Costly information acquisition in a speculative attack: Theory and experiments*

Michal Szkup† Isabel Trevino‡

Abstract

We solve and test experimentally a global game of speculative attack where agents choose, at a cost, the precision of their private signal. We prove existence of a unique equilibrium in the coordination game and explore strategic incentives in information acquisition. In the experiment, we find that subjects follow the strategies suggested by the theory. However, contrary to our predictions, as signals become more precise the actions of subjects move towards efficiency and not risk dominance, which contradicts previous well known results of global games. We document empirically a path to convergence towards the efficient equilibrium under complete information.

Key words: global games, experiments, information acquisition, speculative attack.

JEL codes: C72, C90, D82

1 Introduction

The last decade of the twentieth century and the beginning of this century were characterized by frequent episodes of currency crises.[1] The existing models of currency crises at the time (of which Obstfeld, 1996, remains a classic example) seemed to explain the occurrence of currency crises, but could only offer a limited amount of policy guidance due to the self-fulfilling nature of the outcome, which led to multiplicity of equilibria.

The paper by Morris and Shin (1998) was a breakthrough for the modeling of speculative attacks through the use of global games. As defined by Carlsson and van Damme (1993), a global game is a coordination game where the information structure is perturbed in such a way that the players no longer have certainty over payoffs and over the other players' beliefs over payoffs. This implies that there is no longer common knowledge of the structure of the game, which leads to a unique equilibrium. Morris and Shin (1998) use this global games setup to analyze a reduced form version

*We would like to thank the National Science Foundation via grant SES-105962 and the Center for Experimental Social Science at New York University for financial support. All remaining errors are ours.

†Department of Economics, University of British Columbia, michal.szkup@ubc.ca.

‡Department of Economics, University of California, San Diego, itrevino@ucsd.edu.

[1] Examples of these episodes include the European Exchange Rate Mechanism (ERM) in 1992-93, Mexico in 1994, Asia in 1997-98, and Argentina in 2001.

of the model by Obstfeld (1996). In this model agents get a noisy private signal of the underlying state of the economy and have to decide simultaneously whether to attack the currency or not. If an agent chooses not to attack, he gets a nonnegative payoff with certainty. If he chooses to attack, his payoff depends on the underlying state of the economy and on the number of attacking agents. The global games refinement pins down a unique rationalizable equilibrium characterized by the use of threshold strategies, where an agent attacks if he gets a signal above this threshold, and does not attack otherwise.

The richness and applicability of the global games models have given rise to a vast literature that tries to understand under what conditions the uniqueness result holds, and how far one can extend the setup to better portray economic realities. The information structure has always played a central role and even small changes in the assumptions of the model have important consequences on equilibrium selection.[2] In most global games models, however, the information possessed by agents is typically given to them exogenously, rather than chosen. We develop a global games model of speculative attack where each agent has the possibility to choose the precision of his private information. In our model, two agents covertly choose the precision of the private signal they will observe about the state of the economy, at a cost, and then play the coordination game of speculative attack, as in Morris and Shin (1998), using the information acquired in the first stage. We solve the model and then test the predictions of this model in the laboratory in order to observe how information sets are generated and implemented in a speculative attack.

In the experiment we find that subjects follow the strategies predicted by theory, with over 92% of subjects using threshold strategies and 40% choosing the equilibrium precision level. We observe that subjects' choices are stable over time and do not exhibit significant fluctuations. However, contrary to the theoretical predictions, we find that as subjects choose more precise information they coordinate more often on attacking the currency, and as a result attacks are more successful when agents hold more precise information. Thus, we find that as subjects acquire more precise information their behavior moves towards the efficient equilibrium and away from the risk dominant equilibrium, which contradicts well known results about behavior in global games in the limit, as the noise vanishes. Moreover, as the noise of the private signals vanishes, we observe empirically a path to convergence towards the efficient equilibrium of the complete information version of our model, since subjects behave in accordance to the efficient threshold in a treatment with complete information about fundamentals. These departures from the theory lead to a welfare improvement with respect to the expected payoff in equilibrium and with respect to a constrained efficiency benchmark.

The above results raise an important question about using risk dominance (and the global games predictions in the limit) as an empirically relevant equilibrium selection device. In particular, the

[2] Examples include Hellwig (2002) who characterizes uniqueness of equilibrium based on the relative informativeness of private and public signals. In Hellwig et al (2006) and Angeletos and Werning (2006), equilibrium parity conditions create an endogenous public signal through prices that restores multiplicity of equilibria. In Angeletos et al (2006) multiple equilibria are restored by the introduction of cheap talk from the monetary authority.

results suggest that subjects behave in a similar way under complete information and under incomplete information when their information is very precise. This contrasts the theoretical literature that predicts that players' behavior in coordination games differs dramatically under complete and incomplete information.[3]

We believe that the study of costly private information acquisition in a speculative attack brings the global games model closer to the macroeconomic phenomenon of interest. Investors involved in speculative attacks do not hold symmetric and exogenously given information about the economy. Instead, they continuously make efforts to improve the information they possess, and they are willing to pay for it. Investment groups and individuals pay experts to extract more accurate information about the financial system in order to minimize losses, creating a market for information expertise and financial advising. Therefore, it seems both a natural and a necessary extension to study a speculative attack game where agents get to improve the precision of the information they receive.

Different authors have studied costly information acquisition in coordination games with imperfect information. In Szkup and Trevino (2015) we develop a similar model with a continuum of players, which allows for explicit analytical solutions and a full equilibrium characterization. Hellwig and Veldkamp (2009) investigate what agents choose to observe when there is a cost to acquiring information and how information choices affect equilibrium outcomes in a beauty contest model. In a more general setup, Colombo, Femminis and Pavan (2014) study the welfare effects of endogenous information acquisition in a flexible framework of quadratic loss functions that include beauty contest models as a special case.[4] In a global games context, some studies related to ours are Bannier and Heinemann (2005), in which the monetary authority chooses the precision of the private signals in order to minimize the probability of currency crises, and Yang (2015), where information acquisition in coordination games is modeled using a rational inattention framework. With a notable exception of Yang (2014), none of these papers, however, studied a model with a finite number of agents, where agents directly take into account strategic effects of their choices on the equilibrium outcomes. More importantly, this is the first paper to test the predictions of a coordination game with costly information acquisition experimentally.[5]

[3] Our results support the notion of using payoff-dominance as the equilibrium selection device, as suggested by Harsanyi and Selten (1988), when signals are very precise. The experimental literature on equilibrium selection in coordination games with complete information and Pareto ranked equilibria suggests that subjects do not necessarily coordinate on the efficient equilibrium without additional mechanisms. While the initial studies of Cooper et al. (1990, 1992) and van Huyck et al (1990, 1991) find mostly coordination failure, later studies reported mechanisms to enhance coordination in the efficient equilibrium (e.g. communication). Haruvy and Stahl (1999) compare the predictive power of different selection principles in coordination games with complete information and find that payoff dominance performs the worst, followed by risk dominance. See Devetag and Ortmann (2007) for a comprehensive survey on this literature.

[4] Szkup and Trevino (2015) compares the results of costly information acquisition in a global game with existing results for beauty contest type of games.

[5] In addition, Nikitin and Smith (2008) and Zwart (2008) study costly information acquisition in a Diamond and Dybvig (1983) type of model and in a liquidity run setup, respectively. However, in these last two studies information acquisition is modeled as a binary decision to acquire a private signal with a given precision or not to acquire a signal at all, which contrasts our setup where all agents observe private signals and have to choose their individual precision.

Global games have proven themselves to be very relevant for modelling coordination games with incomplete information not only because they predict a unique equilibrium, but also because their predictions are consistent with observed behavior in the laboratory. Heinemann et al. (2004, 2009) test the model by Morris and Shin (1998) experimentally and find a robust use of threshold strategies, even for treatments with perfect information. Just like Heinemann et al. (2004), we find that subjects follow the strategies predicted by the global games refinement, for any precision chosen. However, we generalize their results by showing how the incidence and success of speculative attacks varies with different precision levels when we endogenize information choices and by studying an empirical path to convergence as the noise of the private signals vanishes. Our results have strong implications for the existing theory of global games by showing that this empirical path to convergence leads to the efficient equilibrium, and not the risk dominant one that is suggested by the theory.

The paper is structured as follows. In section 2 we present and solve the theoretical model for two players and review the theoretical predictions for the set of parameters used in the experiment. The experimental design and results are presented in section 3. We discuss our results in section 4 and we conclude in section 5. All the proofs from the theory and additional tables from the experimental analysis are relegated to the appendix.

2 Theory: Speculative attack with information acquisition

Our framework for the speculative attack game builds on the 2×2 model of Carlsson and van Damme (1993) (CvD93 henceforth), and a discrete version of Morris and Shin (1998). In contrast to these papers, we add an initial stage where agents privately choose the precision of their information before playing the speculative attack game.

Initially, the exchange rate is held fixed at some arbitrary level. Each agent is endowed with one unit of domestic currency and has to decide whether to change it for foreign currency (attack) or to keep it (not attack). If a high enough fraction of the agents attacks the currency, the currency is allowed to float at the underlying (shadow) exchange rate. The economy is characterized by θ, the net gain from attacking the currency in case of a devaluation.[6] The payoff of holding domestic currency is normalized to zero. If an agent chooses to attack, he has to pay a transaction cost T, and gets payoff θ if the attack is successful and devaluation occurs. If the attack is unsuccessful and the currency is not devalued, then the profit of an agent who chooses to attack is $-T$. Thus, the payoffs of the speculative attack for each one of the agents are defined as follows:

This difference has important implications since a unique equilibrium is not guaranteed in the setup of Nikitin and Smith (2008) and Zwart (2008).

[6] One can think of θ as an inverse measure of the strength of fundamentals.

	Devaluation	No Devaluation
Attack	$\theta - T$	$-T$
Not Attack	0	0

Let $A(\theta)$ be the number of agents that attack when the level of fundamentals is θ and $A^*(\theta)$ be the minimum number of attacking agents needed to enforce a devaluation. Since a high θ implies weaker fundamentals, we assume that $A^{*\prime}(\theta) < 0$. A devaluation occurs whenever the number of players that attacked, $A(\theta)$, is greater than $A^*(\theta)$, the minimum number of attacking players needed for the attack to be successful.[7]

There are two agents in the economy. With the use of the function $A^*(\theta)$ we define two cutoff values for the state of the economy, $\{\underline{\theta}, \overline{\theta}\}$, that will determine the number of agents that are needed to attack in order to provoke a devaluation. Specifically, when $\theta \geq \overline{\theta} = A^{*-1}(1)$ one agent alone can provoke devaluation, and when $\theta < \underline{\theta} = A^{*-1}(2)$, the economy is strong enough so that devaluation does not take place even if both agents attack. Therefore, for $\theta \in [\underline{\theta}, \overline{\theta})$ agents have to coordinate on the attack in order for devaluation to occur.

In the game with complete information (common knowledge of θ), multiple equilibria arise:

- If $\theta \geq \overline{\theta}$, it is a dominant strategy to attack.
- If $\theta < \underline{\theta}$, it is a dominant strategy not to attack.
- If $\theta \in [\underline{\theta}, \overline{\theta})$, there are two equilibria in pure strategies and one in mixed strategies.

Following CvD93, to obtain a unique equilibrium we assume that the state θ is a random variable that follows a normal distribution with mean μ_θ and variance σ_θ^2.[8] The actual realization of θ is unobserved by the agents. Instead, each agent $i = 1, 2$ observes a noisy private signal of the state θ,

$$x_i = \theta + \sigma_i \varepsilon_i$$

for $i = 1, 2$ where $\sigma_i > 0$ and ε_i is a random variable, normally distributed with mean zero and variance 1. We assume that ε_i is *i.i.d.* across agents and we denote by $f(\cdot)$ the normal density of ε_i, and $F(\cdot)$ its cumulative distribution function. Note that the precision of the signal that each agent receives is defined by its standard deviation, σ_i.[9] Information acquisition in our model will take place through the choice of σ_i.

[7] Specifically, assume that the government derives a benefit $b > 0$ from defending the exchange rate at the fixed level. However, there is a cost to it, which depends on θ and on the size of the attack A, $c(\theta, A)$. We assume that $c_\theta(\theta, A) > 0$ and $c_A(\theta, A) > 0$. Following Morris and Shin (1998), we do not model explicitly the decision of the central bank here, but it underlies our analysis since we assume that the central bank observes the size of the attack A and mechanically decides to devalue if $A \geq A^*(\theta)$, where $A^*(\theta)$ solves $b = c(\theta, A^*(\theta))$.

[8] We make the additional technical assumption that $\mu_\theta \in [\underline{\theta}, \overline{\theta}]$.

[9] Recall that the precision is defined as the inverse of the variance of a random variable that follows a normal distribution. Therefore, in what follows we will use either of the following terms to refer to the same situation: higher precision, lower variance, or lower standard deviation.

We add to the above one-shot game an initial stage where agents get to choose the precision of the noisy signal they receive. This means that, starting from an initial common standard deviation σ_0, each agent can either keep this level of noise or improve his precision by "buying" a lower σ_i at a cost $C(\sigma_i)$. $C(\cdot)$ is continuous, with $C(\sigma_0) = 0$, $C'(\sigma_0) = 0$, $C'(\sigma) < 0$, $C''(\sigma) > 0$, for all $\sigma \in (0, \sigma_0)$, and $\lim_{\sigma\to\infty} C'(\sigma) = \infty$. We assume that the level of precision that each agent chooses is private information, and hence not observed by the other player.[10]

2.1 Solving stage two: Uniqueness of equilibrium in the speculative attack game

We solve the model by backward induction. At the beginning of the second stage each agent observes his own signal $x_i = \theta + \sigma_i\varepsilon_i$, where σ_i corresponds to agent i's choice of precision in the first stage. Note that the second stage alone can be thought of as a standard 2×2 global game with the notable difference that agents are potentially heterogenous with respect to the precision of their signals. This is a significant departure from the original model of CvD93 and from standard models of global games.[11]

After observing their own signals, agents update their beliefs about the fundamental θ and about the other agent's type.[12] Given signal x_i, agent i's posterior belief regarding θ is given by

$$\theta|x_i \sim N\left(\widehat{\theta}_i, \widehat{\sigma}_i^2\right), \text{ where } \widehat{\theta}_i = \frac{\sigma_i^2\mu_\theta + \sigma_\theta^2 x_i}{\sigma_i^2 + \sigma_\theta^2} \text{ and } \widehat{\sigma}_i = \frac{\sigma_i^2\sigma_\theta^2}{\sigma_i^2 + \sigma_\theta^2}$$

As in the classic global games setup, in the second stage agents follow threshold strategies,[13] i.e. for $i = 1, 2$:

$$a(x_i; \boldsymbol{\sigma}) = \begin{cases} 1 \text{ (attack)} & \text{iff } \; x_i \geq x_i^*(\boldsymbol{\sigma}) \\ 0 \text{ (not attack)} & \text{iff } \; x_i < x_i^*(\boldsymbol{\sigma}) \end{cases}$$

Notice that agents do not necessarily have the same thresholds because they are a function of their chosen precisions.

Next, we derive the expected payoff to agent i from attacking the currency. Suppose that agent i believes that agent j uses a threshold x_j^*. Furthermore, recall that if $\theta < \underline{\theta}$ a devaluation cannot occur even if both players coordinate on the attack, if $\theta \geq \overline{\theta}$ one agent alone is capable of inducing a

[10] This is a common way to model information acquisition in the literature (see Persico, 2000, Colombo et al., 2014, or Szkup and Trevino, 2015, among others).

[11] Corsetti et al. (2004) allow agents to have asymmetric distirbutions of noise. However, in their setup the agents differ also with respect to other characteristics, such as size, leading to very different dynamics than the usually present in standard global games models. See Yang (2014) for another recent exercise where agents might be heterogenous.

[12] In this context the type of an agent is described by the signal he observes and his precision choice. Recall that we assume that agents hold correct beliefs about each other's information choices, so after observing his own signal x_i, agent $i = 1, 2$ only forms beliefs about agent j's signal, x_j.

[13] We first focus on threshold strategies and then verify that the resulting equilibrium is indeed the unique equilibrium of the second stage. In particular, we show that any equilibrium of the game will be in monotone strategies (thresholds), and that there is a unique dominance solvable equilibrium threshold. This is proven in Theorem 1 in the appendix.

devaluation, and if $\theta \in [\underline{\theta}, \overline{\theta})$ a devaluation occurs if and only if both agents attack. Thus, assuming that agent j follows a threshold strategy with a switching point x_j^*, agent i's expected payoff from attacking, conditional on observing signal x_i, is given by:

$$E\left[\theta \Pr\left(x_j \geq x_j^* \mid \theta\right) \mid x_i, \theta \in [\underline{\theta}, \bar{\theta})\right] \Pr(\theta \in [\underline{\theta}, \bar{\theta})|x_i) + E\left[\theta \mid x_i, \theta \in [\bar{\theta}, \infty]\right] \Pr(\theta \in [\bar{\theta}, \infty]|x_i).$$

We denote this payoff by $v_i(x_i, x_j^*; \boldsymbol{\sigma})$. Using our distributional assumptions, we can express $v_i(x_i, x_j^*; \boldsymbol{\sigma})$ as:

$$v_i(x_i, x_j^*; \boldsymbol{\sigma}) = \int_{\underline{\theta}}^{\overline{\theta}} \theta \left(1 - F\left(\frac{x_j^* - \theta}{\sigma_j}\right)\right) \frac{1}{\widehat{\sigma}_i} f\left(\frac{\theta - \widehat{\theta}_i}{\widehat{\sigma}_i}\right) d\theta + \int_{\overline{\theta}}^{\infty} \theta \frac{1}{\widehat{\sigma}_i} f\left(\frac{\theta - \widehat{\theta}_i}{\widehat{\sigma}_i}\right) d\theta - T \qquad (1)$$

where $\widehat{\theta}_i = \frac{\mu_\theta \sigma_i^2 + x_i \sigma_\theta^2}{\sigma_i^2 + \sigma_\theta^2}$, $\widehat{\sigma}_i^2 = \frac{\sigma_i^2 \sigma_\theta^2}{\sigma_i^2 + \sigma_\theta^2}$ and $1 - F\left(\frac{x_j^* - \theta}{\sigma_j}\right)$ is the probability that agent j attacks the currency, conditional on the realization of θ. The next lemma establishes two basic properties of this function.

Lemma 1 *The payoff for agent i of attacking, $v_i(x_i, x_j^*; \boldsymbol{\sigma})$, is increasing in his own signal x_i, and decreasing in the other agent's threshold x_j^*, for $i, j = 1, 2$, $i \neq j$.*

To compute the equilibrium of the model we need to analyze the optimal thresholds. A threshold x_i^* is defined as the value of agent i's signal x_i for which he is indifferent between attacking and not attacking, taking as given the strategy of the other player. For $i = 1, 2$, $i \neq j$, taking x_j^* as given, the optimal threshold of agent i, x_i^*, solves the following equation:

$$v_i(x_i^*, x_j^*; \boldsymbol{\sigma}) = 0. \qquad (2)$$

We now define an equilibrium for the second stage of the game.

Definition 1 *Given (σ_i, σ_j), an equilibrium in monotone strategies for the second stage of the game is a pure strategy profile $a_i(x_i; \boldsymbol{\sigma})$ such that, for $i = 1, 2$*

$$a_i(x_i; \boldsymbol{\sigma}) = \begin{cases} 1 & \text{if } x_i \geq x_i^*(\boldsymbol{\sigma}) \\ 0 & \text{if } x_i < x_i^*(\boldsymbol{\sigma}) \end{cases}$$

where $x_i^*(\boldsymbol{\sigma})$ solves

$$v(x_i^*(\boldsymbol{\sigma}), x_j^*(\boldsymbol{\sigma}); \boldsymbol{\sigma}) = 0$$

Lemma 1 implies that the best response functions are well defined. To prove that there is a unique equilibrium in the second stage of the game, we need to show that there is a unique pair of thresholds that satisfies the above equilibrium conditions, i.e. that there exists a unique combination of (x_1^*, x_2^*) that solves simultaneously equilibrium condition (2) for $i = 1, 2$.[14] This is demonstrated in Theorem 1.

Theorem 1 *There exists a unique, dominance solvable equilibrium of the second stage of the game in which both players use threshold strategies characterized by* (x_1^*, x_2^*) *if either:*[15]

1. $\frac{\sigma_i}{\sigma_\theta} < K_i(\underline{\theta}, \overline{\theta}, \mu_\theta)$, *$i = 1, 2$ holds, for any pair of* (σ_1, σ_2),[16] *or*

2. $\sigma_\theta > \overline{\sigma}_\theta$, *where* $\overline{\sigma}_\theta$ *is determined by the parameters of the model.*

This result is in line with previous work on global games, like Hellwig (2002), who introduces a public signal to the model of Morris and Shin (1998) and shows that uniqueness of equilibrium is preserved when the public signal is noisy enough with respect to the private signals. Our conditions are in the same spirit as those of the above papers, however, due to the strategic effects of individual actions present in our model, they take more complex functional forms. Moreover, in our model as $\sigma_\theta \to \infty$, the coordination game with asymmetric players has always a unique equilibrium. In our context we can think of the prior as being a public signal, since it carries information about the fundamental that is available to all agents at no cost. Finally, it is important to note that the conditions on precisions stated in Theorem 1 are only sufficient but not necessary for uniqueness of equilibrium in the second stage.

[14] To prove this result we make use of the literature on monotone supermodular games. In Appendix 2, we show that the game specified above belongs to the class of Bayesian monotone supermodular games, as defined by Vives and van Zandt (2007). In particular, we extend the result of Vives and van Zandt (2007) to prove existence of the least and the greatest Bayesian Nash Equilibria in monotone strategies in games with unbounded but integrable utility functions. This extension allows us to prove uniqueness of equilibrium in the second stage of our game by showing that the least and greatest BNE of our game coincide, and thus there is a unique equilibrium in threshold strategies.

[15] This result implies that the game specified in the second stage is dominance solvable (Milgrom and Roberts, 1990). While the dominance solvability of symmetric binary global games is well understood (see Morris and Shin, 2003), our contribution is to show that it also applies to an asymmetric game where both thresholds are functions of each other. Corsetti et al (2004) also show dominance solvability in a binary action global game with asymmetric players. Nevertheless, in that setup the optimal threshold of a "large" player is independent of the threshold of the "small" players. Frankel, Morris, and Pauzner (2003) also show dominance solvability for players with bounded utility functions and asymmetric noise distributions with finite support. In our setup, utility functions are unbounded and noise distributions have infinite support.

[16] Where $K_1(\underline{\theta}, \overline{\theta}, \mu_\theta) := \dfrac{1 - F\left(\frac{\overline{\theta} - \underline{\theta}}{\widehat{\sigma}_2}\right) + \frac{1}{\widehat{\sigma}_2} f\left(\frac{\overline{\theta} - \underline{\theta}}{\widehat{\sigma}_2}\right)\kappa_1}{\frac{1}{\sqrt{2\pi}}\left(\frac{1}{\sqrt{2\pi}} + \frac{\overline{\theta}\sigma_1^2\left(1 + \frac{\sigma_1^2}{\widehat{\sigma}_2^2}\right) - \widehat{\sigma}_2^2 \frac{\sigma_2^2}{\sigma_\theta^2}}{\widehat{\sigma}_2^2 + \sigma_1^2}\right)}$,

$$K_2(\underline{\theta}, \overline{\theta}, \mu_\theta) := \frac{1 - F\left(\frac{\underline{\theta} - \widehat{\theta}_{\min}}{\widehat{\sigma}_1}\right) - \left[F\left(\frac{\overline{\theta} - \widehat{\theta}_{\min}}{\widehat{\sigma}_1}\right) - F\left(\frac{\underline{\theta} - \widehat{\theta}_{\max}}{\widehat{\sigma}_1}\right)\right] F\left(\frac{\sigma_2(\widehat{\theta}_{\max} - \underset{\sim}{\theta}_2)}{\widehat{\sigma}_2^2}\right) + \frac{1}{\widehat{\sigma}_1} f\left(\frac{\overline{\theta} - \theta^*_{\min}}{\widehat{\sigma}_1}\right)\kappa_2}{\frac{1}{\sqrt{2\pi}}\left(\frac{1}{\sqrt{2\pi}} + \frac{\overline{\theta}\sigma_2^2\left(1 + \frac{\sigma_2^2}{\widehat{\sigma}_1^2}\right) - \widehat{\sigma}_1^2 \frac{\sigma_1^2}{\sigma_\theta^2}}{\widehat{\sigma}_1^2 + \sigma_2^2}\right)}$$

and $\kappa_i := \overline{\theta} F\left(\dfrac{\sigma_2\left(\theta^*_{\min} - \frac{\sigma_\theta^2\overline{\theta} + \sigma_2^2\mu_\theta}{\sigma_\theta^2 + \sigma_2^2}\right)}{\widehat{\sigma}_2^2}\right) + \underline{\theta}\left(1 - F\left(\dfrac{\sigma_2\left(\theta^*_{\max} - \frac{\sigma_\theta^2\underline{\theta} + \sigma_2^2\mu_\theta}{\sigma_\theta^2 + \sigma_2^2}\right)}{\widehat{\sigma}_2^2}\right)\right)$

where $\theta^*_{\min}$, $\theta^*_{\max}$, $\widehat{\theta}_{\min}$, $\widehat{\theta}_{\max}$, and $\widetilde{\underline{\theta}}_2$ are defined endogenously in detail in the appendix.

2.1.1 Limiting case

We now investigate what happens with the unique equilibrium in the limit, as the noise of the private signals vanishes. This result will be useful for the experiment, since it has been shown (Heinemann et al., 2004) that when agents play the speculative game without information acquisition in the laboratory, subjects use threshold strategies even under complete information.

For $i = 1, 2$, $i \neq j$, we consider what happens when both $\sigma_i \to 0$ and $\sigma_j \to 0$ simultaneously and $\frac{\sigma_i}{\sigma_j} \to c$, for some $c \in \overline{\mathbb{R}}_+$, i.e. when the signal noise goes to zero for both agents while the ratio of their noises tends to some constant c.

Lemma 2 *Suppose that $\sigma_i \to 0$, $\sigma_j \to 0$ and $\frac{\sigma_i}{\sigma_j} \to c$ where $c \in \overline{\mathbb{R}}_+$. If the above game has a unique equilibrium, then this equilibrium converges to the risk-dominant equilibrium of the complete information game, i.e. $x_i^* \to 2T$ and $x_j^* \to 2T$.*

Note that, given the payoff structure, if $\theta \geq 2T$ then attacking is the risk dominant action, and if $\theta < 2T$ not attacking is the risk dominant action. The result of the lemma is of particular importance because it implies that the global games equilibrium converges to the risk-dominant equilibrium of the complete information game (as defined by Harsanyi and Selten, 1988) in the case of players with asymmetric signal distributions. This feature was first pointed out by Carlsson and van Damme (1993) for 2×2 global games with symmetric noise structures. Lemma 2, thus, extends their limiting result for the case of unbounded utility functions and asymmetric noise distributions with infinite support.

2.2 Solving stage one: Optimal information acquisition

Given the equilibrium strategies in the second stage of the game, we now study the first stage where agents get to choose the precision of the private signal they will observe, at a cost. Agents initially have a common prior about the underlying state of the economy, $\theta \sim N\left(\mu_\theta, \sigma_\theta^2\right)$. During this first stage they can improve the precision of the signal they will receive in the second stage. If an agent does not improve his precision he will receive a signal $x_i \sim N(\theta, \sigma_0^2)$. We assume that the ratio $\frac{\sigma_0}{\sigma_\theta}$ satisfies the uniqueness condition for the equilibrium in the second stage of the game derived in section 2.1.[17]

The expected utility of agent i, who chose standard deviation σ_i, and who believes that his opponent chose standard deviation σ_j and that his opponent holds correct beliefs about his own choice of σ_i, is given by:

[17] In general, if the condition for uniqueness is not satisfied, the game played in the second stage will have three equilibria in monotone strategies (which can be Pareto ranked). If we assume that agents play a Pareto dominant (or Pareto dominated) equilibrium, as the precision of their private signals decreases and we enter the dominance region, the equilibrium threshold would experience a jump, creating a discontintuity in the first period expected utility at that point.

$$U_i(\sigma_i, \sigma_j) = \int_{x_i \geq x_i^*} v_i(x_i, x_j^*; \boldsymbol{\sigma}) \frac{1}{\sqrt{\sigma_i^2 + \sigma_\theta^2}} \; f\left(\frac{x_i - \mu_\theta}{\sqrt{\sigma_i^2 + \sigma_\theta^2}}\right) \; dx_i - C(\sigma_i) \tag{3}$$

where $v_i(\cdot)$ is defined in equation (1). $C(\sigma_i)$ is the cost associated with agent i's choice of precision, and $\frac{1}{\sqrt{\sigma_i^2+\sigma_\theta^2}} \; f\left(\frac{x_i-\mu_\theta}{\sqrt{\sigma_i^2+\sigma_\theta^2}}\right)$ is the unconditional density of agent i's signal, given his choice of σ_i.

We now define a pure strategy Bayesian Nash Equilibrium for the full game and proceed to analyze the problem faced by each agent.

Definition 2 (Equilibrium) *A pure strategy Bayesian Nash Equilibrium is a pair of information choices* (σ_i^*, σ_j^*)*, optimal decision rules for the second stage* $a_i(x_i; \boldsymbol{\sigma})$*, and belief functions* $\mu_i : (0, \sigma_0) \to [0, 1]$ *such that for each* $i = 1, 2$ *we have:*

- $U_i(\sigma_i^*, \sigma_j^*) \geq U_i(\sigma_i, \sigma_j^*) \quad \forall \sigma_i \in (0, \sigma_0)$
- *The belief function* μ_i *satisfies* $\mu_i\ (\sigma_j^*) = 1$ *and* $\mu_i\ (\sigma_j) = 0$ *for* $\sigma_j \neq \sigma_j^*$*;*
- *For* $i = 1, 2$*, given the belief function* μ_i*, agent* i*'s decision rule is given by*

$$a_i(x_i; \boldsymbol{\sigma}) = \begin{cases} 1 & \text{if } x_i \geq x_i^*(\sigma_i, \sigma_j^*) \\ 0 & \text{if } x_i < x_i^*(\sigma_i, \sigma_j^*) \end{cases}$$

where $x_i^*(\boldsymbol{\sigma})$ *solves*

$$v(x_i^*(\sigma_i, \sigma_j^*), x_j^*(\sigma_i^*, \sigma_j^*); \boldsymbol{\sigma}) = 0$$

The first condition is an optimality condition and it requires agent i to have no incentives to deviate from his equilibrium precision choice. The second condition requires that an agent assigns positive probability only to the actual equilibrium choice of his opponent. Finally, the last condition requires each agent to act optimally in the second stage, even after unilateral deviations.

It is convenient at this stage to define the benefit of choosing σ_i, i.e. the gross payoff of choosing σ_i:

$$B_i(\sigma_i; \sigma_j) := \int_{x_i > x_i^*} v_i(x_i, x_j^*; \boldsymbol{\sigma}) f(x_i; \sigma_i) dx_i$$

One can show that the marginal benefit of information is always finite and converges to zero (see the appendix). Since $\lim_{\sigma_i \to 0} C'(\sigma_i) = \infty$ it follows that that agents will never find it optimal to choose infinitely precise signals.

We now establish existence of equilibrium for the full game. The proof involves establishing two intermediate results that state that the best response functions are well defined and that agents improve their precision in any equilibrium of the game.

Theorem 2 (Existence) *There exists a symmetric pure-strategy Bayesian Nash Equilibrium of the game with costly information acquisition.*

In the model with two agents, because of the lack of a closed form solution, it is difficult to derive parametric conditions to characterize comparative statics explicitly and to perform a thorough equilibrium characterization.[18] We briefly discuss some results that arise from numerical simulations that will be helpful to understand the predictions of the model for the set of parameters used in the experiment.

For most parameter values there are increasing differences in precision choices, i.e. more information is always beneficial ($B_i'(\sigma_i, \sigma_j) > 0$) and the marginal benefit of increasing individual precisions is increasing in the precision of the other agent $\left(\frac{\partial B_i'(\sigma_i,\sigma_j)}{\partial \sigma_j} > 0\right)$. Therefore, there are strategic complementarities in information acquisition.

The next observation pertains to the way in which optimal thresholds vary with each player's precision choices. As pointed out by Metz (2002), Bannier and Heinemann (2005), and Szkup and Trevino (2015), typically in global games the comparative statics with respect to precisions depend on the agents' prior beliefs. A similar result applies in this case. More precisely, we find that when agents are optimistic about the state of the world and the cost of attacking is low, i.e. when μ_θ is high with respect to T, thresholds are increasing functions of both precisions, irrespective of σ_θ. In the remaining cases (when agents are pessimistic about the state of the world and costs are high, or when costs and prior beliefs are of about the same magnitude, i.e. when μ_θ is very low with respect to T, or when μ_θ is very similar to T) the direction in which precisions affect thresholds depends on the precision of the prior. In particular, when the prior is very noisy (high σ_θ), the threshold of each agent is increasing with the precision of both agents. When the precision of the prior is high (but noisy enough so that a unique equilibrium in the second stage is ensured), the threshold of each player is a decreasing function of the precision of both agents.

In the next subsection we state the equilibrium predictions for the parameters used in the experiment.

2.3 Theoretical predictions for the experiment

The theoretical model is governed by a set of parameters $\Theta = \left\{\mu_\theta, \sigma_\theta, (\underline{\theta}, \overline{\theta}), T, \{\sigma_i\}, \{C(\sigma_i)\}\right\}$. The parameters chosen for the experiment are:

$$\Theta = \{50, 50, (0, 100), 18, \{1, 3, 6, 10, 16, 20\}, \{6, 5, 4, 2, 1.5, 1\}\}$$

In particular:

- The fundamental θ is randomly drawn from a normal distribution with mean 50 and standard deviation of 50.
- The coordination region is for values of $\theta \in [0, 100)$.

[18] For a complete equilibrium characterization of the model with a continuum of agents see Szkup and Trevino (2015).

- The cost of attacking is $T = 18$.
- A discrete choice set of precisions and the cost associated with each precision was presented in the form of a menu of 6 precision levels, standard deviations, and costs:[19]

Precision level	Standard deviation	Cost
1	1	6
2	3	5
3	6	4
4	10	2
5	16	1.5
6	20	1

Table 1: Precision choices

We decided not to have a default precision chosen for subjects in order to avoid status quo biases. The reason to introduce a discrete choice set for precisions was to simplify the choice for subjects and the data analysis. We believe six is a reasonable number of options to observe dynamics in the level of informativeness that subjects choose, without losing statistical power.

Given these parametric assumptions we characterize the predictions of the model in the form of three main hypotheses to be tested with our experiment:

Hypothesis 1 *For any precision choices, subjects use threshold strategies.*

Hypothesis 2 *Thresholds are increasing functions of precisions.*

Hypothesis 3 (Equilibrium) *Subjects coordinate on a precision level of 4 and set a symmetric threshold of 28.31.*

Hypothesis 1 establishes that subjects will use the type of strategy predicted by global games, for any precision choices, as shown in subsection 2.1. We have chosen a situation where the mean of the prior is high with respect to the cost of attacking to ensure that the effect of precisions on thresholds does not depend on the precision of the prior (see Szkup and Trevino, 2015). This allows us to formulate Hypothesis 2, which states that the optimal threshold in stage two for each agent is increasing in the precision of both agents. Finally, Hypothesis 3 aims to test the unique symmetric equilibrium prediction for the parameters used in the experiment, which corresponds to subjects coordinating on their choice of precision level 4 and setting a symmetric threshold at 28.31. Implicit in this prediction is that precision choices are strategic complements, which leads agents to coordinate on both precisions and actions. As an additional feature of the model coming

[19] In the remaining of the paper we will refer to information choices as precision choices to be consistent with the language used in the implementation of the experiment. We will use the term precision as a qualitative measure of informativeness of the signals, i.e. we will compare levels of precision, and not magnitudes of standard deviations.

from section 2.1.1, note that the equilibrium prediction in the limit, as signal noise vanishes, is equal to $2T = 36$, which corresponds to the risk dominant equilibrium for our game. Previous experimental evidence shows that subjects coordinate on this limiting threshold when playing a global game without information acquisition (see Heinemann et al., 2004).[20]

3 The experiment

We present results of a series of laboratory experiments designed to test the implications of the model from section 2. The experiment was conducted at the Center for Experimental Social Science at New York University using the usual computerized recruiting procedures. Each session lasted from 60 to 90 minutes and subjects earned on average $25, including a $5 show up fee. All subjects were undergraduate students from New York University.[21]

3.1 Experimental design

We implement a between subjects design that allows us to directly compare the behavior of subjects across treatments. Each session consists of 50 independent rounds. Our experimental design is closely related to the work of Heinemann et al. (2004) (HNO04 henceforth), who test the predictions of the model by Morris and Shin (1998) in the laboratory (a global game without information acquisition).

There are three main treatments. In the Control treatment subjects play the game without information acquisition (as in HNO04) with an exogenously fixed and commonly known precision for the signal distribution. This precision coincides with the equilibrium precision of the game with information acquisition (level 4). Since our experimental design differs from the one presented in HNO04 in the number of subjects in each group and in the distributional assumptions, we cannot directly compare our results with theirs. Thus, the first treatment serves as a benchmark for our subsequent analysis.

In the second treatment subjects play the speculative attack game with costly information acquisition of section 2. In this treatment subjects choose from a set of discrete precisions with no default option, as described in Table 1, then receive a signal, and finally they choose an action. We refer to this as the treatment with Endogenous Precision and Action choices (EPA).

The purpose of treatment 3 is to observe the evolution of thresholds across rounds. In this treatment subjects play the speculative attack game with information acquisition using the strategy method for the second stage. The precision choice stage is the same as in the EPA treatment, but in the second stage subjects have to choose a cutoff value such that they would attack if their signal

[20] Note that in our case with 2 players the risk dominant equilibrium coincides with the prediction of global games in the limit, as the noise vanishes. Moreover, there is experimental evidence that shows that the risk dominant equilibrium is often selected in 2×2 coordination games (see Cabrales et al, 2000).

[21] Instructions for all treatments can be found in:
http://econweb.ucsd.edu/~itrevino/pdfs/instructions_st.pdf .

was higher than this cutoff and not attack if their signal was lower than the cutoff they report. Hence, this treatment allows us to observe thresholds directly, rather than infer them as in the EPA treatment.[22] We refer to this as the treatment with Endogenous Precision and Strategy method choices (EPS).

On top of these three main treatments we also run additional sessions to assess two alternative hypotheses to Hypothesis 2 that could explain departures from the theory (explained in detail in section 4). For the first alternative hypothesis we run sessions of the Control treatment with exogenous high and low precisions and we refer to these two extra treatments as Control High and Control Low. For each of these treatments there are sessions in which subjects choose actions directly and sessions that use the strategy method for thresholds. We also run an additional session for the original Control treatment (with the equilibrium precision) using the strategy method. To study the second alternative hypothesis discussed in section 4 we run one session with complete information where subjects observe the true value of θ before making a decision and we call this the complete information treatment.

Overall, we ran 15 sessions, leading to a total of 292 participants. Table 2 summarizes our experimental design.

Treatment	# Sessions	# Subjects	Signal precision	Choice of action
1: Control	2	40	Level 4	Direct choice
2: EPA	2	40	Endogenous	Direct choice
3: EPS	3	44	Endogenous	Strategy method
Additional treatments:				
Control - Strategy method	1	24	Level 4	Strategy method
Control High	2	38	Level 1	Direct choice
Control High - Strategy method	1	20	Level 1	Strategy method
Control Low	2	44	Level 6	Direct choice
Control Low - Strategy method	1	20	Level 6	Strategy method
Complete information	1	22	Perfect	Direct choice

Table 2: Experimental design

Subjects are randomly matched in pairs at the beginning of the session and play with the same partner in all rounds.[23] To avoid framing effects, the instructions use a neutral terminology. Subjects are told to choose between two actions A or B, avoiding terms such as "speculative attack". To avoid bankruptcies, subjects enter each round with an endowment of 24 tokens from which they

[22] This feature of our treatment is related to the study of Duffy and Ochs (2012) who use the strategy method to elicit thresholds in coordination games. See Brandts and Charness (2011) for a survey on the strategy method.

[23] We choose fixed pairs, as opposed to random pairs, to be able to study coordination of information choices over time. As is well understood in the experimental literature, there is a trade-off between these two matching protocols. Random matching preserves better the spirit of a one shot game. However, due to the complexity of our setup, subjects might need to learn about the type of opponent they are facing and about the game itself. Our game is a double coordination game in actions and precision choices and we believe fixed pairs are better suited to study such environments. The complexity of our setup also makes it unlikely for subjects to treat it as a repeated game. In particular, we find no evidence of agents using punishment schemes in their strategies across rounds.

subtract their precision cost. From Table 1 we can see that even if subjects choose the most precise information, the lowest payoff they can get in a round is zero (in case of an unsuccessful attack).

Before starting the first paying round, subjects have access to a practice screen where they can generate signals for the different available precisions and they are given an explanation of the payoffs associated to each possible action, given their signal and the underlying state θ. They can spend up to 5 minutes in the practice screen.

Each round of the game with information acquisition consists of two decision stages:

1. Subjects choose from a menu of precisions and associated costs (see Table 1).

2. Subjects observe a signal $x_i \sim N(\theta, \sigma_i^2)$ and simultaneously decide whether to attack (A) or not attack (B).

As stated above, treatment EPS differs from treatment EPA through the use of the strategy method in the second step. For the Control treatment, where subjects do not choose precisions, they receive their private signal drawn independently from the same normal distribution with mean θ and standard deviation of 10, which is the optimal level of precision dictated by the theory for that set of parameters.[24]

After each round, subjects receive feedback about their own choice of precision, their own private signal, their choice of action, the realization of θ, how many people in their pair chose A, whether A was successful or not, and their individual payoff for the round. In addition, each subject can access the history of precision choices made over the previous rounds by pressing a button.

The computer randomly selects five of the rounds played and subjects are paid the average of the payoffs obtained in those rounds, using the exchange rate of 3 tokens per 1 US dollar.

The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007).

3.2 Experimental results

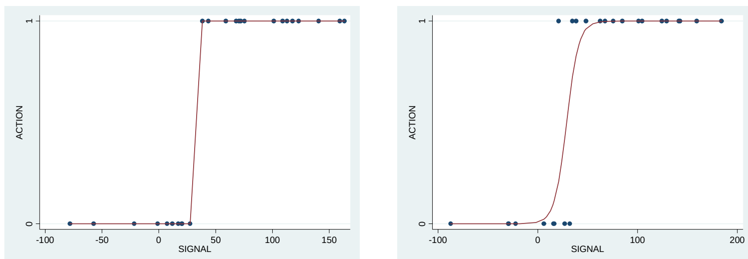
We first address the three hypotheses stated in section 2.3 that derive from the theory. This is followed by a welfare analysis and some additional results about the stability of thresholds over time that enrich the predictions of our model. The analysis is based on behavior in the last 25 rounds of the experiment (once behavior has stabilized), unless otherwise specified.

Hypothesis 1 *For any precision choices, subjects use threshold strategies.*

We find strong support for Hypothesis 1 in the data for both the Control and the EPA treatments. We say that a subject's behavior is consistent with the use of threshold strategies if the subject uses either perfect or almost perfect thresholds. A subject uses a perfect threshold if he does not attack for low values of the signal and attacks for high values of the signal, with exactly

[24] The additional sessions of the control treatment were run with exogenously given precision levels 1 and 6 from Table 1.

one switching point (i.e. the set of signals for which a subject attacks and the set of signals for which he does not attack are disjoint). This type of behavior is illustrated in panel (a) of Figure 1, which has the signal a subject receives on the horizontal axis and a binary value (0 for not attack, 1 for attack) on the vertical axis. For almost perfect thresholds, we allow these two sets to overlap for at most three observations. This means that subjects do not attack for low signal values and attack for high signal values, but these two sets can intersect for at most three observations. Such behavior is portrayed in panel (b) of Figure 1 where we fit a logistic function to the observed data of a specific subject.



(a) Perfect threshold

(b) Almost perfect threshold

Figure 1: Examples of perfect and almost perfect thresholds

In our Control treatment, we find that 92.5% of the behavior is consistent with the use of threshold strategies.[25] In particular, we find that 67.5% of the subjects use perfect thresholds, and 25% use almost perfect thresholds.

Once we have identified the subjects who use threshold strategies, we use two different methods to estimate their mean thresholds. For the first method we pool the data of all the subjects who use thresholds in each treatment and fit a logistic function with random effects (RE) to determine the probability of attacking as a function of the observed signal.[26] The cumulative logistic distribution function is defined as

$$\Pr(A) = \frac{1}{1 + \exp(\alpha + \beta x_i)}$$

By fitting the pooled data to a logistic function, we estimate the mean threshold of the group by finding the value of the signal for which subjects are indifferent between attacking and not attacking, i.e. the value of the signal for which they would attack with probability $\frac{1}{2}$, which is given by $-\frac{\alpha}{\beta}$. As pointed out by HNO04, the standard deviation of the estimated threshold, $\frac{\pi}{\beta\sqrt{3}}$ is a measure of coordination and reflects variations within the group. We call this the Logit (RE) method.

For the second method we take the average, individual by individual, between the highest value of the signal for which a subject chooses not to attack and the lowest value of the signal for which

[25] This result is similar to HNO04. However, HNO04 use a different metric to measure the use of threshold strategies and have 10 decision situations in each round of the experiment.

[26] For the EPA treatment we pool the data of subjects according to the level of precision chosen.

he chooses to attack. This number approximates the value of the signal for which he switches from taking one action to taking the other action. Once we have estimated individual thresholds using this approach, we take the mean and standard deviation of the thresholds in the group. We refer to this estimate as the Mean Estimated Threshold (MET) of the group.

The estimated mean thresholds for the Control treatment are shown in the middle column of Table A.1 in the appendix, corresponding to a medium precision. Standard deviations are reported in parenthesis. These thresholds are not statistically different from each other and coincide with the risk dominant equilibrium predicted by the theory, i.e. the optimal threshold in the limit, as noise in the private signal vanishes.[27] This result is also consistent with the findings of HNO04, therefore, we can conclude that their results extend to our framework.[28]

We now proceed to analyze the use of thresholds for the EPA treatment when subjects choose their precision. Since optimal thresholds depend on the precision chosen, we must ask whether, conditional on the choice of precision, subjects use a threshold strategy.[29] Therefore, to illustrate the use of threshold strategies we first establish stability of individual precision choices and then analyze actions for the precision level that was mostly chosen by each subject.

We find that, on average, in the last 25 rounds of the experiment subjects choose the same level of precision for more than 22 out of the last 25 periods and that the most popular precision choice is the equilibrium level 4. To illustrate this result, in figure 2 we report the transition matrix of precision choices in the last 25 rounds.[30] The entry a_{ij} of the matrix shows the probability of choosing precision level j in $t+1$, given that a subject chose precision level i at t, for $i, j \in [1, 6]$ and $t > 25$.

$$
\begin{array}{c c}
 & \begin{array}{cccccc} \text{Prec 1} & \text{Prec 2} & \text{Prec 3} & \text{Prec 4} & \text{Prec 5} & \text{Prec 6} \end{array} \\
\begin{array}{c} \text{Prec 1} \\ \text{Prec 2} \\ \text{Prec 3} \\ \text{Prec 4} \\ \text{Prec 5} \\ \text{Prec 6} \end{array} &
\begin{pmatrix}
\mathbf{0.95} & 0.03 & 0 & 0 & 0 & 0.02 \\
0.08 & \mathbf{0.74} & 0.08 & 0.05 & 0 & 0.05 \\
0 & 0.02 & \mathbf{0.87} & 0.09 & 0 & 0.02 \\
0 & 0 & 0.04 & \mathbf{0.92} & 0.02 & 0.02 \\
0.01 & 0.01 & 0.10 & 0.15 & \mathbf{0.58} & 0.15 \\
0.01 & 0 & 0.02 & 0.04 & 0.03 & \mathbf{0.90}
\end{pmatrix}
\end{array}
$$

Figure 2: Transition matrix of precision choices in the last 25 rounds, EPA and EPS treatments

By looking at the diagonal entries of the transition matrix, we can see that most precision levels (except for level 5) are absorbent states.[31] This effectively means that precision choices are stable

[27] In particular, we cannot reject the hypothesis that the theshold estimated with the logit is different from the risk dominant threshold to the 1% level of significance.

[28] In our case, because we have two players, this threshold corresponds to the risk dominant equilibrium of the underlying complete information game, see Lemma 2

[29] If individual precision choices were unstable over the last 25 periods, it would be difficult to condition on a precision choice, since it would be constantly changing.

[30] This includes precision choices in treatments EPA and EPS. We aggregate the data because the distributions of precision choices are not statistically different between these two treatments. This was expected since the treatment effect is in the second stage of the game.

[31] Very few subjects choose precision level 5 and their behavior in the second stage is mostly random.

over the last 25 rounds, i.e. that subjects on average choose their precision consistently with respect to their own previous choices.

Given this stability result, we characterize subjects by their preferred precision choice. Table 3 shows the percentage of subjects that choose each precision level for the last 25 rounds of the experiment.[32]

Precision level	Standard deviation	Cost	Precision choices in last 25 rounds
1	1	6	14.7%
2	3	5	3.7%
3	6	4	18.4%
4	10	2	36.9%
5	16	1.5	3.9%
6	20	1	22.4%

Table 3: Precision choices in the last 25 rounds, for EPA and EPS treatments

We now analyze subject behavior in the EPA treatment. We identify three types of subjects: those who use thresholds, those who use a degenerate strategy, and those who act randomly. The use of thresholds can be either by perfect or almost perfect thresholds, as defined above. We say that a subject uses a degenerate strategy if his choice of action is constant and does not vary with the signals (i.e. always attack or never attack). For a subject that exhibits random behavior, the choice of action is independent of his signal realization.

As shown in Table 4, we find that 100% of subjects choosing precision levels 1, 2, or 3 use threshold strategies. For precision level 4, all but one subject use threshold strategies, which corresponds to 93.75% of the subjects. There is only one subject choosing precision level 5, whose behavior is random. For precision level 6, 75% of the subjects use threshold strategies and 25% use degenerate strategies. This suggests that when subjects invest in more precise information their behavior is more likely to be consistent with the use of threshold strategies.

In total, 90% of the subjects in the EPA treatment use threshold strategies for their most preferred precision choice. This implies that the theoretical prediction of a threshold strategy is robust also under endogenous information.

Hypothesis 2 *Thresholds are increasing functions of precisions.*

We first analyze how the individual decision to attack depends on individual precision choices. Then we study behavior at the pair level to establish how the joint decision to attack (and hence the success of attacks) depends on the precision chosen by both pair members.

[32] Precision choices in the first rounds are not very dissimilar to the precision choices portrayed in table 3. In particular, if we compare the choices of the first 5 rounds with the choices of the last 5 rounds of the experiment, we observe the following proportion of choices, by precision level (the first number corresponds to the first 5 rounds and the second to the last 5 rounds). Level 1: 16.2% vs 13.8%; level 2: 11.9% vs 4%; level 3: 25% vs 19%; level 4: 21.4% vs 36.8%; level 5: 4.8% vs 4%; level 6: 20.7% vs 22.4%. We observe the highest shift in precision choices to be in favor of the equilibrium precision level 4.

Precision	Thresholds	Degenerate strategies	Random behavior
1	100%	-	-
2	100%	-	-
3	100%	-	-
4	93.75%	-	6.25%
5	-	-	100%
6	75%	25%	-

Table 4: Strategies followed in the second stage, for EPA treatment

Figure 3 plots the cumulative distribution function (pooled over all subjects in treatment EPA) to illustrate the probability of attacking for each signal realization, by precision levels. The value of the signal for which subjects attack with probability 0.5 determines the threshold. Looking at the intersection of the curves corresponding to the different precision levels with the 0.5 horizontal line, from left to right, we can see that the values of the estimated thresholds are larger for lower precisions.[33] This suggests that the subjects who acquire more precise information attack more often. We also see that as we move towards lower precision levels the slopes of the CDFs are decreasing, indicating that the dispersion of the observations within each precision group increases.

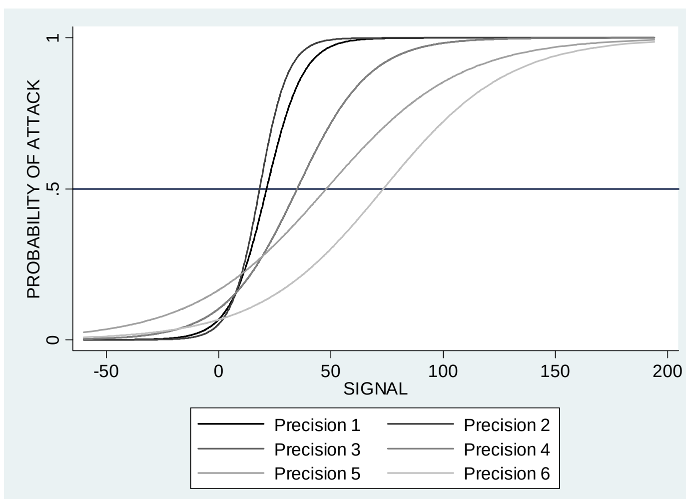


Figure 3: Probability of attack, by precision choices in EPA treatment

In order to better understand this phenomenon, we also run two regressions (one for the EPA and one for the EPS treatment) to determine the statistical effect that each level of precision has on the decision to attack (for EPA) and on the reported thresholds (for EPS).

For EPA, we estimate a random effects logit where the dependent variable is the decision to attack (1) or not attack (0) and the independent variables are dummies for the six precision levels, interacted with the signal realizations. We interact precisions and signals because the decision to

[33] This is true except for precision level 2, but less than 4% of the subjects chose this level in the EPA treatment.

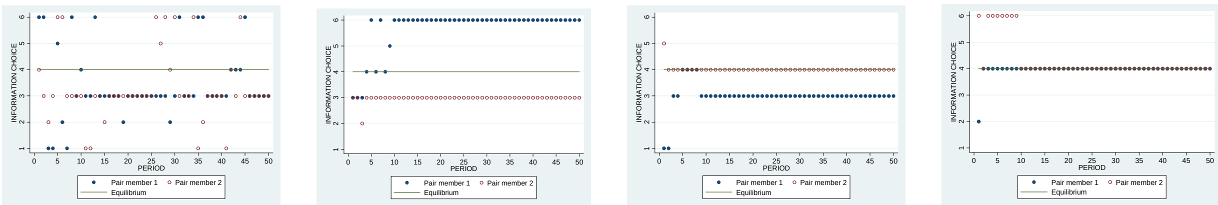
attack is determined by the value of the signal, and the information choice affects how precise the signal is. As shown in Table A.2 in the appendix, we find that all the coefficients for the interacted variables are positive and significant to the 1% level, and that the magnitudes of the coefficients decrease for lower precisions. The positive coefficients mean that subjects attack for higher signal realizations, for all precision levels, which is consistent with the monotonicity implied by threshold strategies. The fact that the magnitudes of the coefficients decrease for less precise information implies that this effect is stronger when subjects observe very precise signals than when they observe noisier signals. If we categorize precision levels into high precision (levels 1 and 2), medium precision (levels 3 and 4), and low precision (level 6),[34] we find that the coefficients for high precision are statistically smaller than the coefficients for medium precision (at the 10% level of significance), and that the coefficients for medium precision are statistically smaller than the coefficient for low precision (at the 1% level of significance). We can therefore conclude that subjects that choose higher precision levels attack more often.

We find further evidence of this effect in the EPS treatment by analyzing how reported thresholds vary with precision choices. Table A.3 in the appendix reports the results of a random effects OLS regression where the dependent variable is the threshold reported by subjects and the independent variables are dummies for each level of precision. Each of these dummies takes the value of 1 if the subject chooses this precision level and 0 otherwise. We find that the reported thresholds depend positively and significantly on the level of precision chosen. In particular, the magnitudes of the coefficients for each precision level increase as we move towards less precise information, suggesting that less precise information gives rise to higher thresholds, i.e. a subject with a lower precision attacks less often than a subject with a higher precision. This corroborates the findings in the EPA treatment and, together, these results are at odds with Hypothesis 2, which states that higher precisions lead to higher thresholds.

In order to fully test Hypothesis 2 we need to study the behavior at the pair level because thresholds depend on the precision choices of both pair members. Therefore, we need to establish first a notion of convergence in precision within a pair to understand how joint precision choices affect the incidence and success of attacks. We separate precision choices within a pair in four categories, which are illustrated in Figure 4. We define individual convergence in precision as a situation where a subject chooses the same precision level for the last 25 rounds, with at most three deviations. We say that a pair exhibits non-stable behavior if at least one of its members does not converge individually in his precision choice (panel (a)). A pair that has stability but not convergence is a pair in which both members converge individually in their own precision choices, but the levels at which they converge are more than one level apart (panel (b)). We define weak convergence as pairs in which both members converge individually to a level of precision and these two precision levels are at most one level away from each other (panel (c)). We say that a pair

[34] We do not include precision level 5 since there is only one subject choosing it, who behaves randomly in the second stage.

exhibits full convergence if both members converge individually to the same level of precision for the last 25 rounds of play (panel (d)).[35]



(a) Non-stable behavior (b) Stability, no convergence (c) Weak convergence (d) Full convergence

Figure 4: Examples of types of convergence in precision

In what follows, we use the notion of weak convergence and we restrict our attention to pairs that coordinate on high precision (levels 1 and 2), medium precision (levels 3 and 4), and low precision (levels 5 and 6). Note that weak convergence includes full convergence. Table 5 summarizes the combinations of precision choices across pairs according to this notion of weak convergence. If we use this qualitative characterization we find that approximately two thirds of the pairs in the endogenous precision treatments exhibit weak convergence in precision choices. Moreover, the majority of the pairs converge to a medium precision, which corresponds to the theoretical prediction.

	High	Medium	Low
H	**10.00%**	13.81%	3.05%
M		**40.00%**	16.67%
L			**16.48%**

Table 5: Weak convergence of precision choices, EPA and EPS treatments

Given this notion of convergence in precision choices, we now characterize how the incidence and success of attacks is determined by precision choices in a pair. In tables 6 and 7 we report, for the EPA and EPS treatments respectively, the number of attacks and the number of successful attacks as proportions of the total number of situations when subjects are in the coordination region $\theta \in [0, 100)$, for each combination of precision choices within pairs.[36] The number of attacks corresponds to the instances when at least one of the pair members decides to attack and $\theta \in [0, 100)$, while the number of successful attacks corresponds to the number of cases where both subjects attack and $\theta \in [0, 100)$.

From tables 6 and 7 we see that in both treatments the incidence and success of speculative attacks increases with the precision at which subjects converge (diagonal entries), which is in line with our findings at the individual level and establishes that our data does not find support for

[35] Table A.4 in the appendix shows all the combinations of precision choices made by the different pairs in our experiment (for both the EPA and EPS treatments). The diagonal entries correspond to the pairs that exhibit full convergence.

[36] We thank readers of a previous version for suggesting these tables.

Hypothesis 2. We provide further evidence for this result in Table A.5 in the appendix where we report the results of a random effects probit regression where we test how the success of an attack is determined by the level of precision to which each pair weakly converges for the EPA treatment. In section 4 we study possible explanations for this departure from the theoretical predictions.

	Number of attacks / Total situations			Number of successful attacks / Total situations		
	High	Medium	Low	High	Medium	Low
H	$\frac{23}{29}$ (79.31%)	$\frac{41}{44}$ (93.18%)	$\frac{2}{2}$ (100%)	$\frac{18}{29}$ (62.07%)	$\frac{35}{44}$ (79.54%)	$\frac{2}{2}$ (100%)
M		$\frac{82}{115}$ (77.39%)	$\frac{36}{68}$ (52.94%)		$\frac{59}{115}$ (51.30%)	$\frac{18}{68}$ (26.47%)
L			$\frac{16}{34}$ (50.00%)			$\frac{5}{34}$ (14.7%)

Table 6: Incidence and success of attacks, by precision choices, EPA treatment

	Number of attacks / Total situations			Number of successful attacks / Total situations		
	High	Medium	Low	High	Medium	Low
H	$\frac{35}{37}$ (94.59%)	$\frac{43}{55}$ (78.18%)	$\frac{7}{19}$ (36.84%)	$\frac{34}{37}$ (91.89%)	$\frac{33}{55}$ (60%)	$\frac{3}{19}$ (15.79%)
M		$\frac{125}{149}$ (83.89%)	$\frac{40}{49}$ (81.63%)		$\frac{99}{149}$ (66.44%)	$\frac{35}{49}$ (71.43%)
L			$\frac{66}{77}$ (85.71%)			$\frac{48}{77}$ (62.34%)

Table 7: Incidence and success of attacks, by precision choices, EPS treatment

Hypothesis 3 (Equilibrium) *Subjects coordinate on a precision level of 4 and set a symmetric threshold of 28.31.*

In Table 8 we compare, for high, medium, and low precision levels, the mean estimated thresholds (MET) and the random effects logit estimations from the EPA treatment, and the mean reported thresholds (MRT) from the EPS treatment with the theoretical predictions. Remember that we define weak convergence to high precision as pairs that converge to precision levels 1 or 2, medium precision as pairs that converge to precision levels 3 or 4, and low precision as pairs that converge to precision levels 5 or 6. Therefore, for each precision level (high, medium, or low) we include the two predictions that correspond to each of the precision levels (1 and 2, 3 and 4, or 5 and 6), as well as the risk dominant equilibrium, i.e. the threshold prediction when the signal noise converges to zero. Standard deviations are reported in parenthesis.

We find that the logit and MET estimates for medium precision in the EPA treatment are not statistically different from the predictions from the theory for the equilibrium precision level 4, which is a threshold of 28.31. This means that the subjects that coordinate on medium precision

	Full information	High precision	Medium precision	Low precision
Logit (RE) (EPA)	22.77	24.99	30.30	74.62
	(11.76)	(7.44)	(12.32)	(21.48)
MET (EPA)	21.07	25.29	27.84	50.65
	(11.85)	(9.27)	(17.65)	(28.65)
MRT (EPS)	-	19.77	38.23	32.13
	-	(1.14)	(13.24)	(12.19)
Theoretical prediction x^*		Info 1 Info 2	Info 3 Info 4	Info 5 Info 6
		35.31 33.88	31.61 28.31	22.82 18.73
Risk dominant equilibrium	36	36	36	36

Table 8: Estimated thresholds and equilibrium predictions

levels behave on average in accordance to the unique equilibrium suggested by the theory, unlike those who converge to either a high or low precision.

Table 8 and Figure 5 illustrate how subjects with high and low precisions set thresholds in the opposite direction of what is predicted by the theory. Figure 5 plots the theoretical thresholds and the mean estimated thresholds of the EPA treatment for each precision level. We can see in this figure that the theory predicts thresholds to increase with higher precisions, while the experimental data suggests the opposite. However, note that the theoretical predictions and the mean estimated thresholds coincide for the equilibrium precision (level 4). This means that subjects who coordinate on precision level 4 understand the trade off between precision and the cost associated to it by choosing a medium level of precision, and then apply this information optimally in the speculative attack game by choosing the threshold level that maximizes their expected profits, given that their opponent also chooses a medium level of precision.

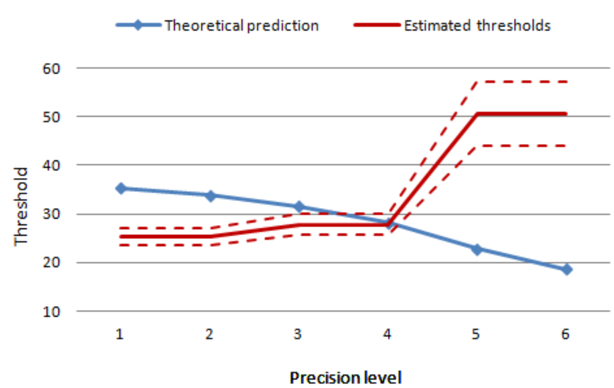


Figure 5: Theoretical and estimated thresholds, EPA treatment

3.2.1 Payoff efficiency

We now analyze whether the actions taken by the subjects in our experiment increase their payoffs with respect to the equilibrium actions predicted by the theory and to two other benchmarks of surplus extraction. Table 9 compares the average realized payoffs of subjects for the treatments with endogenous precision choices (EPA and EPS, separated by precision choices) to three benchmarks.[37] The first one is the average expected payoff that would have arisen if subjects had followed the equilibrium strategy for each realization of the state θ observed in the different sessions of the experiment. The second is a constrained efficiency benchmark where agents truthfully reveal their signals and jointly choose actions to extract the maximum surplus, for each realization of θ observed in the experiment. This means that subjects would still face fundamental uncertainty and would have to purchase precisions for their signals to deal with it, but they would not face strategic uncertainty (it is as if a planner could choose actions for both pair members to extract the maximum surplus). In this case, we find that the optimal precision choice would be level 6. The payoffs we report for these two benchmarks are built using the observed realizations of θ from the experiment. For each θ we get the distribution of signals that would correspond to the precision level that is chosen in each benchmark (levels 4 and 6, respectively). We do not use the observed signals from the experiment to calculate payoffs because for each benchmark a specific precision is assumed to be chosen and each precision gives rise to a different distribution of signals.

The third benchmark corresponds to the average payoffs that would have arisen if subjects had chosen a "first-best" action under complete information, i.e. if they had attacked whenever they could get a positive payoff ($\theta > 18$). This first-best definition is similar to the setup where a social planner can observe the realizations of θ and prescribe the actions that would maximize the payoffs of both agents without informational constraints. This would correspond to the case where subjects would attack as long as $\theta \geq 18$ and would not attack otherwise. Therefore, only the realization of θ in the experiment is used to construct this benchmark and no signals are taken into consideration. For this reason, we assume that subjects do not pay a cost to improve the precision of their signals. Standard deviations are reported in parenthesis.

Clearly, among all of these benchmarks, the highest possible payoffs come from the first-best outcome with complete information, followed by the constrained efficiency benchmark, and finally followed by the equilibrium play. The results in Table 9 show that subjects who choose a high level of precision increase their payoff with respect to the constrained efficiency and the equilibrium play in the EPA treatment. For the EPS treatment we can see that the realized and the equilibrium payoffs are not statistically different. We can conclude that subjects who choose a high precision see a Pareto improvement with respect to equilibrium due to the departure from the theory that leads them to set lower thresholds and attack more often.

For subjects with medium and low precisions equilibrium payoffs are higher than realized payoffs.

[37] We look at net realized payoffs substracting the cost of acquiring information and the cost of attacking, if applicable.

	High precision		Medium precision		Low precision	
Treatment	EPA	EPS	EPA	EPS	EPA	EPS
Realized payoffs	32.94	38.76	26.16	38.08	18.11	37.63
	(7.25)	(7.37)	(8.61)	(4.14)	(12.97)	(4.30)
Expected equilibrium	31.06***	39.40	30.83***	39.20***	33.63***	39.46***
payoffs	(6.65)	(0.84)	(6.66)	(0.66)	(5.75)	(0.90)
Expected constrained	31.21***	39.76**	30.97***	39.56***	33.84***	39.84***
efficient payoffs	(6.83)	(0.75)	(6.84)	(0.58)	(5.90)	(0.79)
First-best complete	35.25***	43.62***	35.02***	43.43***	37.80***	43.70***
information payoffs	(6.61)	(0.71)	(6.62)	(0.56)	(5.71)	(0.74)

Statistically different from realized payoffs at the ***1%; **5%; *10% level of significance

Table 9: Average payoffs

However, for the EPA treatment, realized payoffs for subjects with medium precision are at most 13.4% less than the corresponding equilibrium payoffs, whereas realized payoffs for subjects with low precisions are up to 45.9% less than the corresponding equilibrium payoffs.[38] This effectively means that choosing a low precision leads to the highest loss in individual payoffs with respect to equilibrium.

In our game an increase in the incidence of speculative attacks for subjects who coordinate on a high precision improves individual payoffs because this high precision minimizes the probability of making the mistake of attacking when fundamentals are strong. This analysis is related to the findings of the previous section that show a departure from the theory in the way that precisions determine thresholds. As Table 9 indicates, it is beneficial for subjects who coordinate on a high precision to set a lower threshold than what is predicted by the theory. By doing this, both pair members pay the high cost to get the most precise signal in order to coordinate on attacking for a wider range of fundamentals. In theory, the highest precision is not an equilibrium because there is an individual incentive to deviate and save some of the cost, while still getting the benefit of having a very well informed pair member that will make more accurate decisions. However, by coordinating on the highest precision both pair members get more accurate signals, thus reducing the probability of making the mistake of attacking when an attack is unsuccessful. Clearly, such a cooperative agreement results in a higher payoff than the one predicted in equilibrium.

3.2.2 Welfare

Even though individual investors extract a higher surplus by choosing a higher precision, it is important to keep in mind that, for the specific context of our model, speculative attacks that

[38] Even if subjects who choose equilibrium precisions also set equilibrium thresholds, we can still observe differences in realized payoffs with respect to the expected equilibrium payoffs. This is due to the fact that we are comparing payoffs from the realization of θ for a small sample to the expected payoffs that would arise according to the distribution of a population. By a similar argument, constrained efficiency payoffs should, on average, be higher than realized payoffs, but over finite samples this might not be the case.

lead to a devaluation can be detrimental for society. This means that an increase in an individual investor's welfare is not necessarily beneficial to other members of society (non-investors). If one were to model society's welfare in this context, we could think that the interests of society are aligned to those of the central bank. In this sense, for very bad states of the economy ($\theta \geq 100$) it is not optimal for the central bank to defend the currency, so a devaluation is the best outcome. For very good states of the economy ($\theta < 0$) the central bank always defends the currency, so a devaluation never takes place. However, when devaluations are driven by self-fulfilling beliefs (coordinated attacks for $\theta \in [0, 100)$) they can be detrimental to society because, in principle, the central bank would have been able to defend the currency and avoid the instability caused by the devaluation if an attack had not occurred. To analyze social welfare in this context, we construct a measure with our data by looking at the number of successful attacks as a fraction of total realizations of $\theta \in [0, 100)$. This would serve as an index of society's welfare loss (the higher the number of coordinated attacks, the higher the loss in welfare for society).

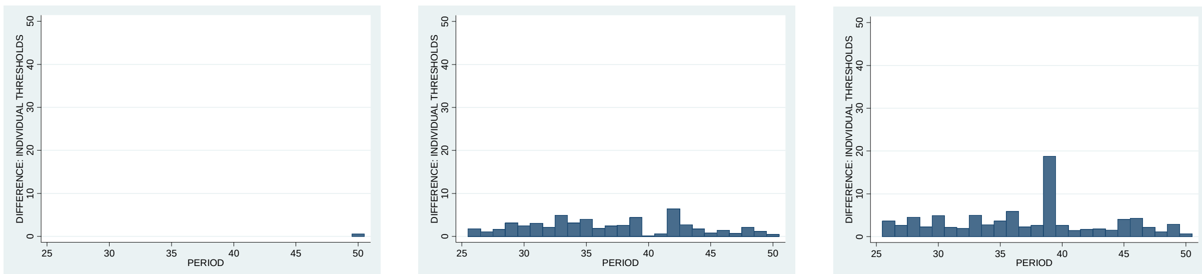
Tables 6 and 7 report the indices of society's welfare loss for treatments EPA and EPS, respectively, for each level of precision. We see that there is a higher welfare loss for society when investors coordinate on a higher precision and that the smallest loss of welfare for society arises when investors choose lower precisions, since a high precision leads to a higher rate of successful attacks. This can be seen by comparing the rate of successful attacks for the EPA treatment to be 62.07% for high precision, 51.3% for medium precision, and 14.7% for low precision. Likewise, for the EPS treatment we observe the rate of successful attacks to be 91.89% for high precision, as opposed to 66.44% and 62.34% for medium and low precisions, respectively. By looking at these numbers it is evident that the interests of individual investors and the rest of society are not aligned in this model, so that when speculators coordinate on a high precision they improve their individual welfare by extracting a higher surplus, but this translates into a reduction of social welfare.[39] This illustrates the trade-off between investors' and society's welfare in episodes of high speculation and endogenous precision choices.

3.2.3 Stability of thresholds

We now analyze the stability of reported thresholds across rounds for the EPS treatment for the pairs that converge to high, medium, and low levels of precision. First, we analyze the evolution of individual thresholds by taking the average period-to-period differences (in absolute value) of the individual reported thresholds in each of the last 25 periods of the experiment, for subjects in pairs that coordinate on high, medium, and low precision. We calculate, for each subject, the difference in absolute value between the threshold chosen in one period with respect to the previous period, and then we take the average of these differences across subjects from pairs that coordinate on

[39] Note that our results can also be applied to different contexts, such as an investment decision or a social revolt. In these cases the interests of society would be aligned to the interests of the agents in the model. For example, a successful investment would be beneficial to both individual investors and society.

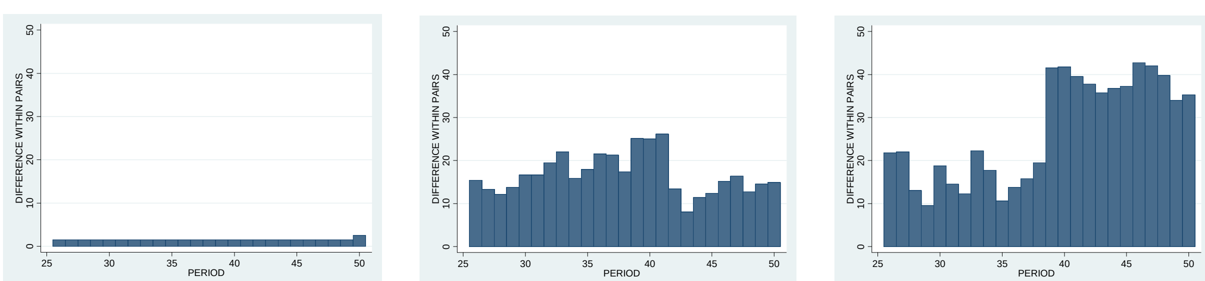
each precision level. This is portrayed in Figure 6. The vertical bar at period t illustrates how much, on average, a subject changed the value of his own threshold in period t with respect to the threshold he reported in period $t-1$. We find that pairs that coordinate on high precision levels converge individually in their thresholds, since the mean difference is consistently zero, except for the last round. For medium precision levels the mean difference is less than 7 in all periods, and for low precision levels it is less than 7 in all but one period. Therefore, convergence in individual thresholds is less stark as subjects choose lower precisions.



High precision Medium precision Low precision

Figure 6: Convergence of individual thresholds, EPS treatment

Given this behavior, we also plot the difference in absolute value between the reported thresholds of the members in each pair, for each precision level. In Figure 7 we see, for each period and each precision level, the average difference in absolute value between the reported thresholds of both members of each pair, i.e. we observe by how much subjects coordinate on the same threshold with their pair member. Recall that subjects do not receive feedback about the threshold reported by their pair member, but only observe whether action A (attack) was successful or not. We can see that subjects from pairs that converge to high precision levels seem to coordinate also on their thresholds, and that this is less clear as we move towards less precise levels of information. Figure 7 suggests thus that convergence to a unique threshold within a pair is more likely as subjects coordinate on more precise information.



High precision Medium precision Low precision

Figure 7: Convergence of thresholds within a pair, EPS treatment

4 Discussion

The most striking finding of our analysis above is that, opposite to the predictions of the theory, subjects who choose and coordinate on the highest precision set lower thresholds than those subjects who coordinate on lower precisions. That is, contrary to the theoretical predictions, our experimental results show that a higher precision level leads subjects to behave closer to the efficient equilibrium under complete information, and further away from the risk dominant one. This is an important departure from the theoretical framework, since it has direct implications about the consequences of more precise information on welfare. It is thus crucial to understand this departure from the theory. Below, we investigate several potential explanations of our experimental results.

4.1 Sunk cost and forward induction

One possible explanation for this departure from the theory is that subjects do not take the cost of precision as a sunk cost. Instead, subjects that choose a high precision might take into account the high cost that they have paid for information when they make their decision to attack, i.e. they attack more often to make up for the high cost paid.[40] Alternatively, subjects might use a high precision to signal to their opponents that they are willing to set lower thresholds and, thus, make the most out of the high cost paid (a forward induction reasoning).

To address these alternative explanations for our results, we run extra sessions with exogenous high (level 1) and low (level 6) precisions to shut down the active role of precision choice on threshold formation. We run sessions where subjects choose actions (as in the Control treatment, where subjects had an exogenously given medium precision, at level 4) and sessions with the strategy method to elicit thresholds directly.[41] We find that the thresholds of subjects that are given either a medium or a low precision exogenously are not statistically different from each other, but when subjects are exogenously endowed with a high precision they choose a significantly smaller threshold than for medium and low precisions, just as in the case of endogenous precision choices. This implies that the sunk cost hypothesis and the forward induction argument do not explain our findings. Instead, these sessions provide more evidence that subjects seem to respond to more precise information in the opposite direction from what is predicted by the theory in a systematic way, i.e. when precision is chosen by them or is given to them exogenously.

4.2 Level-k reasoning

Another possible way to explain our findings is to think about how a higher precision of signals influences the level of reasoning of subjects when forming thresholds. In our model we assume that agents are rational in their belief formation process and this leads them to choose the actions

[40] There is numeruous evidence from psychology that suggests that subjects are willing to take riskier choices to make up for sunk costs, see for example Arkes and Blumer (1985). In economics Thaler (1980) uses prospect theory to explain the sunk cost fallacy.

[41] We report the estimated thresholds for these sessions in the last three columns of Table A.1 in the appendix.

predicted by the theory. However, this might not be the way in which subjects in the experiment reason when making a decision.[42]

Different models of cognition have been proposed to explain the observed departures from the theoretical predictions in strategic settings, such as level-k and cognitive hierarchy (see Nagel, 1995, Costa-Gomes and Crawford, 2006, to name a few). These models suggest that in experimental settings subjects might exhibit a bounded depth of reasoning. In a global games context (without costly information acquisition), Kneeland (2014) develops a theoretical model of level-k thinking to incorporate experimental evidence about the role of private and public information on threshold formation. The main argument of Kneeland (2014) is that the results from HNO04 are better approximated by a model of limited depth of reasoning than by equilibrium play.

We use a level-k model to investigate whether the results in our experiment are due to the subjects' limited depth of reasoning.[43] We assume that the level-0 player chooses randomly whether to attack or not, irrespective of the signal he receives, and that a level-k player best responds to the behavior of a level-$(k-1)$ player. Compared to the model analyzed in Kneeland (2014), in our setup a level-k player is described by a two-dimensional strategy $\{x_k^*, \sigma_k^*\}$ where x_k^* is a threshold set by a player with level-k depth of reasoning, while σ_k^* is his choice of standard deviation. Figure 8 presents the predictions of the level-k model applied to our setup for the parameters used in the experiment. The left panel reports precision choices σ_k^* for each level-k, while the right panel presents the corresponding threshold, $x_k^*(\sigma_k^*)$ for each level-k. We see that, given our parametric assumptions, this level-k model predicts no information acquisition to take place, and, as such, it cannot explain the findings of our experimental data.[44]

4.3 Convergence towards the equilibrium under complete information

As presented in section 2.1.1 and consistent with previous well known results of global games, the theory predicts that in the limit, as the noise of the private signals vanishes, the unique equilibrium of the game coincides with the risk dominant equilibrium of the underlying complete information game. For the parameters used in the experiment, this would imply that as the private signals become more informative the thresholds set by subjects increase towards the risk dominant threshold of 36. As reported above, this is not the case. We observe that as subjects in our experiment

[42] One might think that our results reflect the fact that subjects could be relying less on the prior than what is assumed by the theory, therefore putting more weight on the private signal. However, doing this is equivalent to behaving as if the precision of the private signal was higher than the actual precision chosen, in which case the theory would predict an even higher threshold. Thus, this type of overconfidence cannot explain our results.

[43] Alternatively, we could also think about other belief-based models that would seem appropriate to global games to explain this departure, such as understanding how the reduction in fundamental and strategic uncertainty due to a higher precision affects the belief formation process. While a systematic investigation of these issues is beyond the scope of this paper, we deem this endeavour as an exciting avenue for future research.

[44] This result does not change if we allow a level-k player to best respond to a distribution of players with lower levels of reasoning. However, as in Kneeland (2014), thresholds might be increasing or decreasing on the bound k, depending on the assumptions made for level-0 types. We view this as an important limitation of this particular model.

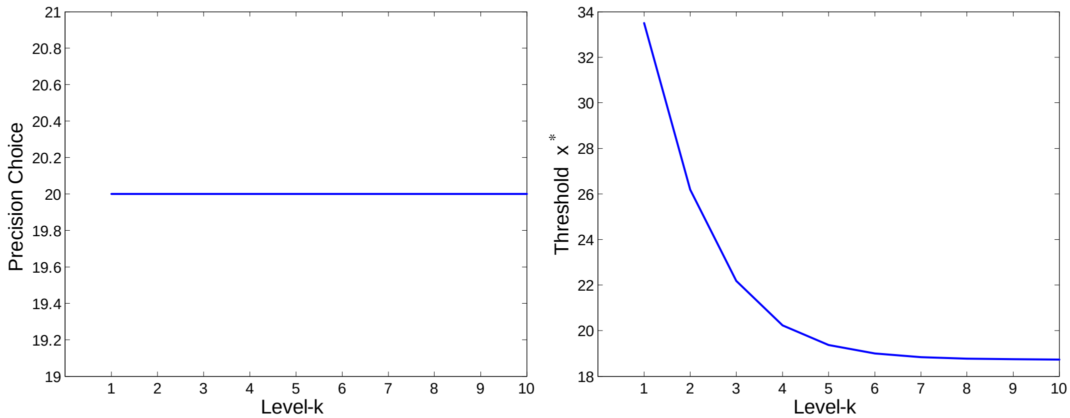


Figure 8: Level-k Thinking - Numerical Results

converge to choosing higher precisions, they set thresholds that are further away from the risk dominant threshold and closer to the efficient threshold of 18. Moreover, this result cannot be explained by a limited depth of reasoning or forward induction arguments, or by a sunk cost fallacy, and it is robust to exogenously given precisions.

In this section we explore a different possibility. We argue that as the precision of the private signals increases, the behavior of agents converges to the behavior of subjects that play the same game under complete information, which is different to the equilibrium suggested by the theory. To test this theory, we run a complete information treatment where subjects observe the state θ perfectly in every round. We then test whether the thresholds of subjects corresponding to (endogenous and exogenous) low, medium and high precisions converge towards the thresholds observed in this complete information session.[45] We find that all of the subjects in the complete information treatment use threshold strategies, and that 77.27% of them, in fact, use exactly the efficient threshold of 18.[46] That is, the vast majority of our subjects not only select one of the two equilibria of the complete information game, but more importantly, they select the efficient equilibrium, and not the risk dominant one. This is a stark contrast from the theoretical predictions for this game.[47]

In Table 10 we compare the mean estimated thresholds of all the subjects in the treatment with complete information (second column) to the mean estimated thresholds (MET) corresponding to a low, a medium and a high precision in the treatments with and without information acquisition and

[45] The details of the complete information treatment are in the last row of table 2.

[46] 86.36% of the subjects use perfect thresholds and 13.64% use almost perfect thresholds.

[47] This argument is related to the findings reported in HNO04. In their treatment with complete information, subjects set a lower threshold than in their treatment with noisy private signals. That is, under complete information, subjects not only choose a threshold strategy (which is by itself a departure from the theoretical prediction of multiplicity of equilibria under complete information), but the threshold they set is closer to efficiency than in the presence of noisy private signals. HNO04 do not invesitgate, however, the convergence behavior of thresholds as the noise in the signals vanishes.

direct action choices (EPA). We see that in the case of endogenous precision choices, the thresholds display a clear convergence towards the threshold obtained under complete information. We observe a similar convergence, albeit less strong, in the case of exogenous information. Furthermore, we can see that the mean estimated threshold from the complete information treatment is significantly smaller than the mean estimated thresholds from the sessions where subjects are exogenously endowed with a high precision and when they endogenously choose a high precision and choose actions directly. Thus, the results reported in Table 10 suggest internal consistency in subjects' behavior across treatments, in the sense that, as the precision of the signals increases, subjects' behavior converges towards the complete information behavior.

	Complete info	High Precision	Medium Precision	Low Precision
Mean estimated threshold (Endogenous Precision)	21.07	25.29	27.84	50.65
Mean estimated threshold (Exogenous Precision)	21.07	27.42	40.37	36.23

Table 10: Mean estimated thresholds, by treatment

This systematic path to convergence illustrates an underlying force in the game that is not captured by the theory. Indeed, the data shows convergence of thresholds in the limit as the noise of private signals vanishes, but the limit that we actually observe does not correspond to the limit predicted by the theory. Thus, our findings suggest that in environments with precise but incomplete information, and contrary to many existing theories of equilibrium selection (including the theory of global games), subjects behave similarly to the case of environments of complete information and common knowledge of fundamentals. This observation can have important consequences for the theoretical modelling of incomplete information. We deem further investigation of this hypothesis a fruitful avenue for future research.

5 Conclusion

We have characterized equilibrium selection in a global game with a discrete number of agents and an endogenous information structure. We show theoretically that when agents choose privately the precision of their signal, uniqueness of equilibrium in the coordination game can be guaranteed as long as the precision of the prior is diffuse enough with respect to the precision of the private signals, which is consistent with previous results in the case of symmetric noise distributions. For most ranges of parameters, strategic complementarities in information acquisition arise, leading to symmetric precision choices in equilibrium. We test the model experimentally in an effort to understand how real agents make decisions in this double coordination game.

The experimental results of our model give us further insights about the behavior during speculative attacks in which subjects are allowed to choose the precision of the signals they observe, at a cost. We find that over 30% of the subjects behave as is predicted in equilibrium since they understand the trade off between precision and cost, and thus choose a medium level of precision and coordinate on this level within their group. These subjects seem to apply correctly the information they acquire on their decision in the speculative attack by using a threshold that coincides with the unique equilibrium predicted by the theory.

However, we also find subjects who acquire high and low precision levels, which is contrary to the theoretical predictions. In our model, the theory suggests that the incidence of speculative attacks should decrease as agents coordinate on higher levels of precision, i.e. that better informed agents should be more careful and attack when they have more certainty of a successful outcome. This theoretical prediction suggests that having more informed agents would decrease the incidence and success of speculative attacks. However, our experimental results show the opposite. In our study, subjects that acquire more precise information attack significantly more often than those who acquire less precise information. As a result, the overall success of attacks increases for better informed market participants: they achieve higher individual payoffs by coordinating on successful attacks for a wider range of fundamentals.

One possible interpretation for this departure from the theory could be that subjects fail to take the cost of precision as a sunk cost, and thus attack more often when they invest in better information. The results from extra sessions that address this hypothesis show that subjects attack more often when faced with a higher precision, even if this precision is given to them exogenously at no cost.

In order to understand the departure from the theory that associates lower thresholds with higher precisions, we also study the behavior of agents under complete information. We find that, contrary to the theoretical predictions under complete information (multiplicity of equilibria), the vast majority of subjects select the efficient equilibrium. Therefore, our results show a systematic path to convergence in threshold formation (in the limit, as the noise of the private signals vanishes) towards the efficient equilibrium under complete information, which contrasts the theory that predicts convergence to the risk dominant equilibrium. This finding has important consequences from an economic theory point of view. Our results suggest that while economic theory provides a good approximation to the observed behavior in coordination games when there is a substantial amount of noise in the environment, it fails to explain the observed behavior when information is very precise or complete. Understanding what drives these differences and how to augment the theory in order to improve its predictive power remains an important challenge for future research.

In terms of welfare, our experimental results show that having better quality of information leads to a significant increase in individual investor's welfare since it allows them to extract higher payoffs than equilibrium play. In the context of a speculative attack, even if this behavior improves the welfare of individuals, an increased incidence and success of speculative attacks is detrimental for society, as proven by the numerous episodes of currency crises in the last two decades. In this sense, the results of the experiment illustrate the trade off between individual and social welfare in episodes of speculative attack when agents can improve their private information at a cost. This finding sheds light on the importance for policy making of taking into account the role that private information acquisition plays on speculative attack outcomes. However, our results can also be applied to different contexts (e.g. an investment decision or a social revolt), in which case different qualitative implications on welfare might apply.

References

[1] Angeletos, G. M. and I. Werning, 2006, "Crises and prices: Information aggregation, multiplicity, and volatility" *American Economic Review* 96(5).

[2] Angeletos, G.-M., C. Hellwig and A. Pavan, 2006, "Signaling in a global game: Coordination and policy traps" *Journal of Political Economy* 114(3).

[3] Arkes, H. and C. Blumer, 1985, "The psychology of sunk cost" *Organizational Behavior and Human Decision Processes* 35.

[4] Ash, R., 2000, *Probability and Measure Theory*, Harcourt Academic Press, 2nd ed.

[5] Bannier, C. and F. Heinemann, 2005, "Optimal Transparency and risk taking to avoid currency crises" *Journal of Institutional and Theoretical Economics* 161(3).

[6] Brandts, J. and G. Charness, 2011, "The strategy versus the direct-response method: a first survey of experimental comparisons" *Experimental Economics* 14(3).

[7] Cabrales, A., W. Garcia-Fontesa, and M. Motta, 2000, "Risk dominance selects the leader: An experimental analysis" *International Journal of Industrial Organization* 18(1).

[8] Cabrales, A., R. Nagel, and R. Armenter, 2007. "Equilibrium selection through incomplete information in coordination games: an experimental study," *Experimental Economics.* 10(3).

[9] Carlsson, H. and E. van Damme, 1993, "Global games and equilibrium selection" *Econometrica* 61(5).

[10] Colombo, L., Femminis, G. and A. Pavan, 2014, "Information acquisition and welfare", *Review of Economic Studies* 82(2).

[11] Cooper, R. W., D.V. DeJong, R. Forsythe, and T. W. Ross 1990, "Selection criteria in coordination games: Some experimental results" *The American Economic Review* 80: 218-233.

[12] Cooper, R. W., D.V. DeJong, R. Forsythe, and T. W. Ross 1992, "Communication in coordination games" *The Quarterly Journal of Economics* 739-771.

[13] Corsetti, G., A. Dasgupta, S. Morris, and H. S. Shin, 2004, "Does one Soros make a difference? A theory of currency crises with large and small traders" *Review of Economic Studies* 71.

[14] Costa-Gomes, M. and V. Crawford, 2006. “Cognition and Behavior in Two-Person Guessing Games: An Experimental Study” *American Economic Review* 96(5).

[15] Duffy, J. and J. Ochs, 2011, “Equilibrium selection in static and dynamic entry games”, *Games and Economic Behavior* 76(1).

[16] Ely, J. and M. Peski, 2006, “Hierarchies of belief and interim rationalizability”, *Theoretical Economics* 1.

[17] Fischbacher, U., 2007, “z-Tree: Zurich Toolbox for Ready-made Economic Experiments”, *Experimental Economics* 10(2).

[18] Frankel, D. M., S. Morris, and A. Pauzner, 2003, “Equilibrium Selection in Global Games with Strategic Complementarities” *Journal of Economic Theory* 108(1).

[19] Harsanyi, J. and R. Selten, 1988. “A General Theory of Equilibrium Selection in Games,” MIT Press Books, The MIT Press.

[20] Haruvy, E., and D. O. Stahl, 2007, ”Equilibrium selection and bounded rationality in symmetric normal-form games” *Journal of Economic Behavior & Organization*, 62(1): 98-119.

[21] Heinemann, F., R. Nagel and P. Ockenfels, 2004, “The Theory of Global Games on Test: Experimental Analysis of Coordination Games with Public and Private Information” *Econometrica* 72(5).

[22] Heinemann, F., R. Nagel and P. Ockenfels, 2009, “Measuring strategic uncertainty in coordination games” *Review of Economic Studies* 76(1).

[23] Hellwig, C., 2002, “Public information, private information, and the multiplicity of equilibria in coordination games” *Journal of Economic Theory*, 107 (2).

[24] Hellwig, C., A. Mukherji and A. Tsyvinski, 2006, “Self-fulfilling currency crises: The role of interest rates” *American Economic Review* 96(5).

[25] Hellwig, C. and L. Veldkamp, 2009, “Knowing what others know: Coordination motives in information acquisition” *Review of Economic Studies* 76(1).

[26] Kagel, J. and A. Roth, editors, “ The Handbook of Experimental Economics, Volumen 2”, forthcoming.

[27] Kneeland, T., 2014, “Coordination under limited depth of reasoning”, mimeo.

[28] Lehmann, E.L., 1988, “Comparing location experiments” *Annals of Statistics* 16(2).

[29] Milgrom, P. and J. Roberts, 1990, “Rationalizability, learning, and equilibrium in games with strategic complementarities”, *Econometrica*, 58(6).

[30] Morris, S. and H. Shin, 1998, "Unique equilibrium in a model of self-fulfilling currency attacks" *American Economic Review* 88(3).

[31] Morris, S. and H. Shin, 2002, "The social value of public information", *American Economic Review* 92(5).

[32] Morris S. and H. Shin, 2003, "Global games: Theory and applications" In Advances in Economics and Econometrics: Theory and applications, Eight World Congress, Cambridge University Press.

[33] Nagel, R., 1995, "Unraveling in Guessing Games: An Experimental Study." *American Economic Review*, 85(5).

[34] Nikitin, M. and T. Smith, 2008, "Information acquisition, coordination, and fundamentals in a financial crisis" *Journal of Banking and Finance* 32.

[35] Persico, N., 2000, "Information acquisition in auctions", *Econometrica* 68(1).

[36] Royden, H.L., 1988, *Real Analysis*, Prentice Hall, 3rd ed.

[37] Thaler, R., 1980, "Toward A Positive Theory of Consumer Choice" *Journal of Economic Behavior and Organization* 1.

[38] Szkup, M. and I. Trevino, 2015, "Costly information acquisition in Global Games of Regime Change" Mimeo.

[39] Topkis, D., 1998 "Supermodularity and Complementarity", Princeton, Princeton University Press.

[40] Van Huyck, J. B., R. C. Battalio, and R. O. Beil, 1990, "Tacit coordination games, strategic uncertainty, and coordination failure" *The American Economic Review,* 80: 234-248.

[41] Van Huyck, J. B., R. C. Battalio, and R. O. Beil, 1991, "Strategic uncertainty, equilibrium selection principles, and coordination failure in average opinion games" *Quarterly Journal of Economics,* 106: 885-911.

[42] van Zandt, T. and X. Vives, 2007, "Monotone equilibria in Bayesian games of strategic complementarities", *Journal of Economic Theory*, 134(1).

[43] van Zandt, T., 2010, "Interim Bayesian Nash equilibrium on universal type spaces for supermodular games", *Journal of Economic Theory,* 145(1).

[44] Varadhan, S., 2001, "Probability theory", Volume 7 of Courant Lecture Notes in Mathematics, New York University Courant Institute of Mathematical Sciences.

[45] Veldkamp, L., 2011, "Information Choice in Macroeconomics and Finance" Princeton University Press.

[46] Vives, X., 2005, "Complementarities and games: New developments" *Journal of Economic Literature*, 43(2).

[47] Yang, M., 2015, "Coordination with flexible information acquisition" *Journal of Economic Theory*, 158: 721-738..

[48] Zwart, S., 2008, "Liquidity runs with endogenous information acquisition", *Economics Letters*, 100.

A Appendix 1 - Tables and figures

	Complete info	High precision	Medium precision	Low precision
Logit (RE) (EPA)	22.01	27.61	40.16	35.79
	(7.15)	(5.86)	(9.13)	(9.00)
MET (EPA)	21.07	27.42	40.37	36.23
	(11.85)	(19.16)	(18.77)	(23.36)
MRT (EPS)		32.98	37.34	37.83
		(16.75)	(18.98)	(15.47)
Equilibrium x*		35.31	28.31	18.73
Risk dominant eq.		36	36	36

Table A.1: Estimated thresholds and equilibrium predictions for treatments without information acquisition

Variable	Attack {0,1}
Precision 1*signal	0.178***
	(0.03)
Precision 2*signal	0.191***
	(0.05)
Precision 3*signal	0.096***
	(0.01)
Precision 4*signal	0.088***
	(0.01)
Precision 5*signal	0.062***
	(0.02)
Precision 6*signal	0.057***
	(0.01)
Constant	-3.559***
	(0.46)
N	1000

Clustered (by subject) standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%

Table A.2: Attack as a function of precision, EPA treatment

Variable	Reported threshold
Precision 2	6.25
	(4.00)
Precision 3	10.65***
	(3.48)
Precision 4	10.78***
	(3.39)
Precision 5	12.87***
	(3.66)
Precision 6	12.54***
	(3.23)
Constant	20.66***
	(5.14)
N	1100

Clustered (by subject) standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%

Table A.3: Reported threshold as a function of precision, EPS treatment

		Choice of pair member 2					
		Prec 1	Prec 2	Prec 3	Prec 4	Prec 5	Prec 6
	Prec 1	**7.24%**	2.19%	5.71%	4.29%	1.24%	1.52%
Choice	Prec 2		**0.57%**	0.29%	3.52%	0.10%	0.19%
of pair	Prec 3			**5.43%**	14.00%	0.57%	5.33%
member	Prec 4				**20.57%**	1.71%	9.05%
1	Prec 5					**0.10%**	4.00%
	Prec 6						**12.38%**

Table A.4: Combination of precision choices, EPA and EPS treatments

Variable	Success of attack {0,1}
High precision*signal	0.194***
	(0.061)
Medium precision*signal	0.119***
	(0.022)
Low precision*signal	0.044***
	(0.01)
Constant	-4.534***
	(1.052)
N	500

Clustered (by pairs) standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%

Table A.5: Success of attack as a function of precision, EPA treatment

B Appendix 2 - Relation to monotone supermodular games

FOR ONLINE PUBLICATION

We first prove that the game specified in the second stage of the game belongs to the class of monotone supermodular games as defined by Vives and van Zandt (2007). Following their notation, define $N = \{1, 2\}$ as the set of players indexed by i. Let the type space of player i be a measurable space $(\Omega_i, \mathcal{F}_i)$. Denote by $(\Omega_0, \mathcal{F}_0)$ a state space that is capturing the residual uncertainty.[48] We let $\mathcal{F}_{-i}$ be the product $\sigma-$algebra $\otimes_{k \neq i} \mathcal{F}_k$. Let player i's interim beliefs be given by a function $p_i : \Omega_i \to M_{-i}$, where M_{-i} is the set of probability measures on $(\Omega_{-i}, \mathcal{F}_{-i})$. Finally, let $A_i = \{0, 1\}$ be the action set of player i, A be the set of action profiles and $u_i : A \times \Omega \to \mathbb{R}$ be the payoff function.

Definition 3 *A game belongs to the class of monotone supermodular games if*

1. *The utility function $u_i(a_i, a_{-i}, \omega)$ is supermodular in own actions, a_i, and has increasing differences in (a_i, a_{-i}) and in (a_i, ω).*
2. *The belief map $p_i : \Omega_i \to M_{-i}$ is increasing with respect to a partial order on M_{-i} of first-order stochastic dominance.*

In our case, the type space is defined as follows: $\Omega_0 = \mathbb{R}$, $\Omega_i = \mathbb{R}$ for $i = 1, 2$, where $\varpi_0 = \theta$, $\varpi_i = \widehat{\theta}_i$, $\varpi_j = \widehat{\theta}_j$ and $\mathcal{F}_i = B(\mathbb{R})$, a Borel σ-algebra on $\mathbb{R}$, $i = 0, 1, 2$. The set of probability measures M_{-i} is simply the set of joint normal probability distributions over $(\Omega_{-i}, \mathcal{F}_{-i})$ conditional on the realization of ϖ_i.The belief mapping $p_i : \Omega_i \to M_{-i}$ maps $\widehat{\theta}_i$ into the posterior distribution of $(\theta, \widehat{\theta}_j)$ using Bayes' rule. Finally, the underlying utility function for agent i is given by

$$u(a_i, a_j, \theta) = 1_{\{a_i=1\}} \left[\theta \left[1_{\{\theta \in [\underline{\theta}, \overline{\theta}]\}} 1_{\{a_j=1\}} + 1_{\{\theta > \overline{\theta}\}} \right] - T \right]$$

and the expected utility of agent i is:

$$v_i(a_i, a_j, \theta) = 1_{\{a_i=1\}} \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \theta \left[1_{\{\theta \in [\underline{\theta}, \overline{\theta}]\}} 1_{\{s_j(\widehat{\theta}_j)=1\}} + 1_{\{\theta > \overline{\theta}\}} \right] \frac{1}{\widehat{\sigma}_i} \frac{1}{\sigma_j} f\left(\frac{\theta - \widehat{\theta}_i}{\widehat{\sigma}_i} \right) f\left(\frac{\sigma_j \left(\widehat{\theta}_j - \widetilde{\theta}_j \right)}{\widehat{\sigma}_j^2} \right) d\widehat{\theta}_j d\theta \right] - T$$

where $s_j : \Omega_j \to A_j$ is a measurable strategy of player j.

The fact that global games belong to the class of monotone supermodular games was noted first by Vives (2005) and Vives and Van Zandt (2007). The following lemma shows that our game in the second stage satisfies the above definition of monotone supermodular games.

[48] In a global games setting, we usually interpret $(\Omega_i, \mathcal{F}_i)$ to be the space of possible signals that agent i receives, while $(\Omega_0, \mathcal{F}_0)$ corresponds to the measurable space of the underlying parameter of the game.

Lemma 3 *The game specified in the second stage of the speculative attack game with information acquisition belongs to the class of monotone supermodular games.*

Proof. Consider the ex-ante payoff function of agent i:

$$1_{\{a_i=1\}}\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\theta\left[1_{\{\theta\in[\underline{\theta},\overline{\theta}]\}}1_{\left\{\widehat{\theta}_j>\widehat{\theta}_j^*\right\}}+1_{\{\theta>\overline{\theta}\}}\right]\frac{1}{\widehat{\sigma}_i}\frac{1}{\sigma_j}f\left(\frac{\theta-\widehat{\theta}_i}{\widehat{\sigma}_i}\right)f\left(\frac{\sigma_j\left(\widehat{\theta}_j^*-\widetilde{\theta}_j\right)}{\widehat{\sigma}_j^2}\right)d\widehat{\theta}_jd\theta-T\right)$$

We see that the underlying utility function is simply

$$u(a_i,a_j,\theta)=1_{\{a_i=1\}}\left[\theta\left(1_{\{\theta\in[\underline{\theta},\overline{\theta}]\}}1_{\{a_j=1\}}+1_{\{\theta>\overline{\theta}\}}\right)-T\right]$$

By example 2.6.2 in Topkis (1998) we conclude that this function is supermodular.

To show that u has increasing differences in (a_i,a_j) we need to show that

$$u(1,0,\theta)-u(0,0,\theta)\leqslant u(1,1,\theta)-u(0,1,\theta)$$

But this follows immediately from the definition of u, namely

$$u(1,0,\theta)-u(0,0,\theta)=\theta 1_{\{\theta>\overline{\theta}\}}-T \text{ and } u(1,1,\theta)-u(0,1,\theta)=\theta 1_{\{\theta>\underline{\theta}\}}-T$$

so that

$$u(1,0,\theta)-u(0,0,\theta)\leqslant u(1,1,\theta)-u(0,1,\theta)$$

The fact that u has increasing differences in (a_i,θ) follows immediately from the above expressions. Increasing θ always weakly increases the difference between attacking and not attacking, regardless of action of the other player.

To show that the belief mapping is increasing with respect to first-order stochastic dominance it is enough to show that $(\theta,\widehat{\theta}_1,\widehat{\theta}_2)$ are affiliated. Let $\theta''>\theta',\widehat{\theta}_1''>\widehat{\theta}_1',\widehat{\theta}_2''>\widehat{\theta}_1'$ and denote by f the joint PDF of $(\theta,\widehat{\theta}_1,\widehat{\theta}_2)$. We need to show that f is log-supermodular (see Milgrom and Weber, 1982). Without loss of generality it is enough to show that

$$f(\theta'',\widehat{\theta}_1'',\widehat{\theta}_2'')f(\theta',\widehat{\theta}_1',\widehat{\theta}_2')\geqslant f(\theta',\widehat{\theta}_1',\widehat{\theta}_2'')f(\theta'',\widehat{\theta}_1'',\widehat{\theta}_2')$$

Note that

$$\begin{aligned}
f(\theta'',\widehat{\theta}_1'',\widehat{\theta}_2'')f(\theta',\widehat{\theta}_1',\widehat{\theta}_2') &= f(\widehat{\theta}_1''|\theta'')f(\widehat{\theta}_2''|\theta'')f(\theta'')f(\widehat{\theta}_1'|\theta')f(\widehat{\theta}_2'|\theta')f(\theta')\\
f(\theta',\widehat{\theta}_1',\widehat{\theta}_2'')f(\theta'',\widehat{\theta}_1'',\widehat{\theta}_2') &= f(\widehat{\theta}_1'|\theta')f(\widehat{\theta}_2''|\theta')f(\theta')f(\widehat{\theta}_1''|\theta'')f(\widehat{\theta}_2'|\theta'')f(\theta'')
\end{aligned}$$

since $\widehat{\theta}_1, \widehat{\theta}_2$ are conditionally independent. Hence, it is enough to show that

$$f(\widehat{\theta}_2''|\theta'')f(\widehat{\theta}_2'|\theta') \geqslant f(\widehat{\theta}_2''|\theta')f(\widehat{\theta}_2'|\theta'')$$

or that $f(\widehat{\theta}_2|\theta)$ has an increasing marginal likelihood ratio. Since $f(\widehat{\theta}_2|\theta) \sim N\left(\frac{\sigma_\theta^2\theta+\sigma_2^2\mu_\theta}{\sigma_\theta^2+\sigma_2^2}, \widehat{\sigma}_2^4\frac{1}{\sigma_2^2}\right)$, it follows that $f(\widehat{\theta}_2|\theta)$ has a monotone likelihood ratio and hence $f(\theta'', \widehat{\theta}_1'', \widehat{\theta}_2'')f(\theta', \widehat{\theta}_1', \widehat{\theta}_2') \geqslant f(\theta', \widehat{\theta}_1', \widehat{\theta}_2'')f(\theta'', \widehat{\theta}_1'', \widehat{\theta}_2')$. Therefore, $(\theta, \widehat{\theta}_1, \widehat{\theta}_2)$ are affiliated and the belief mapping is increasing with respect to first-order stochastic dominance. ■

In order to prove our main result in this section, we prove the following lemma.

Lemma 4 *The following are true about the utility function $u(\cdot,\cdot,\cdot)$:*

1. *u is bounded from below;*

2. *u is integrable with respect to μ_F, a Baire measure implied by F - the distribution function of θ;*

3. *There exists a function h, integrable w.r.t. μ_F, such that $|u| < h$.*

Proof. It is easy to see that u is bounded from below by $-T$. To prove (3) note that

$$\begin{aligned} \int |u|\, d\mu_F &= \int \left| 1_{\{a_i=1\}} \left[\theta\left(1_{\{\theta\in[\underline{\theta},\overline{\theta}]\}}1_{\{a_j=1\}} + 1_{\{\theta>\overline{\theta}\}}\right) - T\right]\right| d\mu_F \\ &\leqslant \int |\theta - T|\, d\mu_F \leqslant \int |\theta|\, d\mu_F + \int |-T|\, d\mu_F < \infty \end{aligned}$$

where $\int |\theta|\, d\mu_F < \infty$ since $\int \theta d\mu_F = \mu < \infty$. This shows that $|u|$ is integrable and since u is measurable it follows that u is also integrable (hence (2) is true).

Finally let $h(\theta) = |\theta| + |T|$. Then $|u| < h$ for all a_i, a_j and θ since

$$\begin{aligned} &\left| 1_{\{a_i=1\}} \left[\theta\left(1_{\{\theta\in[\underline{\theta},\overline{\theta}]\}}1_{\{a_j=1\}} + 1_{\{\theta>\overline{\theta}\}}\right) - T\right]\right| \\ \leq\ &\left| 1_{\{a_i=1\}} \left[\theta\left(1_{\{\theta\in[\underline{\theta},\overline{\theta}]\}}1_{\{a_j=1\}} + 1_{\{\theta>\overline{\theta}\}}\right)\right]\right| + |-T| \\ \leq\ &|\theta| + |T| \end{aligned}$$

We argued above that $\int |\theta|\, d\mu_F + \int |T|\, d\mu_F < \infty$ and hence (3) holds. ■

We proceed now by extending the following result from van Zandt and Vives (2007) for unbounded utility functions.

Proposition 1 *Assume that a game Γ belongs to the class of monotone supermodular games. Furthermore, assume that the following hold:*

1. *Each Ω_k is endowed with a partial order,*

2. A_i is a complete lattice,

3. $\forall a_i \in A_i$, $u_i(a_i, \cdot) : \Omega \to \mathbb{R}$ is measurable,

4. u_i is bounded.

5. u_i is continuous in a_i[49]

Then, there exist a least and a greatest Bayesian Nash Equilibrium of the game Γ and each one of them is in monotone strategies.

Proof. See van Zandt and Vives (2007). ■

Note that in our setup the underlying utility function $u(\cdot)$ is unbounded, namely as $\theta \to \infty$, $u(\theta) \to \infty$. Thus, we cannot apply the above proposition to our problem directly. In the following corollary, we show that, as long as conditions $(1) - (3)$ of Proposition 1 hold, u is bounded from below, and the distribution of θ satisfies assumptions $A1 - A5$, we can still reach the conclusion of the Proposition of Vives and van Zandt (2007) under some further assumptions. The strategy is to use the Dominated Convergence Theorem in the proof of existence of a greatest and a least Bayesian Nash Equilibria instead of the Bounded Convergence Theorem and to use the fact that u is bounded from below to show that the best reply mapping is well defined.

Corollary 1 *Assume that the game Γ belongs to the class of monotone supermodular games. Furthermore, assume that assumptions $(1)-(3)$ of Proposition 1 are satisfied, u is bounded from below, and let v_i satisfy the following assumption:*

(1C) There exists a measurable function h s.t. h is integrable with respect to $p(t_{-i}|t_i)$ for all t_i, and all t_{-i} and $|u| < h$.

Then there exists a least and a greatest Bayesian Nash Equilibrium of the game Γ and each one of them is in monotone strategies.

Proof. We prove this corollary in two steps. First, assuming that the greatest best reply mapping $\overline{\beta}_i$ is well-defined, increasing, and monotone, we show that the greatest Bayesian Nash Equilibrium (BNE) exists. Then, we show that under the above conditions $\overline{\beta}_i$ is indeed well-defined, increasing, and monotone.

Step 1: Suppose that $\overline{\beta}_i$, is well-defined, increasing and monotone and u satisfies assumption (1C). Then we can repeat the argument of van Zandt and Vives (2007) to show that there is a greatest and least BNE in monotone strategies. We can relax the boundedness assumption, since under assumption (1C) we can interchange the order of limit and integration due to the Lebesgue Dominated Convergence Theorem. Since this is the only step in that proof that requires boundedness of the utility function, we are done.

Step 2: Here we need to establish that $\overline{\beta}_i$ is well-defined and increasing. Then, the monotonicity of $\overline{\beta}_i$ will follow from Proposition 11 in van Zandt and Vives (2007). The tricky part of this step

[49] When A_i is finite this condition is vacuous.

is to show that $\overline{\beta}_i$ is well-defined, and more precisely that it is a measurable function of ϖ_i. For this purpose we extend the proof of Lemma 9 in Ely and Peski (2006) to cover general measurable functions. The rest of argument follows from van Zandt (2010).

Fix $a_i \in A_i$ and define $U_i(\varpi_i, \varpi_j) := u_i(a_i, s_j(\varpi_j), \varpi_i, \varpi_{-i})$. We need to show that a function $\pi_i : A_i \times \Omega_i \to \overline{\mathbb{R}}$ defined by

$$\pi_i(a_i, \varpi_i) = \int_{\Omega_{-i}} U_i(\varpi_i, \varpi_{-i}) dp(\varpi_{-i}|\varpi_i)$$

is measurable in t_i. To prove this we use a result by Ely and Peski (2006):

Lemma (Ely and Peski) *Let A and B be measurable sets and $g : A \times B \to [0, 1]$ be a jointly measurable map. If $m : A \to \Delta B$ (where ΔB denotes the set of probability measures defined on B) is measurable, then the map $L^g : A \to R$ defined as $L^g(a) = \int g(a, \cdot)dm(a)$ is measurable.*

Note however, that the proof of their lemma is essentially unchanged if we allow $g : A \times B \to \overline{\mathbb{R}}$, as long as g is measurable and bounded from below. In this case, there exists an increasing sequence of simple functions g_n such that $g_n \uparrow g$, so by the extended Monotone Convergence Theorem (Ash, 2000) we have $\int g_n dv \to \int g dv$ for a measure v defined on $A \times B$. Hence we conclude that $\pi_i : A_i \times \Omega_i \to \overline{\mathbb{R}}$ is a measurable function of ϖ_i. The rest of the proof follows directly from van Zandt (2010) section 7.5. Monotonicity of $\overline{\beta}_i$ follows from Proposition 11 in van Zandt and Vives (2007). ■

C Appendix 3 - Proofs

FOR ONLINE PUBLICATION

In order to facilitate notation when solving the model, we rewrite the condition for threshold strategies in terms of the posteriors that agents hold about the fundamental θ, as in Hellwig (2002). Note that this is straightforward since the posterior about θ held by agent i, $\widehat{\theta}_i$, is a linear, strictly increasing function of the signal he observes, x_i. Therefore, agent i will attack whenever his posterior belief about θ, given his signal realization of x_i, is higher than the posterior of θ that corresponds to agent i's optimal threshold:

$$a(x_i;\boldsymbol{\sigma}) = \begin{cases} 1 & \text{iff } \widehat{\theta}_i \geq \widehat{\theta}_i^*(\boldsymbol{\sigma}) \\ 0 & \text{iff } \widehat{\theta}_i < \widehat{\theta}_i^*(\boldsymbol{\sigma}) \end{cases}$$

where $\widehat{\theta}_i^* = \frac{\mu_\theta \sigma_i^2 + x_i^* \sigma_\theta^2}{\sigma_i^2 + \sigma_\theta^2}$. In order to write the condition of agent j in terms of his posterior belief, notice that

$$\frac{x_j^* - \theta}{\sigma_j} = \frac{\sigma_j \left(\widehat{\theta}_j^* - \widetilde{\theta}_j\right)}{\widehat{\sigma}_j^2}$$

where $\widetilde{\theta}_j = \frac{\sigma_\theta^2 \theta + \sigma_j^2 \mu_\theta}{\sigma_j^2 + \sigma_\theta^2}$.

The expected payoff of attacking for agent $i = 1, 2$, conditional on observing signal x_i and given that the other agent follows a threshold strategy with switching point $\widehat{\theta}_j^*$ is:

$$v_i(x_i, x_j^*; \sigma) = \frac{1}{\widehat{\sigma}_i} \int_{\underline{\theta}}^{\overline{\theta}} \theta f\left(\frac{\theta - \widehat{\theta}_i}{\widehat{\sigma}_i}\right) \left(1 - F\left(\frac{\sigma_j \left(\widehat{\theta}_j^* - \widetilde{\theta}_j\right)}{\widehat{\sigma}_j^2}\right)\right) d\theta + \frac{1}{\widehat{\sigma}_i} \int_{\overline{\theta}}^{\infty} \theta f\left(\frac{\theta - \widehat{\theta}_i}{\widehat{\sigma}_i}\right) d\theta - T \qquad (4)$$

This notation is used for all proofs in this appendix.

Lemma 1 *The payoff for agent i of attacking, $v_i(x_i, x_j^*; \sigma)$, is increasing in his own signal x_i, and decreasing in the other agent's threshold x_j^*, for $i, j = 1, 2$, $i \neq j$.*

Proof. (1)

Note that $\widehat{\theta}_i$ is an increasing function of x_i, i.e. $\frac{\partial \widehat{\theta}_i}{\partial x_i} > 0$. Thus, it is enough to show that the payoff of attacking is increasing with respect to the posterior mean of θ, $\widehat{\theta}_i$.

Taking a partial derivative of (4) wrt $\widehat{\theta}_i$ yields:

$$-\int_{\underline{\theta}}^{\overline{\theta}} \theta \frac{1}{\widehat{\sigma}_i^2} f'\left(\frac{\theta - \widehat{\theta}_i}{\widehat{\sigma}_i}\right) \left(1 - F\left(\frac{\sigma_j \left(\widehat{\theta}_j^* - \widetilde{\theta}_j\right)}{\widehat{\sigma}_j^2}\right)\right) d\theta \; - \int_{\overline{\theta}}^{\infty} \theta \frac{1}{\widehat{\sigma}_i^2} f'\left(\frac{\theta - \widehat{\theta}_i}{\widehat{\sigma}_i}\right) d\theta$$

Consider the second term and apply integration by parts. Then,

$$
\begin{aligned}
-\int_{\bar{\theta}}^{\infty} \theta \frac{1}{\widehat{\sigma}_i^2} f'\left(\frac{\theta-\widehat{\theta}_i}{\widehat{\sigma}_i}\right) d\theta &= -\left[\theta \frac{1}{\widehat{\sigma}_i} f\left(\frac{\theta-\widehat{\theta}_i}{\widehat{\sigma}_i}\right)\right]_{\bar{\theta}}^{\infty} + \int_{\bar{\theta}}^{\infty} \frac{1}{\widehat{\sigma}_i} f\left(\frac{\theta-\widehat{\theta}_i}{\widehat{\sigma}_i}\right) d\theta \\
&= -\left[\theta \frac{1}{\widehat{\sigma}_i} f\left(\frac{\theta-\widehat{\theta}_i}{\widehat{\sigma}_i}\right)\right]_{\bar{\theta}}^{\infty} + \left(1-F\left(\frac{\bar{\theta}-\widehat{\theta}_i}{\widehat{\sigma}_i}\right)\right)
\end{aligned}
$$

Since $f(\cdot)$ is a normal density, this simplifies to:

$$
\bar{\theta}\frac{1}{\widehat{\sigma}_i} f\left(\frac{\bar{\theta}-\widehat{\theta}_i}{\widehat{\sigma}_i}\right) + \left(1-F\left(\frac{\bar{\theta}-\widehat{\theta}_i}{\widehat{\sigma}_i}\right)\right) > 0 \tag{5}
$$

Now consider the first term. Again apply integration by parts:

$$
\begin{aligned}
&-\int_{\underline{\theta}}^{\bar{\theta}} \theta \frac{1}{\widehat{\sigma}_i^2} f'\left(\frac{\theta-\widehat{\theta}_i}{\widehat{\sigma}_i}\right) \left(1-F\left(\frac{\sigma_j\left(\widehat{\theta}_j^*-\widetilde{\theta}_j\right)}{\widehat{\sigma}_j^2}\right)\right) d\theta = \\
&-\left[\theta \frac{1}{\widehat{\sigma}_i} f\left(\frac{\theta-\widehat{\theta}_i}{\widehat{\sigma}_i}\right) \left(1-F\left(\frac{\sigma_j\left(\widehat{\theta}_j^*-\widetilde{\theta}_j\right)}{\widehat{\sigma}_j^2}\right)\right)\right]_{\underline{\theta}}^{\bar{\theta}} \\
&+\int_{\underline{\theta}}^{\bar{\theta}} \frac{1}{\widehat{\sigma}_i} f\left(\frac{\theta-\widehat{\theta}_i}{\widehat{\sigma}_i}\right) \left(1-F\left(\frac{\sigma_j\left(\widehat{\theta}_j^*-\widetilde{\theta}_j\right)}{\widehat{\sigma}_j^2}\right)\right) d\theta + \int_{\underline{\theta}}^{\bar{\theta}} \theta \frac{1}{\widehat{\sigma}_i \sigma_j} f\left(\frac{\theta-\widehat{\theta}_i}{\widehat{\sigma}_i}\right) f\left(\frac{\sigma_j\left(\widehat{\theta}_j^*-\widetilde{\theta}_j\right)}{\widehat{\sigma}_j^2}\right) d\theta
\end{aligned} \tag{6}
$$

Note that the first term of the above expression is of unknown sign while the second and third terms are unambiguously positive.

Putting (5) and (6) together we get:

$$
\begin{aligned}
&\bar{\theta}\frac{1}{\widehat{\sigma}_i} f\left(\frac{\bar{\theta}-\widehat{\theta}_i}{\widehat{\sigma}_i}\right) + \left(1-F\left(\frac{\bar{\theta}-\widehat{\theta}_i}{\widehat{\sigma}_i}\right)\right) - \bar{\theta}\frac{1}{\widehat{\sigma}_i} f\left(\frac{\bar{\theta}-\widehat{\theta}_i}{\widehat{\sigma}_i}\right) \left(1-F\left(\frac{\sigma_j\left(\widehat{\theta}_j^*-\widetilde{\bar{\theta}}_j\right)}{\widehat{\sigma}_j^2}\right)\right) \\
&+\underline{\theta}\frac{1}{\widehat{\sigma}_i} f\left(\frac{\underline{\theta}-\widehat{\theta}_i}{\widehat{\sigma}_i}\right) \left(1-F\left(\frac{\sigma_j\left(\widehat{\theta}_j^*-\widetilde{\underline{\theta}}_j\right)}{\widehat{\sigma}_j^2}\right)\right) + \int_{\underline{\theta}}^{\bar{\theta}} \frac{1}{\widehat{\sigma}_i} f\left(\frac{\theta-\widehat{\theta}_i}{\widehat{\sigma}_i}\right) \left(1-F\left(\frac{\sigma_j\left(\widehat{\theta}_j^*-\widetilde{\theta}_j\right)}{\widehat{\sigma}_j^2}\right)\right) d\theta \\
&+\int_{\underline{\theta}}^{\bar{\theta}} \theta \frac{1}{\widehat{\sigma}_i \sigma_j} f\left(\frac{\theta-\widehat{\theta}_i}{\widehat{\sigma}_i}\right) f\left(\frac{\sigma_j\left(\widehat{\theta}_j^*-\widetilde{\theta}_j\right)}{\widehat{\sigma}_j^2}\right) d\theta > 0
\end{aligned} \tag{7}
$$

since $\bar{\theta}\frac{1}{\widehat{\sigma}_i}f\left(\frac{\bar{\theta}-\widehat{\theta}_i}{\widehat{\sigma}_i}\right) \geqslant \bar{\theta}\frac{1}{\widehat{\sigma}_i}f\left(\frac{\bar{\theta}-\widehat{\theta}_i}{\widehat{\sigma}_i}\right)\left(1-F\left(\frac{\sigma_j\left(\widehat{\theta}_j^*-\widetilde{\theta}_j\right)}{\widehat{\sigma}_j^2}\right)\right)$. Since $\widehat{\theta}_i$ is an increasing linear function of x_i, this proves the first claim.

(2) Similarly, we note that the posterior $\widehat{\theta}_j^*$ is strictly increasing in x_j^*. Hence, it is enough to show that the derivative of the payoff of attacking for agent i with respect to $\widehat{\theta}_j^*$ is negative.

Take a partial derivative of (4) wrt $\widehat{\theta}_j^*$, which is always negative:

$$-\int_{\underline{\theta}}^{\bar{\theta}} \theta \frac{\sigma_j}{\widehat{\sigma}_i\widehat{\sigma}_j^2} f\left(\frac{\theta-\widehat{\theta}_i}{\widehat{\sigma}_i}\right) f\left(\frac{\sigma_j\left(\widehat{\theta}_j^*-\widetilde{\theta}_j\right)}{\widehat{\sigma}_j^2}\right) d\theta < 0$$

Therefore, it holds true for x_j^* as well. ■

Theorem 1 *There exists a unique, dominance solvable equilibrium of the second stage of the game in which both players use threshold strategies characterized by (x_1^*, x_2^*) if either:*

1. $\frac{\sigma_i}{\sigma_\theta} < K_i(\underline{\theta}, \overline{\theta}, \mu_\theta)$, $i = 1, 2$ *holds, for any pair of* (σ_1, σ_2)*, or*

2. $\sigma_\theta > \overline{\sigma}_\theta$*, where* $\overline{\sigma}_\theta$ *is determined by the parameters of the model.*

Proof. As proven in the first section of the appendix, the coordination game belongs to the class of monotone supermodular games and therefore we know that there are a least and a greatest Bayesian Nash Equilibria in monotone strategies. To prove the theorem, we only need to show that these equilibria are the same, i.e. that there is a unique equilibrium in threshold strategies.

For ease of exposition, we will perform the analysis in terms of thresholds over posterior beliefs, $(\widehat{\theta}_1^*, \widehat{\theta}_2^*)$. Uniqueness of these thresholds imply uniqueness of thresholds over signals (x_1^*, x_2^*).

Let, $s_i(\widehat{\theta}_i^*)$ be a threshold strategy of player i with switching point $\widehat{\theta}_i^*$ such that $s_i(\widehat{\theta}_i^*) = 1$ (attack) if $\widehat{\theta}_i \geq \widehat{\theta}_i^*$ and $s_i(\widehat{\theta}_i^*) = 0$ (not attack) if $\widehat{\theta}_i < \widehat{\theta}_i^*$, where $\widehat{\theta}_i$ is the posterior belief that agent i holds about θ after observing signal x_i, for $i = 1, 2$. Then the equilibrium conditions are given by the following equations:

$$v_1(\widehat{\theta}_1^*, \widehat{\theta}_2^*; \boldsymbol{\sigma}) \equiv \frac{1}{\widehat{\sigma}_1}\int_{\underline{\theta}}^{\overline{\theta}} \theta f\left(\frac{\theta-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right)\left(1-F\left(\frac{\sigma_2\left(\widehat{\theta}_2^*-\widetilde{\theta}_2\right)}{\widehat{\sigma}_2^2}\right)\right)d\theta + \frac{1}{\widehat{\sigma}_1}\int_{\overline{\theta}}^{\infty}\theta f\left(\frac{\theta-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right)d\theta - T = 0 \tag{8}$$

$$v_2(\widehat{\theta}_2^*, \widehat{\theta}_1^*; \boldsymbol{\sigma}) \equiv \frac{1}{\widehat{\sigma}_2}\int_{\underline{\theta}}^{\overline{\theta}} \theta f\left(\frac{\theta-\widehat{\theta}_2^*}{\widehat{\sigma}_2}\right)\left(1-F\left(\frac{\sigma_1\left(\widehat{\theta}_1^*-\widetilde{\theta}_1\right)}{\widehat{\sigma}_1^2}\right)\right)d\theta + \frac{1}{\widehat{\sigma}_2}\int_{\overline{\theta}}^{\infty}\theta f\left(\frac{\theta-\widehat{\theta}_2^*}{\widehat{\sigma}_2}\right)d\theta - T = 0 \tag{9}$$

where $\widehat{\theta}_i^* = \frac{\sigma_i^2\mu_\theta+\sigma_\theta^2 x_i^*}{\sigma_i^2+\sigma_\theta^2}$, $\widehat{\sigma}_i = \sqrt{\frac{\sigma_i^2\sigma_\theta^2}{\sigma_i^2+\sigma_\theta^2}}$ and $\widetilde{\theta}_i = \frac{\sigma_i^2\mu_\theta+\sigma_\theta^2\theta}{\sigma_i^2+\sigma_\theta^2}$for $i = 1, 2$. Note that both equations determine $\widehat{\theta}_j^*$ in terms of $\widehat{\theta}_i^*$. Without loss of generality we analyze the behavior of $\widehat{\theta}_2^*$ as a function of $\widehat{\theta}_1^*$ in the $(\widehat{\theta}_1^*, \widehat{\theta}_2^*)$ space and rewrite equations (8) and (9) as:

$$\begin{aligned} v_1(\widehat{\theta}_1^*, w_1(\widehat{\theta}_1^*; \boldsymbol{\sigma}); \boldsymbol{\sigma}) &= 0 \quad (10) \\ v_2(w_2(\widehat{\theta}_1^*; \boldsymbol{\sigma}), \widehat{\theta}_1^*; \boldsymbol{\sigma}) &= 0 \quad (11) \end{aligned}$$

where $\widehat{\theta}_2^* = w_i(\widehat{\theta}_1^*; \sigma)$ for $\widehat{\theta}_2^*$ as defined by the equation that characterized agent i's payoff function, for $i = 1, 2$. Then any $\widehat{\theta}_1^*$ that solves simultaneously both equations defines an equilibrium threshold for player 1 and the associated threshold for player 2 is simply given by $\widehat{\theta}_2^* = w_1(\widehat{\theta}_1^*; \boldsymbol{\sigma})$.

Consider first equation (10). Define $\overline{\theta}_1^*$ as the unique solution to the following equation:

$$\int_{\overline{\theta}}^{\infty} \theta \frac{1}{\widehat{\sigma}_1} f\left(\frac{\theta - \overline{\theta}_1^*}{\widehat{\sigma}_1}\right) d\theta - T = 0$$

Similarly, denote by $\underline{\theta}_1^*$ the unique solution to the following equation:

$$\int_{\underline{\theta}}^{\infty} \theta \frac{1}{\widehat{\sigma}_1} f\left(\frac{\theta - \underline{\theta}_1^*}{\widehat{\sigma}_1}\right) d\theta - T = 0$$

The first of the above conditions corresponds to the situation when player 2 never attacks while the second condition corresponds to the situation where player 2 always chooses to take the risky action. Note that $-\infty < \underline{\theta}_1^* < \overline{\theta}_1^* < \infty$ and therefore it follows that $\widehat{\theta}_1^*$ is finite (and $\widehat{\theta}_1^* \in \left[\underline{\theta}_1^*,\ \overline{\theta}_1^*\right]$). Recall that by lemma 1, the LHS of (10) is increasing in $\widehat{\theta}_1^*$ and decreasing in $\widehat{\theta}_2^*$. It follows then that as $\widehat{\theta}_1^* \to \underline{\theta}_1^*$, $\widehat{\theta}_2^* \to -\infty$ and as $\widehat{\theta}_1^* \to \overline{\theta}_1^*$, $\widehat{\theta}_2^* \to \infty$. Therefore $w_1(\widehat{\theta}_1^*; \sigma)$ has asymptotes at $\underline{\theta}_1^*$ and $\overline{\theta}_1^*$. Similarly define $\underline{\theta}_2^*$ and $\overline{\theta}_2^*$ for agent two. By lemma 1 we conclude that $w_2(\widehat{\theta}_1^*; \sigma)$ is bounded above by $\overline{\theta}_2^*$ and below by $\underline{\theta}_2^*$. Finally, let $\theta_{\min}^* = \min\{\underline{\theta}_1^*, \underline{\theta}_2^*\}$ and $\theta_{\max}^* = \max\left\{\overline{\theta}_1^*, \overline{\theta}_2^*\right\}$ so that $\theta_{\min}^*$ is the smallest and $\theta_{\max}^*$ is the largest threshold that can be rationalized.

Using the implicit function theorem we can find the derivative of $w_1(\widehat{\theta}_1^*; \sigma)$ w.r.t. $\widehat{\theta}_1^*$:

$$\frac{dw_1(\widehat{\theta}_1^*)}{d\widehat{\theta}_1^*} = \frac{\int_{\underline{\theta}}^{\overline{\theta}} \theta \frac{1}{\widehat{\sigma}_1} f\left(\frac{\theta - \widehat{\theta}_1^*}{\widehat{\sigma}_1}\right) \frac{1}{\sigma_2} f\left(\frac{\sigma_2(\widehat{\theta}_2^* - \widetilde{\theta}_2)}{\widehat{\sigma}_2^2}\right) d\theta + \tilde{V}_1}{\frac{\sigma_\theta^2 + \sigma_2^2}{\sigma_\theta^2} \int_{\underline{\theta}}^{\overline{\theta}} \theta \frac{1}{\widehat{\sigma}_1} f\left(\frac{\theta - \widehat{\theta}_1^*}{\widehat{\sigma}_1}\right) \frac{1}{\sigma_2} f\left(\frac{\sigma_2(\widehat{\theta}_2^* - \widetilde{\theta}_2)}{\widehat{\sigma}_2^2}\right) d\theta} > 0$$

where

$$\begin{aligned}
\tilde{V}_1 &= 1-F\left(\frac{\bar{\theta}-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right)+\bar{\theta}\frac{1}{\widehat{\sigma}_1}f\left(\frac{\bar{\theta}-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right)F\left(\frac{\sigma_2\left(\widehat{\theta}_2^*-\frac{\sigma_\theta^2\bar{\theta}+\sigma_2^2\mu_\theta}{\sigma_\theta^2+\sigma_2^2}\right)}{\widehat{\sigma}_2^2}\right) \\
&\quad +\underline{\theta}\frac{1}{\widehat{\sigma}_1}f\left(\frac{\underline{\theta}-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right)\left(1-F\left(\frac{\sigma_2\left(\widehat{\theta}_2^*-\frac{\sigma_\theta^2\underline{\theta}+\sigma_2^2\mu_\theta}{\sigma_\theta^2+\sigma_2^2}\right)}{\widehat{\sigma}_2^2}\right)\right) \\
&\quad +\int_{\underline{\theta}}^{\overline{\theta}}\frac{1}{\widehat{\sigma}_1}f\left(\frac{\theta-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right)\left(1-F\left(\frac{\sigma_2(\widehat{\theta}_2^*-\widetilde{\theta}_2)}{\widehat{\sigma}_2^2}\right)\right)d\theta>0
\end{aligned}$$

is strictly positive (since both $\widehat{\theta}_1^*$ and $\widehat{\theta}_2^*$ are finite - see the discussion above).

Similarly, we calculate the derivative of $w_2(\widehat{\theta}_1^*;\sigma)$ w.r.t. $\widehat{\theta}_1^*$:

$$\frac{dw_2(\widehat{\theta}_1^*)}{d\widehat{\theta}_1^*}=\frac{\frac{\sigma_\theta^2+\sigma_1^2}{\sigma_\theta^2}\int_{\underline{\theta}}^{\overline{\theta}}\theta\frac{1}{\widehat{\sigma}_2}f\left(\frac{\theta-\widehat{\theta}_2^*}{\widehat{\sigma}_2}\right)\frac{1}{\sigma_1}f\left(\frac{\sigma_1(\widehat{\theta}_1^*-\tilde{\theta}_1)}{\hat{\sigma}_1^2}\right)d\theta}{\int_{\underline{\theta}}^{\overline{\theta}}\theta\frac{1}{\widehat{\sigma}_2}f\left(\frac{\theta-\widehat{\theta}_2^*}{\widehat{\sigma}_2}\right)\frac{1}{\sigma_1}f\left(\frac{\sigma_1(\widehat{\theta}_1^*-\tilde{\theta}_1)}{\hat{\sigma}_1^2}\right)d\theta+\tilde{V}_2}>0$$

where $\tilde{V}_2$ a strictly positive constant and is defined in analogously to $\tilde{V}_1$.

Note that a sufficient condition for uniqueness is

$$\frac{dw_1(\widehat{\theta}_1^*)}{d\widehat{\theta}_1^*}>\frac{dw_2(\widehat{\theta}_1^*)}{d\widehat{\theta}_1^*}>0$$

This translates in the following inequality:

$$\frac{\int_{\underline{\theta}}^{\overline{\theta}}\theta\frac{1}{\hat{\sigma}_1}f\left(\frac{\theta-\widehat{\theta}_1^*}{\hat{\sigma}_1}\right)\frac{1}{\sigma_2}f\left(\frac{\sigma_2(\widehat{\theta}_2^*-\tilde{\theta}_2)}{\hat{\sigma}_2^2}\right)d\theta+\tilde{V}_1}{\frac{\sigma_\theta^2+\sigma_2^2}{\sigma_\theta^2}\int_{\underline{\theta}}^{\overline{\theta}}\theta\frac{1}{\hat{\sigma}_1}f\left(\frac{\theta-\widehat{\theta}_1^*}{\hat{\sigma}_1}\right)\frac{1}{\sigma_2}f\left(\frac{\sigma_2(\widehat{\theta}_2^*-\tilde{\theta}_2)}{\hat{\sigma}_2^2}\right)d\theta}>\frac{\frac{\sigma_\theta^2+\sigma_1^2}{\sigma_\theta^2}\int_{\underline{\theta}}^{\overline{\theta}}\theta\frac{1}{\hat{\sigma}_2}f\left(\frac{\theta-\widehat{\theta}_2^*}{\hat{\sigma}_2}\right)\frac{1}{\sigma_1}f\left(\frac{\sigma_1(\widehat{\theta}_1^*-\tilde{\theta}_1)}{\hat{\sigma}_1^2}\right)d\theta}{\int_{\underline{\theta}}^{\overline{\theta}}\theta\frac{1}{\hat{\sigma}_2}f\left(\frac{\theta-\widehat{\theta}_2^*}{\hat{\sigma}_2}\right)\frac{1}{\sigma_1}f\left(\frac{\sigma_1(\widehat{\theta}_1^*-\tilde{\theta}_1)}{\hat{\sigma}_1^2}\right)d\theta+\tilde{V}_2}$$

Doing some algebraic manipulations, we get that the expression above is equivalent to

$$\begin{aligned}
&\frac{\tilde{V}_1}{\int_{\underline{\theta}}^{\overline{\theta}}\theta\frac{1}{\hat{\sigma}_1}f\left(\frac{\theta-\widehat{\theta}_1^*}{\hat{\sigma}_1}\right)\frac{1}{\sigma_2}f\left(\frac{\sigma_2(\widehat{\theta}_2^*-\tilde{\theta}_2)}{\hat{\sigma}_2^2}\right)d\theta}+ \\
&\frac{\tilde{V}_2}{\int_{\underline{\theta}}^{\overline{\theta}}\theta\frac{1}{\hat{\sigma}_2}f\left(\frac{\theta-\widehat{\theta}_2^*}{\hat{\sigma}_2}\right)\frac{1}{\sigma_1}f\left(\frac{\sigma_1(\widehat{\theta}_1^*-\tilde{\theta}_1)}{\hat{\sigma}_1^2}\right)d\theta}+ \\
&\frac{\tilde{V}_1\tilde{V}_2}{\left(\int_{\underline{\theta}}^{\overline{\theta}}\theta\frac{1}{\hat{\sigma}_1}f\left(\frac{\theta-\widehat{\theta}_1^*}{\hat{\sigma}_1}\right)\frac{1}{\sigma_2}f\left(\frac{\sigma_2(\widehat{\theta}_2^*-\tilde{\theta}_2)}{\hat{\sigma}_2^2}\right)d\theta\right)\left(\int_{\underline{\theta}}^{\overline{\theta}}\theta\frac{1}{\hat{\sigma}_2}f\left(\frac{\theta-\widehat{\theta}_2^*}{\hat{\sigma}_2}\right)\frac{1}{\sigma_1}f\left(\frac{\sigma_1(\widehat{\theta}_1^*-\tilde{\theta}_1)}{\hat{\sigma}_1^2}\right)d\theta\right)} \\
>\ &\frac{\sigma_1^2}{\sigma_\theta^2}+\frac{\sigma_2^2}{\sigma_\theta^2}+\frac{\sigma_1^2\sigma_2^2}{\sigma_\theta^4}
\end{aligned}$$

A sufficient condition for this inequality to hold is to have:

$$\frac{\sigma_1^2}{\sigma_\theta^2} < \frac{\tilde{V}_1}{\int_{\underline{\theta}}^{\overline{\theta}} \theta \frac{1}{\widehat{\sigma}_1} f\left(\frac{\theta-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right) \frac{1}{\sigma_2} f\left(\frac{\sigma_2(\widehat{\theta}_2^*-\widetilde{\theta}_2)}{\widehat{\sigma}_2^2}\right) d\theta}$$

and

$$\frac{\sigma_2^2}{\sigma_\theta^2} < \frac{\tilde{V}_2}{\int_{\underline{\theta}}^{\overline{\theta}} \theta \frac{1}{\widehat{\sigma}_2} f\left(\frac{\theta-\widehat{\theta}_2^*}{\widehat{\sigma}_2}\right) \frac{1}{\sigma_1} f\left(\frac{\sigma_1(\widehat{\theta}_1^*-\widetilde{\theta}_1)}{\widehat{\sigma}_1^2}\right) d\theta}$$

Take the first expression, for agent 1 (the result is analogous for agent 2). We want to find a lower bound for the RHS, i.e. a lower bound for the numerator of the RHS and an upper bound for the denominator of the RHS.

We first look at the denominator $\int_{\underline{\theta}}^{\overline{\theta}} \theta \frac{1}{\widehat{\sigma}_1} f\left(\frac{\theta-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right) \frac{1}{\sigma_2} f\left(\frac{\sigma_2(\widehat{\theta}_2^*-\widetilde{\theta}_2)}{\widehat{\sigma}_2^2}\right) d\theta$. By doing some algebraic manipulations and with the help of the properties of a normal distribution, we can rewrite this expression as

$$\begin{aligned}
&\int_{\underline{\theta}}^{\overline{\theta}} \theta \frac{1}{\widehat{\sigma}_1} f\left(\frac{\theta-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right) \frac{1}{\sigma_2} f\left(\frac{\sigma_2(\widehat{\theta}_2^*-\widetilde{\theta}_2)}{\widehat{\sigma}_2^2}\right) d\theta \\
= \quad & \frac{1}{\sqrt{\widehat{\sigma}_1^2+\sigma_2^2}} f\left(\frac{\widehat{\theta}_1^*-\Omega}{\sqrt{\widehat{\sigma}_1^2+\sigma_2^2}}\right) \int_{\underline{\theta}}^{\overline{\theta}} \theta \frac{1}{\sqrt{\frac{\widehat{\sigma}_1^2\sigma_2^2}{\widehat{\sigma}_1^2+\sigma_2^2}}} f\left(\frac{\theta-\frac{\widehat{\sigma}_1^2\Omega+\sigma_2^2\widehat{\theta}_1^*}{\widehat{\sigma}_1^2+\sigma_2^2}}{\sqrt{\frac{\widehat{\sigma}_1^2\sigma_2^2}{\widehat{\sigma}_1^2+\sigma_2^2}}}\right) d\theta \\
\leq \quad & \frac{1}{\sqrt{\widehat{\sigma}_1^2+\sigma_2^2}} \frac{1}{\sqrt{2\pi}} \left[\frac{\theta_{\max}^*\left(\widehat{\sigma}_1^2+\sigma_2^2+\widehat{\sigma}_1^2\frac{\sigma_2^2}{\sigma_\theta^2}\right)-\widehat{\sigma}_1^2\frac{\sigma_2^2}{\sigma_\theta^2}\mu_\theta}{\widehat{\sigma}_1^2+\sigma_2^2} + \frac{1}{\sqrt{2\pi}} \frac{\sqrt{\widehat{\sigma}_1^2\sigma_2^2}}{\widehat{\sigma}_1^2+\sigma_2^2}\right]
\end{aligned}$$

where $\Omega \equiv \frac{(\sigma_\theta^2+\sigma_2^2)}{\sigma_\theta^2}\widehat{\theta}_2^* - \frac{\sigma_2^2}{\sigma_\theta^2}\mu_\theta$.

We now look at the numerator $\widetilde{V}_1$. Note that

$$\begin{aligned}
&\int_{\underline{\theta}}^{\overline{\theta}} \frac{1}{\widehat{\sigma}_1} f\left(\frac{\theta-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right)\left(1-F\left(\frac{\sigma_2(\widehat{\theta}_2^*-\widetilde{\theta}_2)}{\widehat{\sigma}_2^2}\right)\right) d\theta \\
\geq \quad & \left[F\left(\frac{\overline{\theta}-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right) - F\left(\frac{\underline{\theta}-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right)\right]\left(1-F\left(\frac{\sigma_2(\widehat{\theta}_2^*-\underline{\widetilde{\theta}}_2)}{\widehat{\sigma}_2^2}\right)\right)
\end{aligned}$$

where $\underline{\widetilde{\theta}}_2 \equiv \frac{\sigma_\theta^2\overline{\theta}+\sigma_2^2\mu_\theta}{\sigma_\theta^2+\sigma_2^2}$ and therefore

$$\begin{aligned}
\tilde{V}_1 \quad > \quad & 1-F\left(\frac{\underline{\theta}-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right) - \left[F\left(\frac{\overline{\theta}-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right) - F\left(\frac{\underline{\theta}-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right)\right] F\left(\frac{\sigma_2(\widehat{\theta}_2^*-\underline{\widetilde{\theta}}_2)}{\widehat{\sigma}_2^2}\right) \\
& +\overline{\theta}\frac{1}{\widehat{\sigma}_1} f\left(\frac{\overline{\theta}-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right) F\left(\frac{\sigma_2\left(\widehat{\theta}_2^*-\frac{\sigma_\theta^2\overline{\theta}+\sigma_2^2\mu_\theta}{\sigma_\theta^2+\sigma_2^2}\right)}{\widehat{\sigma}_2^2}\right) + \underline{\theta}\frac{1}{\widehat{\sigma}_1} f\left(\frac{\underline{\theta}-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right)\left(1-F\left(\frac{\sigma_2\left(\widehat{\theta}_2^*-\frac{\sigma_\theta^2\underline{\theta}+\sigma_2^2\mu_\theta}{\sigma_\theta^2+\sigma_2^2}\right)}{\widehat{\sigma}_2^2}\right)\right)
\end{aligned}$$

We can bound the last two terms of the RHS of the above expression by

$$\frac{1}{\widehat{\sigma}_1} f\left(\frac{\bar{\theta}-\theta^*_{\min}}{\widehat{\sigma}_1}\right)\left[\overline{\theta} F\left(\frac{\sigma_2\left(\theta^*_{\min}-\frac{\sigma_\theta^2\bar{\theta}+\sigma_2^2\mu_\theta}{\sigma_\theta^2+\sigma_2^2}\right)}{\widehat{\sigma}_2^2}\right)+\underline{\theta}\left(1-F\left(\frac{\sigma_2\left(\theta^*_{\max}-\frac{\sigma_\theta^2\underline{\theta}+\sigma_2^2\mu_\theta}{\sigma_\theta^2+\sigma_2^2}\right)}{\widehat{\sigma}_2^2}\right)\right)\right]$$

Therefore, sufficient conditions for uniqueness are:

$$\frac{\sigma_2^2}{\sigma_\theta^2} < \frac{1-F\left(\frac{\underline{\theta}-\widehat{\theta}_{\min}}{\widehat{\sigma}_1}\right)-\left[F\left(\frac{\bar{\theta}-\widehat{\theta}_{\min}}{\widehat{\sigma}_1}\right)-F\left(\frac{\underline{\theta}-\widehat{\theta}_{\max}}{\widehat{\sigma}_1}\right)\right]F\left(\frac{\sigma_2(\widehat{\theta}_{\max}-\widetilde{\underline{\theta}}_2)}{\widehat{\sigma}_2^2}\right)+\frac{1}{\widehat{\sigma}_1}f\left(\frac{\bar{\theta}-\theta^*_{\min}}{\widehat{\sigma}_1}\right)\kappa_2}{\frac{1}{\sqrt{2\pi}}\left(\frac{1}{\sqrt{2\pi}}+\frac{\overline{\theta}\sigma_2^2\left(1+\frac{\sigma_2^2}{\widehat{\sigma}_1^2}\right)-\widehat{\sigma}_1^2\frac{\sigma_1^2}{\sigma_\theta^2}}{\widehat{\sigma}_1^2+\sigma_2^2}\right)}$$

$$\frac{\sigma_1^2}{\sigma_\theta^2} < \frac{1-F\left(\frac{\overline{\theta}-\underline{\theta}}{\widehat{\sigma}_2}\right)+\frac{1}{\widehat{\sigma}_2}f\left(\frac{\overline{\theta}-\underline{\theta}}{\widehat{\sigma}_2}\right)\kappa_1}{\frac{1}{\sqrt{2\pi}}\left(\frac{1}{\sqrt{2\pi}}+\frac{\overline{\theta}\sigma_1^2\left(1+\frac{\sigma_1^2}{\widehat{\sigma}_2^2}\right)-\widehat{\sigma}_2^2\frac{\sigma_2^2}{\sigma_\theta^2}}{\widehat{\sigma}_2^2+\sigma_1^2}\right)}$$

Where $\kappa_i := \overline{\theta} F\left(\frac{\sigma_2\left(\theta^*_{\min}-\frac{\sigma_\theta^2\bar{\theta}+\sigma_2^2\mu_\theta}{\sigma_\theta^2+\sigma_2^2}\right)}{\widehat{\sigma}_2^2}\right)+\underline{\theta}\left(1-F\left(\frac{\sigma_2\left(\theta^*_{\max}-\frac{\sigma_\theta^2\underline{\theta}+\sigma_2^2\mu_\theta}{\sigma_\theta^2+\sigma_2^2}\right)}{\widehat{\sigma}_2^2}\right)\right)$

If such conditions hold, $0 < \frac{dw_2(\widehat{\theta}_1^*;\boldsymbol{\sigma})}{d\widehat{\theta}_1^*} < \frac{dw_1(\widehat{\theta}_1^*;\boldsymbol{\sigma})}{d\widehat{\theta}_1^*}$ $\forall \widehat{\theta}_1^* \in [\underline{\theta}_1^*, \overline{\theta}_1^*]$. This means that the least and greatest Bayesian Nash equilibria of the game, as described by our Corollary in the first section of the appendix, coincide. Therefore, there is a unique equilibrium in thresholds strategies. This proves the first part of the theorem.

The proof for the second part of the theorem follows directly from the proof of the above result. Namely, recall that to prove uniqueness we have to find conditions under which the functions $w_1(\widehat{\theta}_1^*)$ and $w_2(\widehat{\theta}_1^*)$ (there were defined above) are such that

$$\frac{dw_1(\widehat{\theta}_1^*)}{d\widehat{\theta}_1^*} > \frac{dw_2(\widehat{\theta}_1^*)}{d\widehat{\theta}_1^*}$$

Note that as $\sigma_\theta \to \infty$ we have

$$\lim_{\sigma_\theta\to\infty}\frac{\sigma_\theta^2+\sigma_1^2}{\sigma_\theta^2} \to 1 \text{ and } \lim_{\sigma_\theta\to\infty}\frac{\sigma_\theta^2+\sigma_2^2}{\sigma_\theta^2} \to 1$$

Therefore

$$\frac{dw_1(\widehat{\theta}_1^*)}{d\widehat{\theta}_1^*} = \frac{\int_{\underline{\theta}}^{\overline{\theta}} \theta \frac{1}{\widehat{\sigma}_1} f\left(\frac{\theta-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right) \frac{1}{\sigma_2} f\left(\frac{\sigma_2(\widehat{\theta}_2^*-\widetilde{\theta}_2)}{\widehat{\sigma}_2^2}\right) d\theta + \tilde{V}_1}{\frac{\sigma_\theta^2+\sigma_2^2}{\sigma_\theta^2} \int_{\underline{\theta}}^{\overline{\theta}} \theta \frac{1}{\widehat{\sigma}_1} f\left(\frac{\theta-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right) \frac{1}{\sigma_2} f\left(\frac{\sigma_2(\widehat{\theta}_2^*-\widetilde{\theta}_2)}{\widehat{\sigma}_2^2}\right) d\theta}$$

$$\to \frac{\int_{\underline{\theta}}^{\overline{\theta}} \theta \frac{1}{\sigma_1} f\left(\frac{\theta-x_1^*}{\sigma_1}\right) \frac{1}{\sigma_2} f\left(\frac{x_2^*-\theta)}{\sigma_2}\right) d\theta + \tilde{V}_1}{\int_{\underline{\theta}}^{\overline{\theta}} \theta \frac{1}{\sigma_1} f\left(\frac{\theta-x_1^*}{\sigma_1}\right) \frac{1}{\sigma_2} f\left(\frac{x_2^*-\theta)}{\sigma_2}\right) d\theta} > 1$$

and

$$\frac{dw_2(\widehat{\theta}_1^*)}{d\widehat{\theta}_1^*} = \frac{\frac{\sigma_\theta^2+\sigma_1^2}{\sigma_\theta^2} \int_{\underline{\theta}}^{\overline{\theta}} \theta \frac{1}{\widehat{\sigma}_2} f\left(\frac{\theta-\widehat{\theta}_2^*}{\widehat{\sigma}_2}\right) \frac{1}{\sigma_1} f\left(\frac{\sigma_1(\widehat{\theta}_1^*-\widetilde{\theta}_1)}{\widehat{\sigma}_1^2}\right) d\theta}{\tilde{V}_2 + \int_{\underline{\theta}}^{\overline{\theta}} \theta \frac{1}{\widehat{\sigma}_2} f\left(\frac{\theta-\widehat{\theta}_2^*}{\widehat{\sigma}_2}\right) \frac{1}{\sigma_1} f\left(\frac{\sigma_1(\widehat{\theta}_1^*-\widetilde{\theta}_1)}{\widehat{\sigma}_1^2}\right) d\theta}$$

$$\to \frac{\int_{\underline{\theta}}^{\overline{\theta}} \theta \frac{1}{\sigma_2} f\left(\frac{\theta-x_2^*}{\sigma_2}\right) \frac{1}{\sigma_1} f\left(\frac{x_1^*-\theta}{\sigma_1}\right) d\theta}{\tilde{V}_2 + \int_{\underline{\theta}}^{\overline{\theta}} \theta \frac{1}{\sigma_2} f\left(\frac{\theta-x_2^*}{\sigma_2}\right) \frac{1}{\sigma_1} f\left(\frac{x_1^*-\theta}{\sigma_1}\right) d\theta} < 1$$

since, as we argued in the proof of the first part of the theorem, $\tilde{V}_1$ and $\tilde{V}_2$ are strictly positive. By continuity of the above expressions we conclude that for any σ_1 and σ_2 there exists $\overline{\sigma}_\theta(\sigma_1, \sigma_2)$ such that if $\sigma_\theta > \overline{\sigma}_\theta(\sigma_1, \sigma_2)$ we have a unique equilibrium in the coordination game.

The above bound depends on the information acquisition choices made by players. However, it is easy to show existence of a uniform bound. To do so, recall first that

$$\begin{aligned} \tilde{V}_1 = {} & 1 - F\left(\frac{\overline{\theta}-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right) + \overline{\theta}\frac{1}{\widehat{\sigma}_1} f\left(\frac{\overline{\theta}-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right) F\left(\frac{\sigma_2\left(\widehat{\theta}_2^* - \frac{\sigma_\theta^2\overline{\theta}+\sigma_2^2\mu_\theta}{\sigma_\theta^2+\sigma_2^2}\right)}{\widehat{\sigma}_2^2}\right) \\ & + \underline{\theta}\frac{1}{\widehat{\sigma}_1} f\left(\frac{\underline{\theta}-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right)\left(1 - F\left(\frac{\sigma_2\left(\widehat{\theta}_2^* - \frac{\sigma_\theta^2\underline{\theta}+\sigma_2^2\mu_\theta}{\sigma_\theta^2+\sigma_2^2}\right)}{\widehat{\sigma}_2^2}\right)\right)\Bigg) \\ & + \int_{\underline{\theta}}^{\overline{\theta}} \frac{1}{\widehat{\sigma}_1} f\left(\frac{\theta-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right)\left(1 - F\left(\frac{\sigma_2(\widehat{\theta}_2^*-\widetilde{\theta}_2)}{\widehat{\sigma}_2^2}\right)\right) d\theta > 0 \end{aligned}$$

and in particular

$$\tilde{V}_1 > 1 - F\left(\frac{\overline{\theta}-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right) > 0$$

Recall that $\widehat{\sigma}_1 = \sqrt{\frac{\sigma_1^2\sigma_\theta^2}{\sigma_1^2+\sigma_\theta^2}}$ and that $\sigma_1^2 \leq \sigma_0^2$ where σ_0^2 is the precision of a private signal if player 1 did not acquire any information. Let $\sigma_\theta^{2,B}$ be an arbitrary lower bound on σ_θ^2 such that $\sigma_\theta^{2,B} > 0$. If at $\sigma_\theta^2 = \sigma_\theta^{2,B}$ we have $\frac{dw_1(\widehat{\theta}_1^*)}{d\widehat{\theta}_1^*} > \frac{dw_2(\widehat{\theta}_1^*)}{d\widehat{\theta}_1^*}$ then we are done. Otherwise, we have to show that there exists a bound on σ_θ^2 that is higher than $\sigma_\theta^{2,B}$ for which $\frac{dw_1(\widehat{\theta}_1^*)}{d\widehat{\theta}_1^*} > \frac{dw_2(\widehat{\theta}_1^*)}{d\widehat{\theta}_1^*}$ independent of values of

σ_1 and σ_2.

To do so we start by finding uniform bounds on $\widehat{\theta}_1^*$. Note that, for any σ_1 and σ_θ, the lowest threshold that agent 1 can possible choose (which we denote by $\underline{\theta}_1^*(\sigma_\theta, \sigma_1)$) is determined by equation

$$\int_{\underline{\theta}}^{\infty} \frac{1}{\widehat{\sigma}_1^2} \theta f\left(\frac{\theta - \underline{\theta}_1^*}{\widehat{\sigma}_1}\right) d\theta = T$$

which corresponds to the situation in which the other agent always chooses to attack (and where we suppressed the dependence of $\underline{\theta}_1^*$ on$(\sigma_\theta, \sigma_1)$). This can be written as

$$\underline{\theta}_1^* \left(1 - F\left(\frac{\underline{\theta} - \underline{\theta}_1^*}{\widehat{\sigma}_1}\right)\right) + \widehat{\sigma}_1 f\left(\frac{\underline{\theta} - \underline{\theta}_1^*}{\widehat{\sigma}_1}\right) = T \tag{12}$$

By the implicit function theorem we have

$$\frac{\partial \underline{\theta}_1^*}{\partial \widehat{\sigma}_1} = -\frac{\underline{\theta}_1^* f\left(\frac{\underline{\theta} - \underline{\theta}_1^*}{\widehat{\sigma}_1}\right) \frac{\underline{\theta} - \underline{\theta}_1^*}{\widehat{\sigma}_1^2} + f\left(\frac{\underline{\theta} - \underline{\theta}_1^*}{\widehat{\sigma}_1}\right) + \frac{(\underline{\theta} - \underline{\theta}_1^*)^2}{\widehat{\sigma}_1^2} f\left(\frac{\underline{\theta} - \underline{\theta}_1^*}{\widehat{\sigma}_1}\right)}{\left(1 - F\left(\frac{\underline{\theta} - \underline{\theta}_1^*}{\widehat{\sigma}_1}\right)\right) + \frac{\underline{\theta}_1^*}{\widehat{\sigma}_1} f\left(\frac{\underline{\theta} - \underline{\theta}_1^*}{\widehat{\sigma}_1}\right) + \frac{\underline{\theta} - \underline{\theta}_1^*}{\widehat{\sigma}_1} f\left(\frac{\underline{\theta} - \underline{\theta}_1^*}{\widehat{\sigma}_1}\right)} < 0$$

So as $\widehat{\sigma}_1$ increases $\underline{\theta}_1^*$ decreases which implies that an increase in σ_θ^2 decreases $\underline{\theta}_1^*$ while an increase in σ_1 increases $\underline{\theta}_1^*$. In the same way we can show that the highest threshold that agent 1 can possible choose (denoted by $\overline{\theta}_1^*$) is also decreasing in $\widehat{\sigma}_1$. Therefore, $\underline{\theta}_1^*$ is minimized at $\sigma_1 = \sigma_0$ and $\sigma_\theta \to \infty$. This implies that

$$\tilde{V}_1 > 1 - F\left(\frac{\bar{\theta} - \widehat{\theta}_1^*}{\widehat{\sigma}_1}\right) > K_1$$

where

$$K_1 \equiv 1 - F\left(\frac{\bar{\theta} - \lim_{\sigma_\theta \to \infty} \underline{\theta}_1^*(\sigma_\theta, \sigma_0)}{\widehat{\sigma}_1}\right)$$

This establishes a bound on $\tilde{V}_1$ that is independent of σ_θ, σ_0. Note that by symmetry this is also a lower bound on $\tilde{V}_2$.

We now show that we can also bound other term appearing in the expression for the derivative uniformly. But

$$\int_{\underline{\theta}}^{\bar{\theta}} \theta \frac{1}{\widehat{\sigma}_2} f\left(\frac{\theta - \widehat{\theta}_2^*}{\widehat{\sigma}_2}\right) \frac{1}{\sigma_1} f\left(\frac{\sigma_1(\widehat{\theta}_1^* - \widetilde{\theta}_1)}{\widehat{\sigma}_1^2}\right) d\theta < \frac{1}{2\pi} \left[\bar{\theta} - \underline{\theta}\right] \frac{1}{\widehat{\sigma}_2} \frac{1}{\sigma_1}$$

Below, when we discuss the first stage of the game we show that the benefit from acquiring information tends to zero as $\sigma_i \to 0$ and therefore, given our assumptions on the cost function, i.e. $C'(\sigma_i) > 0$ and $\lim_{\sigma_i \to \infty} C'(\sigma_i) \to \infty$, there is a bound on the precision choice, call it $\sigma_i^{\min}$, such that agent i will never choose to acquire a lower standard deviation than $\sigma_i^{\min}$. Therefore,

$$\begin{aligned}\int_{\underline{\theta}}^{\overline{\theta}} \theta \frac{1}{\widehat{\sigma}_2} f\left(\frac{\theta-\widehat{\theta}_2^*}{\widehat{\sigma}_2}\right) \frac{1}{\sigma_1} f\left(\frac{\sigma_1(\widehat{\theta}_1^*-\widetilde{\theta}_1)}{\widehat{\sigma}_1^2}\right) &< \frac{1}{2\pi}\left[\overline{\theta}-\underline{\theta}\right]\frac{1}{\widehat{\sigma}_2}\frac{1}{\sigma_1} \\ &\leq \frac{1}{\widehat{\sigma}_2}\frac{1}{\sigma_1^{\min}}\frac{1}{2\pi}\left[\overline{\theta}-\underline{\theta}\right]\end{aligned}$$

Finally, note that $\widehat{\sigma}_2$ is increasing in σ_θ and σ_2 and so $\widehat{\sigma}_2$ is minimized at $\sigma_\theta = \sigma_\theta^B$ (our exogenous lower bound on σ_θ) and $\sigma_2 = \sigma_2^{\min}$ and denote by $\widehat{\sigma}_2^{\min}$ the posterior standard deviation of θ for player 2 when $\sigma_\theta = \sigma_\theta^B$ and $\sigma_2 = \sigma_2^{\min}$. Then

$$\int_{\underline{\theta}}^{\overline{\theta}} \theta \frac{1}{\widehat{\sigma}_2} f\left(\frac{\theta-\widehat{\theta}_2^*}{\widehat{\sigma}_2}\right) \frac{1}{\sigma_1} f\left(\frac{\sigma_1(\widehat{\theta}_1^*-\widetilde{\theta}_1)}{\widehat{\sigma}_1^2}\right) \leq K_2$$

where

$$K_2 \equiv \frac{1}{\widehat{\sigma}_2^{\min}}\frac{1}{\sigma_1^{\min}}\frac{1}{2\pi}\left[\overline{\theta}-\underline{\theta}\right]$$

Therefore, we have

$$\begin{aligned}\frac{dw_1(\widehat{\theta}_1^*)}{d\widehat{\theta}_1^*} &= \frac{\int_{\underline{\theta}}^{\overline{\theta}} \theta \frac{1}{\widehat{\sigma}_1} f\left(\frac{\theta-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right) \frac{1}{\sigma_2} f\left(\frac{\sigma_2(\widehat{\theta}_2^*-\widetilde{\theta}_2)}{\widehat{\sigma}_2^2}\right) d\theta + \tilde{V}_1}{\frac{\sigma_\theta^2+\sigma_2^2}{\sigma_\theta^2}\int_{\underline{\theta}}^{\overline{\theta}} \theta \frac{1}{\widehat{\sigma}_1} f\left(\frac{\theta-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right) \frac{1}{\sigma_2} f\left(\frac{\sigma_2(\widehat{\theta}_2^*-\widetilde{\theta}_2)}{\widehat{\sigma}_2^2}\right) d\theta} \\ &= \frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_2^2} + \frac{\tilde{V}_1}{\frac{\sigma_\theta^2+\sigma_2^2}{\sigma_\theta^2}\int_{\underline{\theta}}^{\overline{\theta}} \theta \frac{1}{\widehat{\sigma}_1} f\left(\frac{\theta-\widehat{\theta}_1^*}{\widehat{\sigma}_1}\right) \frac{1}{\sigma_2} f\left(\frac{\sigma_2(\widehat{\theta}_2^*-\widetilde{\theta}_2)}{\widehat{\sigma}_2^2}\right) d\theta} \\ &\geq \frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_2^2} + \frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_2^2}\frac{K_1}{K_2} \\ &\geq \frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_0^2} + \frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_0^2}\frac{K_1}{K_2}\end{aligned}$$

where we used the fact that $\sigma_2^2 \leq \sigma_0^2$. Note that as $\sigma_\theta^2 \to \infty$ we have $\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_0^2} + \frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_0^2}\frac{K_1}{K_2} \to 1 + \frac{K_1}{K_2} > 1$ and therefore there exists a bound on σ_θ^2, call it $\overline{\sigma}_\theta^{2,P1}$ such that if $\sigma_\theta^2 > \overline{\sigma}_\theta^{2,P1}$ then $\frac{dw_1(\widehat{\theta}_1^*)}{d\widehat{\theta}_1^*} > 1$ irrespective of σ_1 and σ_2.[50]

Following the same steps us above we can show that there exists a bound on σ_θ^2, which we denote by $\overline{\sigma}_\theta^{2,P2}$, such that if $\sigma_\theta^2 > \overline{\sigma}_\theta^{2,P2}$ then $\frac{dw_2(\widehat{\theta}_1^*)}{d\widehat{\theta}_1^*} < 1$ and $\overline{\sigma}_\theta^{2,P2}$ is independent of σ_1 and σ_2.

Setting $\overline{\sigma}_\theta = \max\left\{\overline{\sigma}_\theta^{2,P1}, \overline{\sigma}_\theta^{2,P2}\right\}$ proves the second part of the theorem. ■

Lemma 2 *Suppose that $\sigma_i \to 0$, $\sigma_j \to 0$ and $\frac{\sigma_i}{\sigma_j} \to c$ where $c \in \overline{\mathbb{R}}_+$. If the above game has a unique equilibrium then this equilibrium converges to the risk-dominant equilibrium of the complete*

[50] We use the superscript $P1$ to emphasize the fact that this restriction follows from equilibrium condition of player 1.

information game, i.e. $x_i^* \to 2T$ *and* $x_j^* \to 2T$.

Proof. Agent i's expected payoff of attacking is given by:

$$\int_{-\infty}^{\infty} 1_{\{\theta\in[\underline{\theta},\overline{\theta}]\}}\theta\left(1-F\left(\frac{x_j^*-\theta}{\sigma_j}\right)\right)d\alpha_{\sigma_i}+\int_{\overline{\theta}}^{\infty}\theta d\alpha_{\sigma_i}-T$$

where α_{σ_i} is a measure implied by the CDF of a Normal distribution with mean $\widehat{\theta}_i$ and variance $\widehat{\sigma}_i$ (i.e. α_{σ_i} is a Baire measure implied by F). Then the characteristic function of α_{σ_i} is given by:

$$\phi_{\alpha_{\sigma_i}}(t)=e^{it\widehat{\theta}_i-\frac{\widehat{\sigma}_i^2t^2}{2}}$$

where $\widehat{\theta}_i=\frac{\sigma_i^2x_i+\sigma_\theta^2\mu_\theta}{\sigma_i^2+\sigma_\theta^2}$, $\widehat{\sigma}_i^2=\frac{\sigma_i^2\sigma_\theta^2}{\sigma_i^2+\sigma_\theta^2}$ and $x_i=\theta+\sigma_i\varepsilon_i$. Note that, as $\sigma_i\to 0$ then $\widehat{\sigma}_i^2\to 0$, and so

$$\lim_{\sigma_i\to 0}\phi_{\alpha_{\sigma_i}}(t)=e^{itx_i}$$

But e^{itx_i} is the characteristic function of a probability distribution with mass 1 at θ, which we denote by δ_θ. By the Levy-Cramer Continuity theorem (see e.g. Varadhan 2001) this implies that $\alpha_{\sigma_i}\to\delta_{x_i}$ as $\sigma_i\to 0$, i.e. as the standard deviation converges to 0, the Normal distribution converges to the degenerate distribution with all the mass centered at x_i.

Consider first the limit of $\int_{\overline{\theta}}^{\infty}\theta d\alpha_{\sigma_i}$ as $\sigma_1\to 0$:

$$\lim_{\sigma_i\to 0}\int_{\overline{\theta}}^{\infty}\theta d\alpha_{\sigma_i}=\left(1-F\left(\frac{\overline{\theta}-\widehat{\theta}_i^*}{\sigma_i}\right)\right)+\sigma_i f\left(\frac{\overline{\theta}-\widehat{\theta}_i^*}{\sigma_i}\right)=\begin{cases}1 & \text{if } x_i>\overline{\theta}\\ \frac{1}{2}x_i & \text{if } x_i=\overline{\theta}\\ 0 & \text{if } x_i<\overline{\theta}\end{cases}$$

Let $h(\theta;\sigma_j)\equiv\mathbf{1}_{\{\theta\in[\underline{\theta},\overline{\theta}]\}}\theta\left(1-F\left(\frac{x_j^*-\theta}{\sigma_j}\right)\right)$ and denote by $x_j^{*,\infty}$ the limit of x_j^* as $\sigma_i\to 0$, $\sigma_j\to 0$ and $\frac{\sigma_i}{\sigma_j}\to c\in\mathbb{R}$. Note that h is a bounded, Borel measurable function and is continuous except for a finite set of measure zero. Moreover,

$$\lim_{\sigma_j\to 0}h(\theta;\sigma_j)\to\begin{cases}\theta \text{ if } \theta\in\left(x_j^{*,\infty},\overline{\theta}\right]\\ \frac{1}{2}\theta \text{ if } \theta=x_j^{*,\infty}\\ 0 \text{ otherwise}\end{cases}$$

where we assumed that $x_j^{*,\infty} \in \left[\underline{\theta}, \overline{\theta}\right]$.[51] It follows then that

$$\lim_{\substack{\sigma_i \to 0 \\ \sigma_j \to 0 \\ \frac{\sigma_i}{\sigma_j} \to c}} \int_{-\infty}^{\infty} h(\theta)\, d\alpha_{\sigma_i} \quad \to \quad \int_{-\infty}^{\infty} h(\theta)\, d\delta_{x_i} = \int_{-\infty}^{\infty} \mathbf{1}_{\left\{\theta \in \left[x_j^*, \overline{\theta}\right]\right\}} \theta \left(1 - F\left(\frac{x_j^* - \theta}{\sigma_j}\right)\right) d\delta_{x_i}$$

$$= \begin{cases} x_i^* & \text{if } x_i^* \in \left(x_j^{*,\infty}, \overline{\theta}\right] \\ \frac{1}{2} x_i^* & \text{if } x_i^* = x_j^{*,\infty} \\ 0 & \text{if } x_i^* \notin \left[x_j^{*,\infty}, \overline{\theta}\right] \end{cases}$$

The optimal threshold is defined as the value of a signal that makes agent i indifferent between attacking and not attacking, i.e. it has to solve

$$v_i\left(x_i^*, x_j^*; \sigma_i, \sigma_j\right) = 0$$

Since the same is true for agent j it has to be the case (by the symmetry of the problem) that in the limit $x_i^{*,\infty} = x_j^{*,\infty}$ and so $x_i^{*,\infty}$ solves

$$\begin{aligned} \frac{1}{2} x_i^{*,\infty} - T &= 0 \\ x_i^{*,\infty} &= 2T \end{aligned}$$

It is easy to verify that $2T$ corresponds to a threshold such that if $\theta \geq 2T$ then attacking is a risk-dominant action and if $\theta < 2T$ not attacking is a risk-dominant function.

Above we assumed that $x_j^{*,\infty} \in \left[\underline{\theta}, \overline{\theta}\right]$. Suppose that this is not the case. If $x_j^{*,\infty} > \overline{\theta}$ then following the same argument as before we would find that in the limit

$$x_i^{*,\infty} = \overline{\theta} < x_j^{*,\infty}$$

which contradicts optimality of $x_j^{*,\infty}$ (if agent i follows a threshold $x_i^{*,\infty} = \overline{\theta}$ then the optimal threshold for agent j is $x_j^{*,\infty} = \overline{\theta}$). A symmetric argument establishes that it cannot be the case that $x_j^{*,\infty} < \underline{\theta}$. ■

Theorem (Existence) *There exists a symmetric pure-strategy Bayesian Nash Equilibrium of the game with information acquisition.*

We first prove two simple claims and two corollaries that will make the proof of existence straightforward.

Claim 1 *$B_i(\sigma_i; \sigma_j)$ is a decreasing function of σ_i, that is $B_{i1}(\sigma_i; \sigma_j) \leq 0$.*

Proof. Since $u(\theta, a_i)$ has the single crossing property in (θ, a_i), and that the signal x_i and the unknown parameter θ are affiliated, the claim then follows from Theorem 1 in Persico (2000). ■

Claim 2 *The marginal benefit of increasing the precision of agent i converges to zero as the*

[51] We argue below that only $x_j^{*,\infty} \in \left[\underline{\theta}, \overline{\theta}\right]$ is consistent with equilibrium.

signal noise for agent i vanishes, i.e. $\lim_{\sigma_i \to 0} \frac{\partial}{\partial \sigma_i} B_i(\sigma_i; \sigma_j) = 0$.

This proof is very lengthy and can be obtained from the authors by request. It requires to show that in the limit $\frac{\partial}{\partial \sigma_i} x_i^*(\sigma_i, \sigma_j)$ is bounded and to verify that all the integrals in the expression for the marginal benefit converge to zero.

From the above results we have the following immediate corollaries:

Corollary 1 *The best response functions for both agents are well defined.*

Proof. Since the cost function is strictly decreasing in σ_i and tends to infinite as $\sigma_i \to 0$, and since $B_i(\sigma_i, \sigma_j)$ is positive and stays bounded for each σ_j, we know that for each σ_j there is a unique choice of σ_i, holding beliefs of both players constant. ■

Corollary 2 *In any equilibrium of the game both agents choose to acquire information (increase the precision of their signals).*

Proof. This follows from the fact that the marginal cost of acquiring information is continuous and equal zero at $\sigma_i = \sigma_0$, together with the fact that the marginal benefit of lowering σ_i is strictly positive for $\sigma_i > 0$.[52] ■

Existence. Suppose that agent j believes that whenever he chooses a precision σ_j, agent i will make the same choice. Holding agent j's beliefs fixed, we showed above that the best response function

$$\sigma_i^*(\sigma_j) = \max_{\sigma_i \in [0, \sigma_0]} U_i(\sigma_i, \sigma_j)$$

is well defined. Since $U_i(\sigma_i, \sigma_j)$ is continuous in σ_i and σ_j, by the Theorem of the Maximum we conclude that $\sigma_i^*(\sigma_j)$ is a continuous function. $\sigma_i^*(\cdot)$ is also a self-map: $\sigma_i^* : [0, \sigma_0] \to [0, \sigma_0]$. Hence, by Brower's Fixed Point Theorem, $\sigma_i^*(\cdot)$ has a fixed point. This implies that there exists a σ_j such that if agent j believes that agent i chooses $\sigma_i = \sigma_j$ agent i will find it optimal to choose such a σ_i, that is $\sigma_i^*(\sigma_j) = \sigma_j$. ■

[52] Note that above we established that $B_{i1} \leq 0$ and $\lim_{\sigma_i \to 0} B_{i1} = 0$. It can be shown that $B_{i1}(s_i, \sigma_j) \neq B_{i1}(s_i', \sigma_j)$ $\forall s_i \neq s_i'$. It follows then that $B_{i1}(\sigma_i, \sigma_j) < 0$ whenever $\sigma_i > 0$. That is, decreasing σ_i (increasing the precision) strictly increases the gross payoff for agent i.