dominance of $\mathcal{W}$ over $\mathcal{A}$. ■