Diffusing Coordination Risk

By Deepal Basak and Zhen Zhou

Online Appendix

One time Viability Test

Violation of Acharya and Ramsay (2013) sufficient condition

Acharya and Ramsay (2013) (henceforth AR) consider a global game of regime change game with one-sided dominance. The model in AR can be mapped to our setting by interpreting Cooperate as Attacking ($a_i = 0$) and Defect as Not Attacking ($a_i = 1$). In AR, there is no dominance region where an agent will Cooperate regardless of the opponent's action. Hence, it is similar with PNV which eliminates the dominance region of attacking. Also, in AR, a lower signal will encourage the agents to Defect while in our setting a higher signal encourages agents to Not attack. Hence, the actual inequality in Condition (AR) below is reverse in their paper. Finally, the payoff specification is such that their $\frac{b_i}{b_i+w_i} = p$ in out setting.

Mapping AR's Assumption A1(ii) to our model, we have the following condition. For any $\hat{s}$,

$$\text{(AR)} \qquad \mathbb{P}(s_{-i} \geq \hat{s} | s_i = \hat{s}, \theta \geq 0) > p$$

which guarantees that no attack is the unique cutoff equilibrium when there are two players. The intuition is simple. Let us define the set of private signal at which an agent may attack -

$$A_i = \{s_i | \text{ Attacking } (a_i(s_i) = 0) \text{ is not eliminated as never best response}\}.$$

Take $\hat{s} = \sup_i \cup A_i$. Then for any agent receiving private information $s > \hat{s}$ will not attack because attacking is eliminated as a dominated strategy. Think about any agent with private information $\hat{s}$. Given Condition (AR), the probability that the other agent receives a signal higher than $\hat{s}$ and thus does not attack is higher than p. Thus, not attacking is strictly preferred, which violates the definition of $\hat{s}$. That is why no attacking is the only possible equilibrium.

We take the standard global game information structure and assume that F is log-concave. We did not assume the Condition (AR). In our setting, what we need is that for any $\hat{s}$,

$$\text{(CI)} \qquad \mathbb{P}(\text{succ when agents follow cutoff strategy } \hat{s} | s_i = \hat{s}, \theta \geq 0) > p,$$

where "succ" refers to the success of the regime. We argue in lemma 1 that when α is small enough, the above inequality (CI) holds. However, as shown below, condition (AR) may not hold true under the assumption of small group size.

Claim A.1 *For any α and $p > 1/2$, there exists $\hat{s} < \alpha + \sigma/2$ such that*

$$\mathbb{P}(s_{-i} \geq \hat{s} | s_i = \hat{s}, \theta \geq 0) \leq p.$$

PROOF:

$$\begin{aligned}
& \mathbb{P}(s_{-i} \geq \hat{s} | s_i = \hat{s}, \theta \geq 0) \\
= & \int_{\max\{0,\hat{s}-\sigma/2\}}^{\hat{s}+\sigma/2} \mathbb{P}(s_{-i} \geq \hat{s} | \theta) \mathbb{P}(\theta | s_i = \hat{s}, \theta \geq 0) d\theta \\
= & \int_{\max\{0,\hat{s}-\sigma/2\}}^{\hat{s}+\sigma/2} (1 - F(\sqrt{\tau}(\hat{s} - \theta))) \cdot d \left(\frac{1 - F(\sqrt{\tau}(\hat{s} - \theta))}{F(\sqrt{\tau}\hat{s})} \right) \\
= & \frac{1}{2F(\sqrt{\tau}\hat{s})} \cdot \left[(1 - F(\sqrt{\tau}(\hat{s} - \theta)))^2 \right]_{\max\{0,\hat{s}-\sigma/2\}}^{\hat{s}+\sigma/2} \\
= & \begin{cases} \frac{1}{2}, \text{ for } \hat{s} \in [\sigma/2, 1 + \sigma/2] \\ 1 - \frac{F(\sqrt{\tau}\hat{s})}{2} \text{ for } \hat{s} \in [-\sigma/2, \sigma/2] \end{cases}
\end{aligned}$$

When $p > 1/2$, there is some $\hat{s} \in [\sigma/2, \alpha + \sigma/2]$ for which condition (AR) is violated.[1] It holds true only when $\hat{s} < \sigma F^{-1}(2(1 - p))$. □

In the absence of viability test, an agent believes that half of the agents have received higher signal than him and the other half has received lower signals than him. PNV does not change this belief uniformly. For an agent who receives a low private signal, PNV convinces him that there is a larger fraction of other agents who receive higher signals than him. For example, consider the extreme case - an agent who receives signal $s_i = -\sigma/2$. Upon receiving the PNV, he understands that $\theta = 0$ and all other agents will receive higher private signals. On the other hand, PNV will not have an impact on the belief of an agent with private signal $s_i \geq \sigma/2$ regarding other agents' signals. He always believes that 1/2 mass of agents have received higher signal. Thus, the Condition (AR) is restrictive and does not hold for all $\hat{s}$ when $p > 1/2$.

It is important to point out that a small group size α does not affect agent i's belief about the private signal of agent j. As discussed in Section II, a small

[1] Note that if agents receive private information $s \geq \alpha + \sigma/2$, he understands that $\theta \geq \alpha$ and thus the regime succeeds even if all agents (with measure α) attack. In this case, the condition AR is redundant.

group size α reduces the success criteria. The combined effect of a sufficiently small group size and the PNV makes condition (CI) satisfied.

Repeated Viability Tests

Cutoff Equilibrium

Given a diffused policy J, a cutoff equilibrium specifies a threshold signal $\hat{s}_j$ for any $j = 1, 2 \dots J$, such that an agent i in group j attacks the regime that passes the jth viability test iff $s_i < \hat{s}_j$.

This implies that for any θ, the aggregate attack from any group j is $\alpha\mathbb{P}(s_i < \hat{s}_j|\theta) = \alpha F(\sqrt{\tau}(\hat{s}_j - \theta))$, which is decreasing in θ. Define $\underline{\theta}_{j-1}$ such that if $\theta \geq \underline{\theta}_{j-1}$, then the regime will pass the j-th viability test. Conversely, if $\theta < \underline{\theta}_{j-1}$, then the regime will fail at or before group j. Then $\underline{\theta}_{j-1}$ is such that

$$\text{(PNV)} \qquad \underline{\theta}_{j-1} - \alpha \sum_{j'=1}^{j-1} F(\sqrt{\tau}(\hat{s}_{j'} - \underline{\theta}_{j-1})) = 0.$$

Note that $\underline{\theta}_0 = 0$ and define

$$\text{(A.1)} \qquad \underline{\theta}_J = \hat{\theta}.$$

This means when $\theta = \hat{\theta}$, the regime has the exact strength to meet the aggregate attack from all groups. Hence, any regime with $\theta \geq \hat{\theta}$ will succeed.

Since viability tests remove the lower dominance region, it is always possible that in equilibrium the agents stop attacking a viable regime after the j-th test, for some $j = 1, 2, \dots, J$. Suppose that in equilibrium the agents attack a viable regime until group $\overline{j} \leq J$, for some $\overline{j}$. $\overline{j} = 0$ means no agent attacks the regime. If $\overline{j} > 0$, then $\underline{\theta}_{j-1}$ does not change after $\overline{j}$, i.e., once the regime survives the attack from the first $\overline{j}$ groups, it survives in the end. Then, any agent i in group $j \leq \overline{j}$ who receives the cutoff signal $\hat{s}_j$ must be indifferent between attacking and not attacking - i.e.,

$$(I^v) \qquad \mathbb{P}(\theta \geq \hat{\theta}|\hat{s}_j, \theta \geq \underline{\theta}_{j-1}) = \frac{F(\sqrt{\tau}(\hat{s}_j - \hat{\theta}))}{F(\sqrt{\tau}(\hat{s}_j - \underline{\theta}_{j-1}))} = p.$$

Thus, $\{\hat{s}_j\}_{j=1}^{J}$ constitutes a Perfect Bayesian Equilibrium in monotone strategies if there exists $\overline{j} \leq J$ such that

1) for all $j > \overline{j}$, $\hat{s}_j = \underline{\theta}_{\overline{j}} - \frac{\sigma}{2}$ and consequently, it follows from Condition (PNV) that $\underline{\theta}_{\overline{j}} = \underline{\theta}_{\overline{j}+1} = \dots = \underline{\theta}_J$;

2) for all $j \leq \overline{j}$, $\hat{s}_j$ satisfies the indifference condition (I^v), where the cutoffs $\{\underline{\theta}_{j-1}\}_{j=1}^{J+1}$ satisfy condition (PNV) and $\hat{\theta}$ satisfies equation (A.1).

Limited Diffusion

Repeated Stress Tests

Suppose that the principal has some capability constraint in the sense that she cannot repeat the viability test as many times as required to eliminate the coordination risk. Instead, assume that the principal can commit to the repeated stress tests instead of a repeated viability test. As mentioned before, Inostroza and Pavan (2017) or Goldstein and Huang (2016) argue that a one-time tough enough stress test can be persuasive. We will investigate the case in which the principal repeats such stress tests.

Definition A.1 *A J* ***times repeated*** *k**-stress test** publicly discloses at regular intervals of $\alpha = 1/J$ starting at 0, whether the residual per capita fundamental strength exceeds $k \in [0, 1)$ - i.e., if*

$$\frac{\theta - \alpha \sum_{l=1}^{j-1} w_l}{1 - (j-1)\alpha} \geq k$$

or not, where w_l is the proportion of agents who attack from group l.

Note that the toughness of the test k is on a per-capita basis. For example, if $J = 2$ ($\alpha = 0.5$) and $k = 0.2$, then the first stress test is about whether $\theta \geq k = 0.2$ or not, and the second one is about whether the residual strength $\theta - \alpha w_1 \geq \alpha k = 0.1$ or not, in which w_1 stands for the share of agents in the first group who decide to attack.

When there exist multiple cutoff equilibria, similar to the above mentioned papers, we assume that the principal designs her policy anticipating that the agents play the worst equilibrium in which agents take the most aggressive attacking strategy. Note that when agents publicly know that the regime fails the stress test, all attacking is the most aggressive attacking strategy and in that case, the regime fails for sure. Hence, as in the main paper, we will only focus on agents' strategies given that the regime has passed the test. If a J times repeated k-stress test is persuasive, then no agent in the first group attacks the regime when it passes the first stress test. This means that the regime will pass the second test and no agent will attack, and so on. Thus, a regime with fundamental strength $\theta \geq k$ will succeed in the end. We ask the following question - If the principal runs the stress test more often, can weaker stress tests (lower k) be persuasive? An affirmative answer to the above question will establish the value of repeating stress tests.

Proposition A.1 *For any J, there is a $\hat{k}(J)$ such that a J times repeated k-stress test is persuasive whenever $k > \hat{k}(J)$. $\hat{k}(J)$ is decreasing in J for $J < J^*$.*

PROOF:

Suppose that an agent in group j believes that the agents in group $l < j$ have played a cutoff strategy $\hat{s}_l$ for $l = 1, 2, \dots, (j-1)$. Then, the regime will pass the jth stress test test if

$$\frac{\theta - \sum_{l=1}^{j-1} \alpha F(\sqrt{\tau}(\hat{s}_l - \theta))}{1 - (j-1)\alpha} \geq k.$$

Suppose that $\theta = \underline{\theta}'_{j-1}$ solves the above with equality. Thus, for any cutoff strategies the agents may have played in the past, if the regime passes the stress test it means $\theta \geq \underline{\theta}'_{j-1}$ for some $\underline{\theta}'_{j-1}$. Suppose that the agents in group j follow some cutoff strategy $\hat{s}_j$. Let us define $B_j(\hat{s}_j, \alpha, \underline{\theta}'_{j-1}, k)$ such that

$$B_j(\hat{s}_j, \alpha, \underline{\theta}'_{j-1}, k) = \sum_{l=1}^{j-1} \alpha F(\sqrt{\tau}(\hat{s}_l - \underline{\theta}'_{j-1})) + \alpha F(\sqrt{\tau}(\hat{s}_j - B_j(\hat{s}_j, \alpha, \underline{\theta}'_{j-1}, k))) + (1 - j\alpha)k.$$

Substituting $\underline{\theta}'_{j-1}$, we can rewrite B_j as

$$(B_j^{\alpha,k}) \qquad B_j(\hat{s}_j, \alpha, \underline{\theta}'_{j-1}, k) = \underline{\theta}'_{j-1} - \alpha k + \alpha F(\sqrt{\tau}(\hat{s}_j - B_j(\hat{s}_j, \alpha, \underline{\theta}'_{j-1}, k)))$$

As shown by condition $(B_j^{\alpha,k})$, B_j could be lower than $\underline{\theta}'_{j-1}$, which means a regime that has passed the jth stress test ($\theta \geq \underline{\theta}'_{j-1}$) would automatically pass the next stress test if the equilibrium $\hat{s}_j$ is not very high, i.e., $k \geq F(\sqrt{\tau}(\hat{s}_j - B_j(\hat{s}_j, \alpha, \underline{\theta}'_{j-1}, k)))$. Let us further define

$$A_j(\hat{s}_j, \alpha, \underline{\theta}'_{j-1}, k) = \max\{B_j(\hat{s}_j, \alpha, \underline{\theta}'_{j-1}, k), \underline{\theta}'_{j-1}\}$$

Given the statement M_j, i.e., no one attacks the regime if it passes the $(j+1)$th stress test (regardless of the cutoff strategies played by agents before group $j+1$), A_j serves as a sufficient condition for the regime to succeed. As in the proof of Lemma 3, whenever $\theta \geq A_j(\hat{s}_j, \alpha, \underline{\theta}'_{j-1}, k)$, the regime succeeds. Consider the marginal agent in group j with cutoff signal $\hat{s}_j$. He believes that the regime will succeed with probability at least

$$\frac{\mathbb{P}(\theta \geq A_j(\hat{s}_j, \alpha, \underline{\theta}'_{j-1}, k) | \hat{s}_j)}{\mathbb{P}(\theta \geq \underline{\theta}'_{j-1} | \hat{s}_j)} = \frac{F\left(\sqrt{\tau}(\hat{s}_j - A_j(\hat{s}_j, \alpha, \underline{\theta}'_{j-1}, k))\right)}{F\left(\sqrt{\tau}(\hat{s}_j - \underline{\theta}'_{j-1})\right)}.$$

Note that if $A_j = \underline{\theta}'_{j-1}$, the above probability is 1. In this case, no agent will attack the regime regardless of his private information and group size α. In the following discussion, we will focus on the case in which $A_j = B_j > \underline{\theta}'_{j-1}$. This is the only case in which the cutoff strategy $\hat{s}_j$, the group size α and the history dependent parameter $\underline{\theta}'_{j-1}$ may matter. Substituting the Aggregate Condition

$(B_j^{\alpha,k})$, we get

$$\frac{\left(\frac{A_j-\underline{\theta}'_{j-1}}{\alpha}+k\right)}{F\left(F^{-1}\left(\frac{A_j-\underline{\theta}'_{j-1}}{\alpha}+k\right)+\alpha\sqrt{\tau}\left(\left(\frac{A_j-\underline{\theta}'_{j-1}}{\alpha}+k\right)-k\right)\right)}$$

Note that $\frac{A_j-\underline{\theta}'_{j-1}}{\alpha}+k$ is the required per-capital fundamental for success, which is greater than k given the regime passes the stress test. For $x \in [k,1]$, let us define

$$G(x,\tau,\alpha,k) := \frac{x}{F(F^{-1}(x)+\alpha\sqrt{\tau}(x-k))}$$

For any $x \in [k,1]$, $G(x,\tau,\alpha,k)$ is increasing in k (given α), decreasing in α (given k) and $G(x,\tau,\alpha,1)=1$. Let us define

$$y(\tau,\alpha,k) \equiv \min_{x\in[k,1]} G(x,\tau,\alpha,k).$$

It is easy to check that the minimal value of $G(x,\tau,\alpha,k)$, or $y(\tau,\alpha,k)$, is increasing in k and decreasing in α. Since we already know that $y(\tau,\frac{1}{J},k=0)>p$ for all $J>J^*=1/\alpha^*(p,\tau)$. For any $J \leq J^*$, we can find $\hat{k}(J)$ such that

$$y\left(\tau,\frac{1}{J},\hat{k}(J)\right)=p$$

This means, given the principal can only run stress tests for J times, if she can conduct a $k>\hat{k}(J)$ stress test, then for any possible $\underline{\theta}'_{j-1}$ and any possible cut-off strategy $\hat{s}_j$ (or equivalently for any required per-capita success requirement $\frac{A_j-\underline{\theta}_{j-1}}{\alpha}+k \in [k,1]$), given argument M_j holds true, the marginal agent believes the regime will succeed with probability strictly higher than p. Since this argument is true for any $\hat{s}_j$, under the J times repeated k stress tests, if an agent in group j attacks a regime that passes the jth stress test, then it violates sequential rationality. This validates inductive statement N_j.

Following the same induction argument, we can show that a J times repeated k - stress test is persuasive if $k>\hat{k}(J)$.

Finally, Since $y(\tau,\alpha,k)$ is decreasing in α and increasing in k, for any $J<J'<J^*=\frac{1}{\alpha^*}$,

$$y\left(\tau,\frac{1}{J'},\hat{k}(J)\right)>y\left(\tau,\frac{1}{J},\hat{k}(J)\right)=p$$

and hence $\hat{k}(J')<\hat{k}(J)$. [2] □

Note that if the stress test is not tough enough - i.e., $k \leq \hat{k}(J)$, then there

[2]Note that in Theorem 1 in our paper, we proved that when $J>J^*$, $y(\tau,\frac{1}{J},0)>p$.

exists $\hat{x} \in (k, 1]$ such that $G(\hat{x}, \tau, \alpha, k) = p$. Accordingly, one can construct a cutoff equilibrium in which the regime succeeds iff

$$\theta \geq \hat{\theta} = \alpha(\hat{x} - k) + k = \alpha\hat{x} + (1 - \alpha)k > k.$$

In this equilibrium, the first group of agents would attack given the regime passes the initial k-stress test if their private information is lower than $\hat{s}_1$ and no one in later groups would attack if the regime passes later k-stress tests. $\hat{s}_1$ and $\hat{\theta}$ satisfy the following equilibrium conditions

$$\mathbb{P}(\theta \geq \hat{\theta} | \hat{s}_1, \theta \geq k) = \frac{F\left(\sqrt{\tau}(\hat{s}_1 - \hat{\theta})\right)}{F\left(\sqrt{\tau}(\hat{s}_1 - k)\right)} = p,$$

and

$$\hat{\theta} - \alpha\mathbb{P}(s < \hat{s}_1 | \theta = \hat{\theta}) = (1 - \alpha)k$$

The first condition says that the marginal agent in the first group with signal $\hat{s}_1$ is indifferent between attacking and not attacking. The second condition says that if the regime starts with the fundamental strength $\hat{\theta}$, the attack from the first group will exactly leave $(1-\alpha)k$ residual strength to pass the second k-stress test and thus the regime will remain viable till the end since no one after the first group will attack a regime that passes the next k-stress test. The solution to above two equations is

$$\hat{\theta} = \alpha\hat{x} + (1 - \alpha)k, \ \hat{s}_1 = \frac{1}{\sqrt{\tau}} F^{-1}(\hat{x}) + \hat{\theta}.$$

Thus, J times repeated k-stress test for $k \leq \hat{k}(J)$ may not be persuasive. A J-time repeated stress tests (in regular intervals) is persuasive if and only if $k > \hat{k}(J)$. In this sense, we can make the comparison between the required strength of stress test when considering different frequencies and reach the conclusion that as the principal repeats the tests more often, repeated stress tests with weaker strength can be persuasive.

This result reconciles our main result and the one in Inostroza and Pavan (2017) and Goldstein and Huang (2016), who show that a one-time stress test with strength $k > \hat{k}(1)$ can be persuasive. Proposition A.1 shows when such stress tests can be repeated, stress tests with weaker strength can be persuasive, i.e., for $J < J^*$, $\hat{k}(J)$ is decreasing in J. The main result from our paper is that if the stress tests can be repeated with sufficient frequency, i.e., $J > J^*$, then the weakest stress test with $\hat{k}(J) = 0$ - viability test - can be persuasive. Figures 1a and 1b plot the required strength $\hat{k}(J)$ that makes a J times repeated stress test persuasive, and confirms our theoretical results.

These two figures also confirm our result of comparative statistics: (1) if agents are more reluctant the $\hat{k}(J)$ function shifts upwards - meaning that it will be more difficult to persuade the agents to not attack, (2) if the private signals are more precise, $\hat{k}(J)$ function shifts upwards - meaning that it will be harder to persuade agents to ignore their private signal and follow the principal's recommendation.

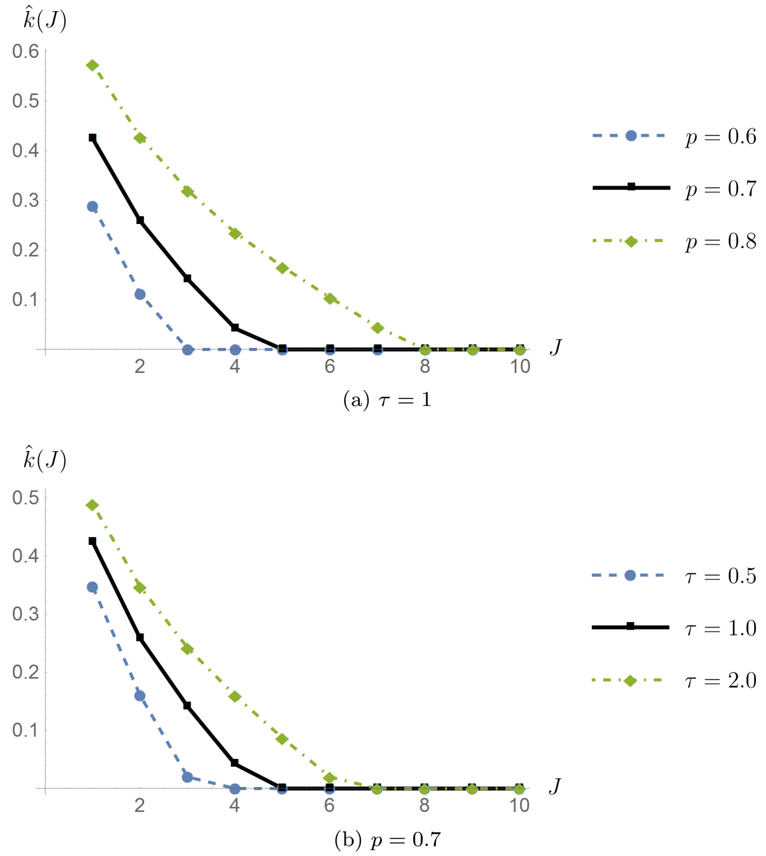

Figure 1. : $\hat{k}(J)$ - The toughness required to make the J times repeated stress tests persuasive

Discussion and Extensions

Time varying fundamental

In our model, nature chooses the fundamental $\theta \in U[\underline{\theta}, \bar{\theta}]$ at time 0 and θ can only decreases from time 0 to 1 if there are endogenous attacks. We show that sufficient diffusion can dissuade the agents from attacking. In practice, the fundamental θ can vary over time. In particular, if the agents are worried that some negative shocks in future will fail the regime, then persuading them not to attack will be more difficult.

Suppose that some negative shocks can arrive over time. Let Z_t be the number of negative shocks that can arrive by time t. Z_t follows a Poisson arrival process with parameter λ - i.e.,

$$\mathbb{P}(Z_t = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t},$$

where λ is the arrival rate of shock. We assume that λ is sufficiently small.

Assumption A.1 $e^{-\lambda} > p$.

For simplicity, we further assume that the shocks are large enough in size in the sense that once a shock hits, the regime fails. This means that when the regime passes the jth viability test, the agents know that a shock has not arrived yet. Note that, given these assumptions, if no agent attacks a viable regime from time 0 to 1, the probability the regime will succeed is greater than p. In other words, if Assumption A.1 is violated, even if all agents decided not to attack, the probability that a viable regime will succeed is lower than p, independent of the realization of θ. Hence, attacking is the dominant action, regardless of agents' information about the underlying fundamental θ and past actions. Thus, there is no room for persuasion. Below, we show that viability tests with sufficient frequency could persuade the agents not to attack under these two assumptions (whenever there is some room for persuasion).

Claim A.2 *Under Assumption A.1, there exists $\alpha^*(e^{\lambda}p, \tau) > 0$ such that a diffused policy $J > \frac{1}{\alpha^*(e^{\lambda}p,\tau)}$ is persuasive.*

PROOF:

We will show the induction statement N_j holds true as in Lemma 3. Then following the same inductive steps as in Theorem 1 we can arrive at the result.

For any group j, the probability that no shock will arrive until 1 is

$$\mathbb{P}(Z_1 = Z_{(j-1)\alpha}) = \mathbb{P}(Z_1 - Z_{(j-1)\alpha} = 0) = \mathbb{P}(Z_{1-(j-1)\alpha} = 0) = e^{-\lambda(1-(j-1)\alpha))}.$$

Consider the agent in group j who receives the cutoff signal $\hat{s}_j$ (as in Lemma 3). As before, he believes that the regime passes the jth viability test if $\theta \geq \underline{\theta}_{j-1}$ and it will survive the attacks from group j if $\theta \geq A_j(\hat{s}_j, \alpha, \underline{\theta}_{j-1})$ (A_j is defined as in Condition (A_j^{α})). The statement M_j is the same as in the paper. Under M_j, no one in group $j' > j$ attacks a regime passes the $j+1$th viability test (independent of the cutoff strategies played by agents in early groups), which means a viable regime can only fail if negative shock arrives after time $j\alpha$.

Given M_j holds true, this means that if $\theta \geq A_j(\hat{s}_j, \alpha, \underline{\theta}_{j-1})$, then in the absence of any future shock, the regime will succeed. Therefore, after receiving PNV, he believes that the regime will succeed with probability at least

$$\mathbb{P}(\text{No future shock}) \cdot \frac{\mathbb{P}(\theta \geq A_j(\hat{s}_j, \alpha, \underline{\theta}_{j-1}))|\hat{s}_j, \text{No future shock})}{\mathbb{P}(\theta \geq \underline{\theta}_{j-1}|\hat{s}_j, \text{No future shock})}$$

$$= e^{-\lambda(1-(j-1)\alpha))} \cdot \frac{F\left(\sqrt{\tau}(\hat{s}_j - A_j)\right)}{F\left(\sqrt{\tau}(\hat{s}_j - \underline{\theta}_{j-1})\right)}.$$

Substituting the Aggregate Condition, we get

$$e^{-\lambda(1-(j-1)\alpha))} \cdot \frac{\left(\frac{A_j - \underline{\theta}_{j-1}}{\alpha}\right)}{F\left(F^{-1}\left(\frac{A_j - \underline{\theta}_{j-1}}{\alpha}\right) + \alpha\sqrt{\tau}\left(\frac{A_j - \underline{\theta}_{j-1}}{\alpha}\right)\right)}$$

$$= e^{-\lambda(1-(j-1)\alpha))} \cdot G\left(\left(\frac{A_j - \underline{\theta}_{j-1}}{\alpha}\right), \tau, \alpha\right).$$

If $\alpha < \alpha^*(e^{\lambda}p, \tau)$, then $G(x, \tau, \alpha) > e^{\lambda}p$ for all $x \in [0, 1]$. Therefore, the agent who receives the cutoff signal $\hat{s}_j$ believes that the regime will succeed with probability strictly higher than

$$e^{-\lambda(1-(j-1)\alpha))} \cdot e^{\lambda}p = e^{(j-1)\alpha}p \geq p.$$

Therefore, the marginal agent who receives the cutoff signal strictly prefers not attacking. This implies that there is no cutoff equilibrium in which an agent in group j will attack the regime that passes the jth viability test. Given M_J is true, it follows that no agent will attack a viable regime as in Theorem 1. □

Due to the exogenous negative shocks, the regime could fail with a higher probability. However, this claim shows that, if the probability of having such shock is small, our diffused policy could still guarantee that, for any regime that passes the viability test, there is no endogenous attack against a viable regime.

Arrival of new information

In our baseline model, we assume that the agents move sequentially, but they do not know the past actions by other agents. Agents have some private information about the underlying fundamental θ. It is the principal who dynamically discloses some partial information based on the fundamental and the past attacks, to dissuade the agents from attacking.

However, in practice, the agents may get some new private signals on their own, while the principal runs the viability tests. This may make it harder to persuade the agents to ignore their private information and follow the principal's recommendation. Can we say that sufficient diffusion will still be persuasive?

To keep it analytically tractable, we assume that the agents have improper prior over θ and the noise ϵ follows a standard normal distribution. This satisfies log-concavity, but the support is not bounded as we assumed in the main paper. This assumption is conventional in the global game literature (See Morris and Shin (2003) or Angeletos, Hellwig and Pavan (2006)). Under these two assumptions, by replacing the F by Φ in the definition of G (see equation (4)), we can define $\alpha^*_{\Phi}(p, \tau)$ and prove the main result as we did in the paper.

Recall the diffused policy can work as long as $J > \frac{1}{\alpha^*_\Phi(p,\tau)}$, in which the critical group size α^*_Φ depends on the precision of private information τ. As we have shown in the comparative statics in the discussion section, α^*_Φ decreases with τ. If new noisy private signals arrive over time, then the agents who move later will be more informed about θ. With normal distribution of noises, the precision of private information that consists of multiple pieces of noisy signals is just the summation of the precision of each signal. If the arrival of new information is independent of the diffused policy, then the principal can design the policy targeting at the most informed agent, who has the highest precision of private information τ_{max}. A diffused policy with $J > \frac{1}{\alpha^*_\Phi(p,\tau_{max})}$ is persuasive.

The above argument is not necessarily true when the precision of the best informed agent's private information depends on the policy J. Below, we consider a simple example and show that it will be more difficult to persuade the agents if they become better informed about the underlying fundamental θ with a more diffused policy. Nevertheless, they can be persuaded.

Precision increases with diffusion.– Suppose that the principal adopts a diffused policy J. Then, when the principal runs the jth viability test, for any $j > 1$, each agent i receives an additional noisy private signal

$$s_i^j = \theta + \sigma\epsilon_i^j,$$

where $\epsilon_i^j \sim N(0,1)$. Note that the first viability test does not disclose anything new about the past attacks. We use the convention $s_i^1 = s_i$, where s_i is the initial noisy private signal as in our baseline model.

Thus, the agents in group j receive j different noisy private signals. Given the improper prior and Normal error, the updated belief of an agent i in group j is

$$\theta|\{s_i^l\}_{l=1}^j \sim N\left(\frac{\sum_{l=1}^j \tau s_i^l}{\sum_{l=1}^j \tau}, \frac{1}{\sum_{l=1}^j \tau}\right) = N\left(\frac{1}{j}\sum_{l=1}^j s_i^l, \frac{1}{j\tau}\right).$$

This shows that that the later groups are better informed about θ. More importantly, as the policy becomes more diffused (J increases and α decreases), the precision of this private information can become as high as $J\tau = \frac{\tau}{\alpha}$.

For a policy to be persuasive without new arrival of private signals, we need $\alpha < \alpha^*_\Phi(p,\tau)$. Hence, it is possible that $\alpha > \alpha^*_\Phi(p,\frac{\tau}{\alpha})$, which means, with new arrivals of private signals, the group size may not be small enough to make $G(x,\alpha) > p$ for all $x \in [0,1]$. Thus, this type of diffused policy may no longer be persuasive when agents have additional private information. But is this true for any α? If so, then our main result will be violated.

In the above simple example, we can show there exists $\alpha^{**}(p,\tau)$ such that a diffused policy with $J > \frac{1}{\alpha^{**}}$ is persuasive.

To understand this, let us consider the last group who is the best informed one. Agents in group J have private information with precision $\frac{\tau}{\alpha}$. Following the arguments in the proof of Lemma 3, we need the group size α to be sufficiently small so that the agent in group J who receives the cutoff signal believes that

$$\min_{x\in[0,1]} \frac{x}{\Phi(\Phi^{-1}(x)+\sqrt{\alpha}\sqrt{\tau}x)} > p$$

Recall that $x = \frac{A_J-\theta_{J-1}}{\alpha} \in [0,1]$. The above probability is still decreasing in α. Hence, it is easy to see that whenever $\alpha < (\alpha^*_\Phi(p,\tau))^2$,

$$\min_{x\in[0,1]} \frac{x}{\Phi(\Phi^{-1}(x)+\sqrt{\alpha}\sqrt{\tau}x)} > \min_{x\in[0,1]} \frac{x}{\Phi(\Phi^{-1}(x)+\alpha^*\sqrt{\tau}x)} = p.$$

Therefore, there exists $\alpha^{**} = (\alpha^*_\Phi)^2$ such that for any $\alpha < \alpha^{**}$, argument N_J is true. For $j < J$, since the precision is smaller, $\alpha < \alpha^{**}$ can guarantee N_j is true. Hence, we can apply the inductive argument from M_J and show that the diffused policy $J > \frac{1}{\alpha^{**}}$ is persuasive.

In the above simple example, when J increases, the precision of accumulated information $J\tau$ also increases at a rate proportional to J. One can see from the proof that as long as the precision of the accumulated private information is growing at a rate lower than J^2 (or $O(J^2)$), a sufficiently diffused policy (or a sufficiently high J) can still be persuasive.

When the precision of private signal increases at a sufficiently high rate as more diffused policy is adopted, then no matter how frequent the principal runs the viability tests, the PNV cannot be effective enough to overcome the effect of more precise private information. Below, we use a simple example to confirm this logic.

A Counter Example.– Consider the case in which the agents not only receives additional information, but the additional information becomes more precise over time. Let us assume that $\tau > 1$ and the jth private signal precision is τ^j.

$$s_i^j = \theta + \frac{1}{\sqrt{\tau^j}}\epsilon_i^j,$$

Then the updated belief

$$\theta|\{s_i^l\}_{l=1}^j \sim N\left(\frac{\sum_{l=1}^j \tau^l s_i^l}{\sum_{l=1}^j \tau^l}, \frac{1}{\sum_{l=1}^j \tau^l}\right) = N\left(\xi_i^j, \frac{1}{\hat{\tau}^j}\right),$$

$$\text{where } \hat{\tau}^j \equiv \sum_{l=1}^j \tau^l = \frac{\tau^{j+1}-\tau}{\tau-1}, \text{ and } \xi_i^j \equiv \sum_{l=1}^j \frac{\tau^j}{\hat{\tau}^j} s_i^j.$$

Following the same steps as in the previous example, we can say that agent in group J who receives the cutoff signal believes that the regime will succeed with probability

$$\frac{x}{\Phi(\Phi^{-1}(x)+\alpha\sqrt{\hat{\tau}^J}x)}.$$

It is easy to check that $\alpha^2\hat{\tau}^J = \frac{\tau}{\tau-1}\frac{\tau^J-1}{J^2}$ is increasing in J when J is sufficiently large (or $J > \frac{2}{\ln\tau}$). Hence, unlike in the previous example, when J increases and thus α decreases, the marginal agent becomes more pessimistic about the success of the regime. Thus, for any critical group size α^{**}, we can find $\alpha < \alpha^{**}$ such that there exist some cutoff equilibrium in which the marginal agent believes the regime will succeed with probability p. Accordingly, one can construct a cutoff equilibrium with positive attack against a viable regime. Hence, the arrival of new information in the dynamic setting may break our result. Our result is only robust to the case in which the new arrival of information (induced by a more diffused policy) would not increase the precision of agents' private information very rapidly.

Non-uniform Prior

We proved our main theorem based on the assumption of uniform prior. The following claim shows that uniform prior is not essential for the main theorem. However, it is worth highlighting that the uniform prior assumption is important for the monotonicity of $G(x,\tau,\alpha)$ with respect to α. Without this assumption, the intuitive comparative statics results (Proposition 3) may not hold true.

Claim A.3 *Suppose agents have the common prior on $[\underline{\theta},\bar{\theta}]$ with a continuous density function $\pi(\theta)$ and private information $s_i = \theta + \sigma\epsilon_i$, in which ϵ_i is conditionally independently and identically distributed with a differentiable and log-concave CDF $F : [-\frac{1}{2},\frac{1}{2}] \to [0,1]$. There exists $\alpha^*_\pi(p,\tau)$ such that one-time viability test can be persuasive if the group size $\alpha < \alpha^*_\pi$.*

PROOF:

As in the proof for Lemma 1, when agents play a cutoff strategy $\hat{s}$, the regime succeeds when $\theta \geq A_\pi(\hat{s},\alpha)$, in which $A_\pi(\hat{s},\alpha)$ is defined as

$$\alpha F(\sqrt{\tau}(\hat{s} - A_\pi(\hat{s},\alpha))) = A_\pi(\hat{s},\alpha). \tag{A.2}$$

The agent with private signal $\hat{s}$ believes that the regime succeeds (given it passes the viability test) with probability at least [3]

$$\mathbb{P}(\theta \geq A_\pi|\hat{s},\theta\geq 0) = \frac{\mathbb{P}(\theta\geq A_\pi|\hat{s})}{\mathbb{P}(\theta\geq 0|\hat{s})} = \frac{\int_{A_\pi}^{\hat{s}+\frac{\sigma}{2}}\pi(\theta)f(\sqrt{\tau}(\hat{s}-\theta))d\theta}{\int_0^{\hat{s}+\frac{\sigma}{2}}\pi(\theta)f(\sqrt{\tau}(\hat{s}-\theta))d\theta}$$

[3] Recall that, given any $\hat{s}$, $\theta \in [\hat{s}-\frac{\sigma}{2},\hat{s}+\frac{\sigma}{2}]$. By definition, $\hat{s} \leq 1+\frac{\sigma}{2}$ since the dominant strategy for agents with $s > 1+\frac{\sigma}{2}$ is not attacking. Because $\bar{\theta} > 1+\sigma$ by assumption, $\hat{s}+\frac{\sigma}{2} < \bar{\theta}$ and thus the upper bound of the integral is $\hat{s}+\frac{\sigma}{2}$.

Plugging $\hat{s}$ from (A.2) to the above probability, we have

$$\mathbb{P}(\theta \geq A_\pi | \hat{s}, \theta \geq 0) = \frac{\int_{A_\pi}^{A_\pi + \sigma F^{-1}(\frac{A_\pi}{\alpha}) + \frac{\sigma}{2}} \pi(\theta) f(\sqrt{\tau}(A_\pi - \theta) + F^{-1}(\frac{A_\pi}{\alpha}))d\theta}{\int_0^{A_\pi + \sigma F^{-1}(\frac{A_\pi}{\alpha}) + \frac{\sigma}{2}} \pi(\theta) f(\sqrt{\tau}(A_\pi - \theta) + F^{-1}(\frac{A_\pi}{\alpha}))d\theta}$$

Let us define (replacing $\frac{A_\pi}{\alpha}$ by x)

$$G^\pi(x, \tau, \alpha) \equiv \frac{\int_{\alpha x}^{\alpha x + \sigma F^{-1}(x) + \frac{\sigma}{2}} \pi(\theta) f(\sqrt{\tau}(\alpha x - \theta) + F^{-1}(x))d\theta}{\int_0^{\alpha x + \sigma F^{-1}(x) + \frac{\sigma}{2}} \pi(\theta) f(\sqrt{\tau}(\alpha x - \theta) + F^{-1}(x))d\theta}$$

Next, we want to show that, for any $p \in (0, 1)$, there exists $\alpha_\pi^*(p, \tau)$ such that for all $\alpha < \alpha_\pi^*$,

$$\min_{x \in [0,1]} G^\pi(x, \tau, \alpha) > p.$$

When $x \to 0$, $\alpha x + \sigma F^{-1}(x) + \frac{\sigma}{2} \to 0$. Applying L'Hospital rule, we have

$$\begin{aligned}
\lim_{x \to 0} G^\pi(x, \tau, \alpha) &= \lim_{x \to 0} \frac{(\alpha + \frac{\sigma}{f(F^{-1}(x))})\pi(\alpha x + \sigma F^{-1}(x) + \frac{\sigma}{2}) f(-\frac{1}{2}) - \alpha \pi(\alpha x) f(F^{-1}(x))}{(\alpha + \frac{\sigma}{f(F^{-1}(x))})\pi(\alpha x + \sigma F^{-1}(x) + \frac{\sigma}{2}) f(-\frac{1}{2})} \\
&= 1 - \frac{\alpha \pi(0) f(-\frac{1}{2})}{(\alpha + \sigma \frac{1}{f(-\frac{1}{2})})\pi(0) f(-\frac{1}{2})} \\
&= \frac{1}{\frac{\alpha}{\sigma} f(-\frac{1}{2}) + 1}.
\end{aligned}$$

Therefore, given $f(-\frac{1}{2}) < \infty$,

$$\lim_{\alpha \to 0} \lim_{x \to 0} G^\pi(x, \tau, \alpha\sqrt{\tau}) = 1.$$

For any $x \in (0, 1]$, it is easy to see that

$$\lim_{\alpha \to 0} G^\pi(x, \tau, \alpha) = \frac{\int_0^{\sigma(F^{-1}(x) + \frac{1}{2})} \pi(\theta) f(\sqrt{\tau}(-\theta) + F^{-1}(x))d\theta}{\int_0^{\sigma(F^{-1}(x) + \frac{1}{2})} \pi(\theta) f(\sqrt{\tau}(-\theta) + F^{-1}(x))d\theta} = 1$$

Define

$$y^\pi(\tau, \alpha) := \min_{x \in [0,1]} G^\pi(x, \tau, \alpha).$$

By theorem of maximum, $y^\pi(\tau, \alpha)$ is continuous in α. Also, $\lim_{\alpha \to 0} y^\pi(\tau, \alpha) = 1$. Therefore, for any $p < 1$, there exists α_π^* such that $y^\pi(\tau, \alpha) > p$ for all $\alpha < \alpha_\pi^*$. The rest of the proof is exactly the same as in the one for Lemma 1. □

We skipped the proof of the inductive argument N_j, since it is a natural extension of the above proof (the same as proving Lemma 3 based on Lemma 1). As discussed above, although we can still prove our main theorem under general prior, the argument depends on limit ($\alpha \to 0$) and not the monotonicity on α. Thus, the intuitive comparative statics results may not hold true.

Endogenous Timing with Costless Delay

Let us first consider the exact same model, with the only difference that the agents can decide when to attack, if attack at all (rather than exogenously deciding at one time). In this sense, attacking is an irreversible action while not attacking is reversible. Suppose there is no cost of delay. Consider the following strategy under policy J. An agent attacks at $(j-1)\alpha$ if the regime fails the jth viability test for any $j < J$, attacks at $(J-1)\alpha$ if either the regime fails the Jth test or if it passes all the viability tests but his private signal $s_i < \hat{s}$, and never attacks if the regime passes all the viability tests and $s_i \geq \hat{s}$, where $G(A(\hat{s},1),\tau,1) = p$. It is easy to check that this constitutes a cutoff equilibrium in which a viable regime may fail. The intuition is that when delayed attack is not costly, the agents may wait until the end. Therefore, regardless of however diffused the policy is, the viability test only discloses that $\theta \geq 0$, and this may not be enough to dissuade mass 1 of agents from attacking.

Endogenous Timing with Costly Delay

Let us consider a simple example in which attacking at a later date can be costly. We slightly modify the policy - there is a finite set of dates $\mathcal{T} := \{t_1, t_2, \ldots, t_J\}$ at which the principal can run viability tests. At time 0, before any agent makes his decision, the principal decides at which dates she will run the viability tests.

Consider the following payoff:[4] If an agent attacks right away he gets 1, if he waits until time t to attack, the he gets $1 - c(t)$, where $c(t) \in [0,1]$ is increasing in t. If he does not attack at all, then he gets $(1+b)$ if the regime succeeds ($b > 0$) and 0 otherwise.

Suppose that the principal announces that she will run a one time viability test at some date $\hat{t}$. The agent will either attack right away or wait for the test. As before, if the regime fails the test, then it is the dominant strategy for agents who have not attacked to attack. If the regime passes the test, the agents may or may not attack.

Claim A.4 *If an agent waits for the viability test, then he will not attack the regime that passes the viability test.*

PROOF:

Suppose for contradiction, there is some signal realization s_i such that an agent decides to wait for the test and then attack even when the regime passes the

[4]This is a simplified version of the payoff structure for illustration purpose. One can easily extend it to the one we assumed in the paper and the result remains the same.

test. Recall that it is the dominant strategy to attack if the regime fails the test. Therefore, the agent knows he is going to attack regardless of whether the regime succeeds or not. Since delayed attack is costly, he is better off by attacking right away rather than wait for the viability test. Hence, contradiction. □

This above claim shows a "screening property" - agents who will wait for the test have no intention of attacking once the regime passes the test. Therefore, any regime that passes the viability test at $\hat{t}$ will succeed. For this reason, repeating viability test is unnecessary.

The remaining question is how agents would choose between the two options: (1) attack early and (2) wait and then attack only if the regime fails the test.

Consider a cutoff strategy $\hat{s}$ such that an agent attacks immediately if $s_i < \hat{s}$. Thus, for any θ, the aggregate immediate attack is $F(s_i < \hat{s}|\theta)$. We can define $\hat{\theta}$ such that

$$\hat{\theta} - F(s_i < \hat{s}|\hat{\theta}) = 0.$$

Hence, if $\theta \geq \hat{\theta}$, $\theta - F(s_i < \hat{s}|\theta) \geq 0$ and the regime will pass the viability test. The expected payoff from attacking immediately at 0 is 1, while the expected payoff from waiting is

$$(1 + b)\mathbb{P}(\theta \geq \hat{\theta}|s_i) + (1 - c(\hat{t}))(1 - \mathbb{P}(\theta \geq \hat{\theta}|s_i)).$$

Therefore, if an agent believes that $\mathbb{P}(\theta \geq \hat{\theta}|s_i) \geq \frac{c(\hat{t})}{b+c(\hat{t})}$ then he will not attack immediately. In equilibrium, the marginal agent who receives the cutoff signal $\hat{s}$ must be indifferent between attacking immediately and waiting for the test. Therefore,

$$\mathbb{P}(\theta \geq \hat{\theta}|\hat{s}) = F(\sqrt{\tau}(\hat{s} - \hat{\theta})) = \hat{\theta} = \frac{c(\hat{t})}{b + c(\hat{t})}.$$

Note that if there was no test (and no other information transmission), there will be no delayed attack and the regime would succeed if and only if $\theta \geq \frac{1}{b+1}$ (as in standard regime change game). Thus, a regime with strength $\theta \in [\frac{c(\hat{t})}{b+c(\hat{t})}, \frac{1}{b+1})$ fails in the absence of the viability test, but agents with moderate private signals decide to wait for the test rather than attack immediately. This makes such a moderately strong regime successful. Note that $\hat{\theta}$ is increasing in $c(\hat{t})$ and $c(\hat{t})$ is increasing in $\hat{t}$. Therefore to minimize $\hat{\theta}$, the optimal policy is to do the viability test once at the lowest possible $\hat{t}$ in $\mathcal{T}$.

Hence, considering the endogenous timing of action with costly delay would not mute the effectiveness of viability test. Actually, it makes the viability test more effective in the sense that the principal does not even need to repeat the viability tests. However, as you can see that this is not just a simple extension of our insight from this paper. We think the screening property arising from the interaction between endogenous timing and information disclosure is an independent research topic and we plan to study it in more detail in the future.

REFERENCES

Acharya, Avidit, and Kristopher W Ramsay. 2013. "The calculus of the security dilemma." *Quarterly Journal of Political Science*, 8(2): 183–203.

Angeletos, George-Marios, Christian Hellwig, and Alessandro Pavan. 2006. "Signaling in a Global Game: Coordination and Policy Traps." *Journal of Political Economy*, 114(3): 452–484.

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

Inostroza, Nicolas, and Alessandro Pavan. 2017. "Persuasion in Global Games with Application to Stress Testing." `http://faculty.wcas.northwestern.edu/~apa522/persuasion-GG.pdf`.

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.