CSEF

Centre for Studies in Economics and Finance

WORKING PAPER NO. 470

Terrorism, Counterterrorism and Optimal Striking Rules

Giovanni Immordino, Gülen Karakoç-Palminteri and Salvatore Piccolo

April 2017

University of Naples Federico II

University of Salerno

Bocconi University, Milan

CSEF - Centre for Studies in Economics and Finance
DEPARTMENT OF ECONOMICS – UNIVERSITY OF NAPLES
80126 NAPLES - ITALY
Tel. and fax +39 081 675372 – e-mail: csef@unisa.it
ISSN: 2240-9696

WORKING PAPER NO. 470

Terrorism, Counterterrorism and Optimal Striking Rules

Giovanni Immordino[♥], Gülen Karakoç-Palminteri[♠] and Salvatore Piccolo[♦]

Abstract
We study a simple mechanism design problem that describes the optimal behavior of a country targeted by a foreign terrorist group. The country is uncertain about the terrorists' strength and may decide to acquire such information from the community hosting the terrorists. We highlight a novel trade-off between target hardening – i.e., mitigating the incidence of an attack by strengthening internal controls and improving citizens' protection – and preemptive military measures aimed at eradicating the problem at its root – i.e., a strike in the terrorists' hosting country. We show that, conditional on being informed about the terrorists' strength, the country engages in a preemptive attack only when it faces a sufficiently serious threat and when the community norms favoring terrorists are weak. Yet, in contrast with the existing literature, we show that it is optimal for the country to acquire information only when these norms are strong enough and when its prior information about the terrorists' strength is sufficiently poor.

Classification JEL: D02, D74, D73, D78

Keywords: Counter-terrorism, Striking Rules, Target Hardening, Terrorism

Acknowledgement: We are indebted to David Martimort for his insightful comments and suggestions. Comments received by the audience at the Mannheim Graduate Students Seminar are gratefully acknowledged.

[♥] Università di Napoli Federico II, Department of Economics and CSEF.

[♠] Università di Milano (Bicocca).

[♦] Università di Bergamo and CSEF.

Table of contents

1 Introduction

After September 11, the war on terror has became a primary objective of the western world. Since then, a wide campaign against religious terrorism (primarily Al-Qaeda) has been undertaken, involving open and covert military operations, new security legislation, efforts to block the financing of terrorism, etc. Yet, despite remarkable efforts, the problem seems far from being solved. After the death of Osama bin Laden in 2011, the rise of the ISIS (Islamic State of Iraq and Syria) and the recent dramatic attacks to the heart of Europe have increased again the alert, and thrown serious doubts on the way the risks associated with the new threat have been assessed. After more than 15 years, the major obstacle to the implementation of timely and effective measures is still the 'hidden face' of terror. When will they strike? Where? And, how aggressively?

Answering these questions is difficult, especially in the absence of reliable information on terrorists' activities, network, strength, equipment, etc. This is why, in many cases, it could be useful to collaborate with the local communities hosting the terrorists. Noncombatants, indeed, are likely to own insider knowledge that would hardly be acquired by intelligence agencies — see, e.g., Kalyvas (2006). According to Berman *et al.* (2011), noncombatants are responsive and active actors. Popkin (1979) argues that they make rational decisions regarding the direction and degree of their cooperation, while Galula (1964) and Petersen (2001) show that their propensity to do so varies at the individual level and shifts across space and time.

Hence, the interaction between target countries, terrorists, and the communities whose cooperation they compete for can be best understood by accounting for their preferences and incentives (Nagl, 2002; Sepp, 2005; Petraeus, 2006; Fridovich and Krawchuk, 2007; Cassidy, 2008; McMaster, 2008). To investigate these incentives we study a simple mechanism design problem that describes the optimal behavior of a country targeted by a foreign terrorist group. The country is uncertain about the terrorists' strength (measured by the magnitude of the damage produced by an attack) and may decide to acquire such information from the community hosting the terrorists. As a prize for collaboration, the target country provides the community with basic public goods, infrastructures, technological knowledge, etc.

We highlight a novel trade-off between target hardening, which requires the country to invest public funds to mitigate the incidence of the attack — e.g., by strengthening internal controls and improving citizens' protection — and preemptive attacks aimed at eradicating the problem at its root. Specifically, we show that, conditional on being informed about the terrorists' strength, the country engages in a preemptive attack only when the threat that it faces is sufficiently serious — i.e., the potential damage to civilians and public infrastructures is relatively large — and when the community norms of noncombatants favoring terrorists are weak — i.e., when the terrorists are not very much rooted in the country hosting them. The key point hinges on the mechanism design approach we adopt: in order to elicit truthful information about the strength of the terrorist group, the target country has to grant an information rent to the noncombatants, who would otherwise have an incentive to overstate the strength of the terrorists in order to amplify the risk of retaliation (which occurs when the strike fails) and induce the country to provide larger provisions of public goods to guarantee their collaboration.

Hence, in order to reduce these rents, the country is forced to engage in preemptive attacks relatively more often than in the first-best (i.e., the scenario in which the terrorists' strength is common knowledge) when facing strong terrorist groups. Indeed, *ceteris paribus*, being aggressive to strong terrorist groups mitigates (or even nullifies) the incentive of the noncombatants to overstate the risk of retaliation: a *rent saving* effect. Yet, a preemptive strike is also costly because a military intervention would encroach on the community's internal norms, whose members might either be resilient to foreign interference with their territorial sovereignty, or have ties with the terrorists: a *norm breaching* effect. The optimal policy trades off these two effects. Clearly, although the county has an incentive to strike strong terrorist groups, it does less so as the community features stronger norms.

Building on this insight we then turn to characterize the country's information acquisition decision. The benchmark against which we compare the country's expected payoff from acquiring information is that in which there is no interaction with noncombatants, and the policy is chosen *behind the veil of ignorance* — i.e., without knowing the terrorists' strength. We show that acquiring information is optimal only when: (*i*) the community features strong enough internal norms or non-negligible ties with the terrorist group; (*ii*) the target country's prior information about the terrorists' strength is sufficiently poor. Intuitively, the country's decision between acquiring information and remaining uninformed is shaped by the following trade-off.

On the one hand, by not acquiring information the country bears the risk of making a decision that is inefficient ex-post. On the other hand, when the country decides to remain uninformed, whatever decision is taken, it saves on the cost of acquiring information, which as argued above depends on the strength of the community norms and the rent that the community has to be granted in order to truthfully report its private information.

Surprisingly, when the community of noncombatants features sufficiently weak internal norms, the country prefers not to acquire information. The intuition is straightforward. In this case, conditional on acquiring information, the country is most likely to strike. Yet, when it does so, its ex post payoff is lower than the payoff it obtains when a preemptive strike is conducted behind the veil of ignorance: no public good has to be provided to the community since no information is transmitted. By contrast, as the community norms become stronger, the country is relatively less likely to strike upon acquiring information. Indeed, the value of acquiring information lies in the ability of the informed country to tailor its target hardening choice to the terrorists' damage because this eliminates the risk of making a wrong decision (as opposed to the case of no information acquisition). As a result, the larger is the set of contingencies (states of nature) in which the country decides not to strike, the higher is the value of information. Since it is relatively more appealing for the country to invest in target hardening as the community norms become stronger, the value of information increases too.

Hence, the model delivers an inverse relationship between the incentive to acquire information about the strength of terrorist groups and the norms featured by the communities hosting these terrorists. This result contrasts with the equilibrium characterization offered by Berman *et al.* (2011) who find that norms favoring rebel control reduce incentives to disclose information in a model where information is not elicited through incentive compatible mechanisms but it is spontaneously disclosed by the community in a 'cheap talk' fashion. Their model features an equilibrium in which communities disclose information if and only if norms are sufficiently weak. By contrast, in our environment countries have an incentive to acquire this information if and only if norms are sufficiently strong.

Finally, notice that although our model is very stylized, we believe that it captures some important aspects of reality. First, the distinction between defensive and offensive strategies and their optimal use contingent on the severity of the terrorists potential damage, is a natural aspect of the problem in real life. Second, allowing the target country to decide not to acquire information — i.e., the extensive margin of the information acquisition problem — complements the cheap talk approach featured in earlier contributions, and seems more compelling when there is a substantial imbalance between the bargaining power of the country and the community of noncombatants. Third, our simple set-up is flexible enough to accommodate additional aspects of the problem that will be discussed in Section 4 and that we hope to address in future research.

The paper is organized as follows. In the next section (Section 2) we relate our paper to the existing literature. In Section 3 we set up the model. In Section 3.1, we analyze the first-best benchmark. The analysis under asymmetric information is developed in Section 3.2, where we characterize the optimal behavior of the uninformed country and the informed one. In Section 3.3, we identify the incentive of the target country to get informed and the relevant comparative statics. Section 4 discusses possible extensions of the model, traces a research agenda for future work and concludes.

2 Related literature

There exists a fairly developed literature in economics studying terrorism and counterterrorism, to which our paper relates. The closest contribution is Berman *et al.* (2011), who provide a model that captures a three-way interaction between insurgent organizations, government forces and a local community. As already said, an important difference between our model and Berman *et al.* (2011) is that we take a mechanism design approach, while they analyze a game in which the community decides how much information to disclose in a 'cheap-talk' fashion. Hence, in this sense the two models are complementary. Nevertheless, while we make a formal difference between target hardening and military measures, they model the country's counterterrorism activity as a unidimensional decision variable. In contrast with our results, in their model norms favoring rebel control reduce the community incentive to disclose information.

Like us, Powell (2006) considers a model where a target country can engage in preemptive attacks. He adopts a bargaining approach and focus on commitment problems to argue that even if violence occurs in equilibrium, it is inefficient in some cases. In our model, this is not the case when the country acquire information, while it can well be the case with an uninformed country. The main trade-off between acquiring information and acting behind the veil of ignorance hinges precisely on this tension.

As for target hardening, Hastings and Chan (2013) develop a simple cheap-talk model that highlights the relationship between target hardening and the value that a terrorist group derives from attacking it. They compare how the expected value of attacking a hardened target varies depending on whether terrorists just maximize the physical damage inflicted to the target, or if they also attach a symbolic value to the attack (even when it fails). They find that there are some benefits in hardening a target since it decreases the probability of an attack and, almost by definition, raises the loss of the terrorist group. Differently, they also stress that the marginal benefits of hardening the target decreases because of the significant rise in the symbolic value of the target to terrorist group. We do not model this interaction and assume that terrorists always attack in order to produce the most harmful damage to the target. However, we focus on the information acquisition aspect that is neglected by Hastings and Chan (2013).

Finally, as far as the uncertainty about the terrorists preferences is concerned, Wang and Bier (2011) consider a model of incomplete information in which the defender is uncertain about the attacker's weights on which target to attack. They consider a dynamic game with incomplete information in which the defender first chooses how to allocate his defensive resources, and then an attacker chooses which target to attack according to his preferences. The defender's uncertainty about attacker's preferences is modeled by a subjective distribution representing both defender uncertainty about the attacker weights on the various known attributes and also defender ignorance about any unobserved attributes. In our model there is only one target, and the uncertainty is about the potential harm produced by the attack. Yet, while they allow only to defend the target, we also consider the possibility of a preemptive attack.

3 The model

A country is targeted by a foreign terrorist group whose attack produces an harm (damage) $\mathcal{H}(\theta, d) \leqslant 0$, with $\mathcal{H}_\theta(\theta, d) < 0$ and $\mathcal{H}_d(\theta, d) \geqslant 0$. The parameter θ is a random variable, distributed on the compact support $\Theta \triangleq [\underline{\theta}, \overline{\theta}] \subseteq \Re^+$ with cdf $F(\theta)$ and pdf $f(\theta)$, which measures (other things being equal) the severity (incidence) of the harm. As a convention, we assume that the higher θ, the more harmful the attack, which explains why $\mathcal{H}(\cdot)$ is decreasing in θ. The interpretation of θ is that it reflects the military strength (violence) of the terrorist group. In order to shield against the attack, the country can invest public funds in counterterrorism activities $d \geqslant 0$ (target hardening) that mitigate the incidence of the attack, which explains why $\mathcal{H}(\cdot)$ is increasing in d — e.g., introduction of metal detectors, mandatory passenger screening, fortification and protection of government buildings, etc. For simplicity, we assume that the monetary cost of defense is linear, and it simply equals d.

In addition to counterterrorism activities, the country can also engage in a preemptive attack striking the terrorists in the hosting country. We denote by $s \in \{0, 1\}$ the striking decision: $s = 1$ if a strike occurs, $s = 0$ otherwise. Again, we assume that the monetary cost of conducing a military campaign abroad is linear and equal to λ. We posit that λ is larger than 1 to capture the idea that — differently from the cost of defense — the actual cost of the military campaign weights also non-pecuniary aspects such as political dissent stemming from the possible loss of lives during the campaign, and so on.

The country has no *a priori* information about the strength of the terrorist group. Yet, this information can be acquired by dealing with the community of noncombatants hosting the terrorists. These people own private information about θ and are willing to disclose it as long as the target country provides public goods g as a prize for collaboration.

Let $\alpha = \Pr[s = 1] \in [0, 1]$ be the probability of a strike. The community is risk neutral and has a utility function

$$u(g, \alpha, \theta) \triangleq g - \alpha x - \theta(1 - \alpha).$$

The parameter x can be interpreted as a measure of the weight that the community assigns to its territorial sovereignty or, in other words, as the strength of the norms against foreign control of their territory. Alternatively, x could simply reflect a measure of how rooted the terrorist group is within the community: the higher x, the more rooted in the community the terrorists are. This cost is paid when the strike occurs — i.e., with probability α. By contrast, when the country decides not to strike, the terrorist group exerts retaliation over the community who bears a loss equal, for simplicity, to the strength θ of the terrorist group. Indeed, obvious reputation concerns may induce terrorists to discipline the community when it collaborates with the country.[1]

[1] Kalyvas (2006) argues that rebels and terrorist groups typically engage in violence against non-

As we will explain below, acquiring information allows the country to always take the best (interim) decision. The country's payoff is

$$\mathcal{V}(\alpha, d, g, \theta) \triangleq (1-\alpha)\mathcal{H}(\theta, d) - d - \lambda\alpha - g,$$

Hence, the country must first decide whether to acquire information about θ or not, and then it chooses the intensity (probability) of the strike and how much to invest in counterterrorism.

Without loss of generality we assume that, if the country decides to acquire information, it offers by the Revelation Principle (Laffont and Martimort, 2002) a direct mechanism

$$\mathcal{M} \triangleq \{g(m), \alpha(m)\}_{m\in\Theta},$$

which specifies an amount of public goods $g(\cdot)$ and a probability of strike $\alpha(\cdot)$ both contingent on the community report m about the strength of the terrorist group θ. Of course, given $g(\cdot)$ and $\alpha(\cdot)$, the country will optimally set the counterterrorism activity $d(\cdot)$ so as to maximize its (expected) payoff. As we will explain below, since d has no direct impact on the community's payoff, there is no loss of generality in restricting attention to mechanisms where the country does not commit to a defense level *vis-à-vis* the community.[2]

For tractability, and without loss of insights, throughout we assume that

- **(A1)** The harm is quadratic

$$\mathcal{H}(\theta, d) \triangleq -\frac{1}{2}[\theta - d]^2.$$

A quadratic loss function is typically used in the literature to obtain (tractable) closed form solutions.

- **(A2)** $F(\theta)$ exhibits an increasing (inverse) hazard rate — i.e., $h(\theta) \triangleq F(\theta)/f(\theta)$ is increasing in θ. Moreover, it features mean μ and variance σ^2, with $\Delta\theta \triangleq \overline{\theta} - \underline{\theta} \geqslant 0$.

Assuming an increasing (inverse) hazard rate is standard in most screening models.

combatants to discourage their collaboration with target countries.

[2] Considering such an extended mechanism would also seem unrealistic since the country's actual choice of d is hardly verifiable by the community.

- **(A3)** The harm is not too strong in expectation — i.e., $\mu < \frac{1}{2} + \lambda$. Moreover, $\sigma^2 \geqslant \sqrt{2\lambda}$ and
$$\underline{\theta} \geqslant \max\left\{\sqrt{2\lambda}, \sqrt{2(x+\lambda)+1} - 1\right\}.$$

These assumptions simply guarantee that, there exist a non-empty region of the model parameters, such that the country strikes regardless of whether it is informed or not. The complementary region of parameters in which there is no strike is uninteresting for our purposes.

The solution concept is Subgame Perfect Nash Equilibrium (SPNE).

3.1 The first-best benchmark

As a benchmark, consider first the case in which the country knows θ. In that case, there is no need to waste funds in providing public goods to the community since the strength of the terrorist group is common knowledge. Hence, for every θ the optimal d and α solve

$$\max_{\alpha\in[0,1],d\geqslant 0} \mathcal{V}(\alpha, d, \theta) \triangleq \max_{\alpha\in[0,1],d\geqslant 0} (1-\alpha)\,\mathcal{H}(\theta, d) - \alpha\lambda - d,$$

where, abusing slightly notation, we have defined $\mathcal{V}(\alpha, d, \theta) \triangleq \mathcal{V}(\alpha, d, g = 0, \theta)$.

Differentiating with respect to α we have

$$\frac{\partial \mathcal{V}(\alpha, d, \theta)}{\partial \alpha} = \underbrace{-\mathcal{H}(\theta, d)}_{\text{Harm avoidance}} - \lambda \lesseqgtr 0. \tag{1}$$

Clearly, a higher probability of striking lowers the harm since, by assumption, the attack eradicates the terrorists' threat (harm avoidance). Hence, it is optimal to strike if and only if the benefit that the country obtains from the avoidance of the harm is larger than the cost λ of the military campaign.

Differentiating with respect to d we have

$$\frac{\partial \mathcal{V}(\alpha, d, \theta)}{\partial d} = \underbrace{(1-\alpha)\,\mathcal{H}_d(\theta, d)}_{\text{Target hardening}} - 1 \leqslant 0. \tag{2}$$

The optimal defense level trades off two effects: the cost of investing public funds in counterterrorism activities against the reduced incidence of the harm due to these activities (target hardening). The $\leqslant$ sign comes from the fact that $\mathcal{H}(\theta, d)$ is strictly concave in d. Then, the solution is either 0 — i.e., (2) holds with a strict inequality — or interior — i.e., (2) holds with equality.

Combining (1) and (2), we can establish the following result.

Proposition 1 *When the country knows θ (complete information), the optimal policy is such that $\alpha^{FB}(\theta) = 1$ and $d^{FB}(\theta) = 0$ if and only if*

$$\theta \geqslant \theta^{FB} \triangleq \frac{1}{2} + \lambda.$$

Otherwise, for every $\theta \leqslant \theta^{FB}$, it is $\alpha^{FB}(\theta) = 0$ and $d^{FB}(\theta) = \theta - 1 > 0$.

The solution of the first-best problem is clearly 'bang-bang' since the country's payoff is linear in α. As intuition suggests, when there is no uncertainty about the terrorists' strength, it is optimal to strike only when the harm is sufficiently large, namely when the terrorists are strong enough. Clearly, other things being equal, as the harm becomes more severe — i.e., as θ grows large — the country has more incentive to strike, while it has less incentive to do so as the cost of the military campaign rises — i.e., as λ grows large. Of course, when the strike does not take place, the optimal level of public funds invested in target hardening is increasing with harm.

3.2 Asymmetric information

Suppose now that the country has no information about θ. As explained before, there are two viable options. The country can either base its strategy on the prior information it owns about θ, or it can acquire information from the community and make a decision that is ex post optimal. Yet, in order to ensure participation of the community to the deal and elicit a truthful report, the country has to give up an information rent to the community, which might either not accept the contract or misreport θ. In what follows we characterize the optimal counterterrorism policy for each information acquisition choice and then compare the two outcomes.

Uninformed country. If the country decides to be uninformed, its optimization problem is similar to that solved in the complete information benchmark with the difference that the harm has to be taken in expected value. Hence, the optimal policy under no information acquisition solves the following maximization problem:

$$\max_{\alpha \in [0,1], d \geqslant 0} \mathcal{V}(\alpha, d) \triangleq \max_{\alpha \in [0,1], d \geqslant 0} \int_{\theta} \mathcal{V}(\alpha, d, \theta)\, dF(\theta).$$

Differentiating with respect to α we have again a trade off between the harm avoidance effect (in expectation though) and the cost of striking — i.e.,

$$\frac{\partial \mathcal{V}(\alpha, d)}{\partial \alpha} = -\int_\theta \mathcal{H}(\theta, d)\, dF(\theta) - \lambda \lesseqgtr 0.$$

Differentiating with respect to d we have a trade off between the target hardening effect (again in expectation) and the cost of defense — i.e.,

$$\frac{\partial \mathcal{V}(\alpha, d)}{\partial d} = (1-\alpha)\int_\theta \mathcal{H}_d(\theta, d)\, dF(\theta) - 1 \lesseqgtr 0.$$

Hence, it is not difficult to guess that the solution of the uninformed country's maximization problem has a structure that is similar to that of the first-best. The difference being that, since the terrorists' strength is unknown, the costs and benefits associated with the use of each policy instrument must be taken in expected terms. Because the country is risk averse — i.e., the harm $\mathcal{H}(\cdot)$ is quadratic — this means that the variance of θ now plays an important role in the analysis.

Proposition 2 *When the country is uninformed about θ, the optimal counterterrorism policy is such that $\alpha^N = 1$ and $d^N = 0$ if and only if*

$$\sigma^2 \geqslant \sigma_0^2 \triangleq 1 - 2(\mu - \lambda).$$

and $\alpha^N = 0$ and $d^N = \mu - 1$ otherwise.

The behavior of an uninformed country is rather simple, and is shaped by the following trade-off. On the one hand, investing in defense only entails some risk for the country because defense is costly and it does not completely neutralize the harm. To see why, notice that the expected harm evaluated at d^N is

$$\begin{aligned}
\int_\theta \mathcal{H}\left(\theta, d^N\right) dF(\theta) &\triangleq -\int_\theta \frac{[\theta - (\mu - 1)]^2}{2} dF(\theta) \\
&= -\int_\theta \frac{[\theta - \mu + 1]^2}{2} dF(\theta) \\
&= -\frac{\sigma^2 + 1}{2},
\end{aligned}$$

which is decreasing in σ^2 since the country is risk averse (i.e., the harm is quadratic in θ). On the other hand, a strike completely neutralizes the harm, but the military campaign is costly. Hence, the higher σ^2 relative to λ, the more appealing the strike.

Finally, we can compute the expected payoff of the uninformed country:

$$\mathcal{V}^N \triangleq \begin{cases} \frac{1}{2} - \mu - \frac{\sigma^2}{2} & \Leftrightarrow \sigma^2 \leqslant \sigma_0^2 \\ -\lambda & \Leftrightarrow \sigma^2 \geqslant \sigma_0^2 \end{cases},$$

which is weakly decreasing in λ, μ and σ^2.

As explained before, the uninformed country strikes, and completely eradicates the threat, if and only if there is enough uncertainty about the terrorists' strength ($\sigma^2 \geqslant \sigma_0^2$) — e.g., because the group of terrorists is new and very little is known about them. In this case, the country's expected payoff is simply equal to the expected cost of the military campaign. Differently, when the uncertainty about the strength of the terrorist group is low ($\sigma^2 \leqslant \sigma_0^2$), the country prefers to take the risk of being attacked and invests only in defense to reduce the incidence of the attack. Hence, the expected payoff falls with the average strength of the terrorists μ and with its variance, which measures the risk to which the country is exposed when there is no strike.

Informed country. Suppose now that the country acquires information from the community before deciding whether to strike or not. The information released by the community is truthful if and only if the mechanism $\mathcal{M}$ is incentive compatible. Let

$$u(\theta, m) \triangleq g(m) - x\alpha(m) - \theta(1 - \alpha(m)),$$

denote the community's expected payoff when it reports m and the true state of nature is θ. And, abusing slightly notation, denote by

$$u(\theta) \triangleq u(\theta, m = \theta)$$

the community's information rent when it truthfully reveals the terrorists' strength. An incentive compatible policy requires the following standard implementability conditions

$$\left.\frac{\partial u(\theta, m)}{\partial m}\right|_{m=\theta} = 0 \quad \Longrightarrow \quad u(\theta) = u(\overline{\theta}) + \int_{\theta}^{\overline{\theta}} (1 - \alpha(z))\, dz, \tag{3}$$

and

$$\left.\frac{\partial^2 u(\theta, m)}{\partial m^2}\right|_{m=\theta} \leqslant 0 \quad \Longrightarrow \quad \dot{\alpha}(\theta) \geqslant 0. \tag{4}$$

Condition (3) simply reflects the local incentive compatibility constraint: it guarantees that the community has no incentive to misreport locally the strength of the terrorist group.[3] Essentially, this condition delivers the information rent that the

[3] It can be easily shown that incentive compatibility is satisfied also globally if it holds locally (see, e.g., Laffont and Martimort, 2002).

country has to give up to the community in order to elicit truthful information. Notice that the community's rent is decreasing with the strength of the terrorist group — i.e., $\dot{u}(\theta) \leqslant 0$ — which means that communities hosting weaker terrorist groups are able to extract higher rents when collaborating with target countries. The intuition is that stronger terrorists retaliate more harshly. Hence, the community has an incentive to overstate the risk of retaliation in order to be offered a better deal (i.e., more public good provision). In addition, it is useful to observe that the rent enjoyed by the community is decreasing in the probability of a strike. Indeed, the larger is $\alpha(\cdot)$, the more protected the community feels against the risk of retaliation, and the less costly it is for the country to elicit truthful information.

Condition (4), instead, reflects the so called 'monotonicity' constraint which states that any incentive feasible mechanism must be such that the stronger is the terrorist group, the more likely it is that a strike will occur. This requirement is equivalent to the standard Spence-Mirleess single-crossing condition.

The country's maximization problem is

$$\max_{\alpha(\cdot)\in[0,1],g(\cdot)\geqslant 0,d(\cdot)\geqslant 0} \int_{\theta} \mathcal{V}(\alpha(\theta), d(\theta), g(\theta), \theta)\, dF(\theta),$$

subject to (3), (4) and

$$u(\theta) \geqslant 0 \quad \forall \theta \in \Theta. \tag{5}$$

Neglecting the monotonicity condition (4) and using a standard change of variable — i.e., optimizing with respect to $u(\cdot)$ instead of $g(\cdot)$ — simple integration by parts of the rent in (3) (see, Laffont and Martimort, 2002) allows to rewrite the country's maximization problem as

$$\max_{\alpha(\cdot)\in[0,1],d(\cdot)\geqslant 0} \int_{\theta} \{\mathcal{V}(\alpha(\theta), d(\theta), \theta) - \alpha(\theta)x - (1-\alpha(\theta))[\theta + h(\theta)]\}\, dF(\theta). \tag{6}$$

Notice that, as usual in these models, we have optimally set $u(\overline{\theta}) = 0$ because the community hosting the most violent terrorists can only under-report its retaliation loss, but this is clearly not optimal. As a result, the $\overline{\theta}$-type is left with no rent.

Again, the objective function of the country is linear in $\alpha(\cdot)$ and its structure reflects the standard trade off between efficiency and rent-extraction. Essentially, in order to elicit truthful information, the country must give up an information rent to the community, which will in turn affect the optimal striking rule. Yet, relative to the first-best benchmark, dealing with the community is costly and the structure of this cost depends on the probability of a strike. Differentiating with respect to $\alpha(\cdot)$ we have

$$\underbrace{-\mathcal{H}(\theta, d(\theta)) - \lambda}_{\text{First-best rule}} - \underbrace{x}_{\text{Norm breaking}} + \underbrace{\theta + h(\theta)}_{\text{Rent saving}} \lessgtr 0. \tag{7}$$

The new trade off that the country faces when it acquires information is essentially between the cost of breaching the community's norms and the information rent that it has to pay when deciding to invest in target hardening only. First, since the community is averse to hosting a foreign army, the level of public good that has to be provided by the country when it strikes must compensate the community for the loss due to the breach of its norms. Second, when the country does not strike, retaliation by the terrorist group occurs. Hence, the public good that has to be provided in this case exceeds the participation level θ, because it also takes into account the rent that is necessary to pay in order to induce the community to report truthfully the terrorists' strength. Other things being equal, the higher θ the more willing the country is to strike because a relatively higher probability of striking in state θ mitigates the incentive to mimic of the community with type below θ, as reflected by the increasing inverse hazard rate $h(\theta)$ that measures the mass of types below θ.

Differentiating with respect to $d(\cdot)$ it is easy to show that the same condition as in the first-best obtains. This echoes the dichotomy result in Laffont and Tirole (1993, Chapter 3). Essentially, since counterterrorism does not affect the community's payoff, only the striking decision is used as a screening device, while the defense level is set at its first-best rule. Notice that this same outcome would occur if the community committed to an extended mechanism that includes also the defense level as a screening instrument, which explains why it is without loss of generality to exclude d from the mechanism $\mathcal{M}$.

We can thus establish the following result:

Proposition 3 *Define*

$$0 < \underline{x} \triangleq 2\underline{\theta} - \frac{1}{2} - \lambda < \overline{x} \triangleq \underline{x} + 2\Delta\theta + h\left(\overline{\theta}\right).$$

When the country acquires information, the optimal policy is such that:

- *If* $x \leqslant \underline{x}$ *then* $\alpha^I(\theta) = 1$, $d^I(\theta) = 0$ *and* $g^I(\theta) = x$ *for every* $\theta \in \Theta$.

- *If* $x \in (\underline{x}, \overline{x})$ *then:* $\alpha^I(\theta) = 1$, $d^I(\theta) = 0$ *and* $g^I(\theta) = x$ *when* $\theta \geqslant \theta^*$. *Otherwise,* $\alpha^I(\theta) = 0$, $d^I(\theta) = \theta - 1$ *and*

$$g^I(\theta) = \theta + \int_{\theta}^{\theta^*} \left(1 - \alpha^I(\theta)\right) dz = \theta^*.$$

  The threshold $\theta^* \in \left(\underline{\theta}, \overline{\theta}\right)$ *is the unique solution of*

$$2\theta + h(\theta) = \frac{1}{2} + x + \lambda,$$

  and it is increasing in x *and* λ.

- *If* $x \geqslant \overline{x}$, *then* $\alpha^I(\theta) = 0$ *and* $d^I(\theta) = \theta - 1$ *for every* $\theta \in \Theta$.

The intuition behind this result is rather simple. As in the first-best, linearity of the country's payoff implies a bang-bang solution also when getting informed is costly. Yet, with asymmetric information, the optimal strategy depends not only on the terrorists' actual strength, but also on the extent of the community norms. More precisely, depending on the magnitude of x either a pooling or a separating outcome may occur. Notice that, in both cases, the monotonicity condition (4) is satisfied.

First, when the community's internal norms are sufficiently weak ($x \leqslant \underline{x}$) — e.g., because its members are not too resilient to host a foreign army or because they have weak ties with the terrorists — the optimal policy requires a pooling outcome in which the country strikes no matter how strong the terrorist group is. As intuition suggests, the threshold $\underline{x}$ is decreasing in λ and increasing in $\underline{\theta}$. Intuitively, as λ grows large it becomes more costly to strike and therefore a pooling outcome such that the country always strikes regardless of θ becomes less appealing. Differently, as the harm becomes more severe — i.e., as the support of θ shifts to the right (higher values of $\underline{\theta}$) — the information rent enjoyed by the community when a strike does not occur increases, and this makes the country (other things being equal) less willing to take the risk of investing only in target hardening.

Second, when the community norms are not so weak to induce the country to strike no matter what, but neither so large to make a strike not worth at all — i.e., $x \in (\underline{x}, \overline{x})$ — the optimal policy is such that the country intervenes militarily if and only if the terrorists are sufficiently strong. That is, a strike occurs when θ exceeds the threshold θ^* that is determined endogenously to trade off the cost that the country has to pay in order to compensate the community members for violating their norms and the information rents, induced by the risk of retaliation, that it would have to give up when the information received by the community is used to fine tune defense only. As intuition suggests, the threshold θ^* is increasing in x and λ since is relatively less profitable to strike if the community features stronger internal norms and/or if the military cost of the campaign is higher.

Third, when the community norms are very strong — i.e., $x \geqslant \overline{x}$ — striking becomes excessively costly, and the optimal policy is such that the country prefers a strategy based on target hardening only. The threshold $\overline{x}$ is increasing with the width of the terrorists' types $\Delta\theta$ since a wider support of types is associated with larger rents, and it is increasing with the (inverse) hazard rate function which is also a direct measure of rents as explained above.

Interestingly, the level of public goods provided by the country is increasing in the norms' severity x and (weakly) increasing in the cost of the strike λ. That is, the more the terrorists are rooted in the local communities that host them, the larger is the public goods provision requested in order to reveal truthful information. Similarly, the investment in public goods is higher when the target country faces a higher military cost to implement a successful strike because, in this case, it is more likely that it will invest only in target hardening leaving the community exposed to retaliation by the terrorists.

Finally, by inspecting (7) we can compare the first-best with the behavior of an uninformed country which decides to acquire information. Letting θ_0 be the value of the terrorists' strength at which the solution of the second- and first-best problems are equivalent, that is, the solution of

$$x = \theta + h(\theta),$$

it can be immediately shown that

Corollary 4 *For any $\theta > \theta_0$ (resp. $<$) in the second-best striking becomes more (resp. less) appealing than in the first-best.*

Hence, when the country acquires information, weak terrorist groups are attacked relatively less often than in the first-best, while the country behaves more aggressively *vis-à-vis* stronger and more violent groups. Obviously, although these distortions will not affect the defense level on the intensive margin, they affect the extensive margin — i.e., the extent to which each instrument is used.

Summing up, we can compute the country's expected payoff from acquiring information. The result stated below will turn quite useful when we will provide the interpretation for the comparison between the country's expected payoff with and without information acquisition.

Proposition 5 *When the country acquires information, its expected payoff is*

$$\mathcal{V}^I \triangleq \underbrace{-\lambda}_{\textit{Striking cost}} + \underbrace{\int_{\underline{\theta}}^{\theta^*} F(\theta)\, d\theta + F(\theta^*)\, h(\theta^*) - x}_{\textit{Rent saving}},$$

which is increasing in θ^*.

Hence, the country's expected payoff from acquiring information can be written as the cost of striking, plus the expected rent-saving effect that a strike brings about. That is, the community's expected information rent

$$\int_{\underline{\theta}}^{\theta^*} F(\theta)\, d\theta + F(\theta^*)\, h(\theta^*),$$

minus the cost x of breaching its internal norms. Of course, the larger is the set of states in which it is optimal to invest in target hardening only, the stronger is the rent saving effect, whereby increasing the benefit that the country obtains when it acquires information.

3.3 Optimal information acquisition rule

Now we turn to analyze the information acquisition decision of the country. The comparison between V^I and V^N yields (see the Appendix):

$$\mathcal{V}^I - \mathcal{V}^N \triangleq \begin{cases} \int_{\underline{\theta}}^{\theta^*} F(\theta)\, d\theta + F(\theta^*)\, h(\theta^*) - x & \Leftrightarrow \sigma^2 \geqslant \sigma_0^2 \\ \int_{\underline{\theta}}^{\theta^*} F(\theta)\, d\theta + F(\theta^*)\, h(\theta^*) - (x+\lambda) - \frac{1}{2} + \mu + \frac{\sigma^2}{2} & \Leftrightarrow \sigma^2 \leqslant \sigma_0^2 \end{cases}.$$

Hence, we can establish the following result:

Proposition 6 *The optimal information acquisition rule has the following features:*

- *For $\sigma^2 \geqslant \sigma_0^2$ acquiring information is optimal — i.e., $\mathcal{V}^I \geqslant \mathcal{V}^N$ — only if $\overline{\theta} \leqslant \sigma_0^2/2$, and if $x \geqslant x^*$, with $x^* \in (\underline{x}, \overline{x})$ being the unique solution in x of*

$$\int_{\underline{\theta}}^{\theta^*} F(\theta)\, d\theta + F(\theta^*)\, h(\theta^*) - x = 0.$$

  Otherwise, it is never optimal to acquire information.

- *For $\sigma^2 \leqslant \sigma_0^2$ acquiring information is optimal — i.e., $\mathcal{V}^I \geqslant \mathcal{V}^N$ — only if $\overline{\theta} \leqslant \sigma_0^2/2$ and $x \geqslant x^*$, and if*

$$\sigma^2 \geqslant \sigma_0^2 - 2\left[\int_{\underline{\theta}}^{\theta^*} F(\theta)\, d\theta + F(\theta^*)\, h(\theta^*) - x\right].$$

Otherwise, it is never optimal to acquire information.

The country's decision between acquiring information and remaining uninformed is shaped by the following trade-off. On the one hand, by not acquiring information the country bears the risk of making a decision that is inefficient ex-post. On the other hand, when the country decides to remain uninformed, whatever decision is taken, it saves on the cost of acquiring information, which as argued above depends on the strength of the community norms and the rent that the community has to be granted in order to truthfully report its private information.

Surprisingly, when the community's norms are sufficiently weak the country prefers not to acquire information. To see why recall that θ^* is increasing in x. Hence, for low values of x the country strikes relatively more often and, when it does, its ex post payoff is $-(x+\lambda)$. Therefore, conditional on striking, the country is always worse off when it acquires information — i.e., $-\lambda > -(\lambda + x)$ — which is the case also in expectation when x is low enough to induce θ^* being close to $\underline{\theta}$. This suggests that acquiring information should be optimal if the rent saving effect discussed above is strong enough — i.e., if $x \geqslant x^*$. Yet, x sufficiently large is not the only condition that is needed in order for the country to be willing to acquire information: $\overline{\theta}$ must also be not too large. The reason is simple: strong norms (high x) imply that the set of contingencies in which the informed country invests in target hardening only is large, and in these contingencies a rent is paid to the community. Now, if $\overline{\theta}$ is large, these rents are large too (see, e.g., equation 3), hence acquiring information becomes costly. In the region of parameters where the uninformed country does not strike, the condition for the country to acquire information becomes even stronger: in addition to a sufficiently high x and a not too large $\overline{\theta}$ it must also be σ^2 not too small. The reason is simple: if there is little uncertainty (σ^2 small) about the terrorists' strength, then acquiring information becomes too costly relative to being uninformed because in the latter case the risk to which the country is exposed when it invests in target hardening only is small.

Summing up, the model delivers an inverse relationship between the incentive to acquire information about the strength of terrorist groups and the norms featured by the communities hosting these terrorists. This result contrasts with the equilibrium characterization obtained by Berman *et al.* (2011) who find that norms favoring rebel control reduce incentives to disclose information. Their model features an equilibrium in which communities disclose information if and only if norms are sufficiently weak. By contrast, in our model countries have an incentive to acquire this information if and only if norms are sufficiently strong.

Turning to the comparative statics, we can show the following result.

Proposition 7 *x^* is decreasing in λ.*

Intuitively, as λ grows large, there is more incentive to acquire information (other things being equal) because by not doing so, the country is more exposed to the risk of making the wrong choice.

4 Concluding remarks

In this paper we have developed a mechanism design approach to study the relationship between terrorism, counterterrorism and information acquisition. By so doing, we have highlighted a novel tension between target hardening and preemptive attacks. Specifically, we showed that a country targeted by a group of terrorists engages in preemptive attacks, which eradicate the threat at its root, only when it faces sufficiently strong terrorists and when the community of noncombatants from which information about terrorists is elicited features weak norms, or if these people have poor connections with the terrorists. Yet, in contrast with the existing literature, in our model it is optimal for the target country to acquire information about the strength of the terrorists only when noncombatants feature strong enough internal norms and when there is enough uncertainty about the terrorists' strength. This suggests that using public goods provision as a way to attract communities of noncombatants in order to elicit from them information about terrorists is not always ex ante efficient.

Although the model is highly stylized, we believe it captures some important aspects of reality and that, in addition, it provides a solid basis for a broader research agenda. First, the distinction between defensive and offensive strategies and the fact that each of these instruments is used depending on the severity of the terrorists' potential damage, seem to be natural aspects of the real life problem. Second, allowing the target country to gather information about the terrorists' strength — i.e., the extensive margin on the information acquisition decision — seems in some cases more realistic than the cheap-talk approach featured in earlier contributions, especially when the community hosting terrorists is relatively smaller than the target country. Finally, the model also provides a flexible framework to address — in future research — related issues that have not been addressed here. Namely, the decision making problem of a country targeted by more than one terrorist group, the introduction of violence as one of the terrorists' decision variables, the design of a coalition against terror and the related common agency problem, the effects of dynamics, etc.

Appendix

Proof of Proposition 1. Differentiating the objective function $\mathcal{V}(\cdot)$ with respect to α we have

$$\frac{1}{2}[\theta - d]^2 - \lambda. \tag{8}$$

Differentiating with respect to d we have

$$(1-\alpha)[\theta - d] - 1. \tag{9}$$

First, notice that if $\alpha = 1$, then (9) is negative, so that $d = 0$. Hence, $\alpha = 1$ is admissible if and only if

$$\frac{\theta^2}{2} - \lambda \geqslant 0 \quad \forall \theta, \tag{10}$$

which is always true as long as $\underline{\theta} \geqslant \sqrt{2\lambda}$.

Next, we show that if $d > 0$, then $\alpha = 0$. Suppose, by contradiction, that this is not the case. For given $\alpha < 1$, the objective function is strictly concave in d. Hence, if $d > 0$ it must be equal to

$$\widehat{d}(\theta) \triangleq \theta - \frac{1}{1-\alpha}.$$

Evaluating the objective function at $\widehat{d}(\theta)$ we have

$$\mathcal{V}(\alpha, \widehat{d}(\theta), \theta) = \frac{1}{2(1-\alpha)} - \theta - \lambda\alpha. \tag{11}$$

Moreover, notice that if $d > 0$ then $\alpha > 0$ if and only if (8) is positive. Therefore,

$$\alpha \geqslant 1 - \frac{1}{\sqrt{2\lambda}}. \tag{12}$$

Maximizing (11) with respect to α subject to (12) we have

$$\frac{1}{2(1-\alpha)^2} - \lambda + \eta = 0, \tag{13}$$

where $\eta \geqslant 0$ is the Lagrangian multiplier associated to (12). If $\eta > 0$ then (12) binds and (13) implies $\eta = 0$: a contradiction. Moreover, if $\eta = 0$ and $\alpha > 0$, the first order condition (13) yields another contradiction since $\alpha = 1 - \frac{1}{\sqrt{2\lambda}}$. Hence, $\alpha = 0$ whenever $d > 0$.

Finally, to show the result we need to sign the difference

$$\mathcal{V}(\alpha = 1, d = 0, \theta) - \mathcal{V}(\alpha = 0, d = \theta - 1, \theta) = \theta - \lambda - \frac{1}{2},$$

which yields immediately the result. ■

Proof of Proposition 2. Differentiating the objective function $\mathcal{V}(\alpha, d)$ with respect to α we have

$$\frac{1}{2}\int_{\theta}(\theta-d)^{2}\,dF(\theta)-\lambda, \tag{14}$$

while differentiating with respect to d we have

$$(1-\alpha)\int_{\theta}(\theta-d)\,dF(\theta)-1. \tag{15}$$

First, notice that if $\alpha = 1$, then (15) is negative, so that $d = 0$. Hence, $\alpha = 1$ is admissible if and only if

$$\frac{\sigma^{2}}{2}-\lambda \geqslant 0, \tag{16}$$

which is true as long as $\sigma^2 \geqslant 2\lambda$.

In what follows, we show that if $d > 0$, then $\alpha = 0$. Suppose, by contradiction, that this is not the case. Then for given $\alpha < 1$, the objective function is strictly concave in d. Hence, if $d > 0$ it must be equal to

$$\widehat{d} \triangleq \mu-\frac{1}{1-\alpha}.$$

Substituting for $\widehat{d}$ into the objective function, we have

$$\mathcal{V}(\alpha,\widehat{d})=\frac{1}{2(1-\alpha)}-\lambda\alpha-\mu. \tag{17}$$

Next, notice that if $d > 0$, then $\alpha > 0$ if and only if (14) is positive. Hence, if the following holds

$$\alpha \geqslant 1-\frac{1}{\sqrt{2\lambda}}. \tag{18}$$

Now, maximizing (17) with respect to α subject to (18), one gets

$$\frac{1}{2(1-\alpha)^{2}}-\lambda+\eta=0, \tag{19}$$

where $\eta \geqslant 0$ is the Lagrangian multiplier associated to (18). If $\eta > 0$ then (18) binds and (19) implies that $\eta = 0$: a contradiction. Moreover, if $\eta = 0$ and $\alpha > 0$, then the condition (19) implies that $\alpha = 1 - \frac{1}{\sqrt{2\lambda}}$. Hence, $\alpha = 0$ whenever $d > 0$.

Finally, to show the result we just need to sign the following difference

$$\mathcal{V}(\alpha=1,d=0)-\mathcal{V}(\alpha=0,d=\mu-1)=\mu+\frac{\sigma^{2}}{2}-\lambda-\frac{1}{2},$$

which yields immediately the result. In fact,

$$\mathcal{V}(\alpha=1,d=0)\geqslant\mathcal{V}(\alpha=0,d=\mu-1)\quad\Leftrightarrow\quad\sigma^{2}\geqslant\sigma_{0}^{2}\triangleq 1-2(\mu-\lambda),$$

where it is immediate to verify that $\sigma_0^2 > 0$ under Assumption **A3**. ■

Proof of Proposition 3. Optimizing pointwisely the objective function in (6) with respect to $\alpha(\cdot)$, we obtain

$$\frac{[\theta - d(\theta)]^2}{2} - (x + \lambda) + \theta + h(\theta), \tag{20}$$

while optimizing with respect to $d(\cdot)$, we have

$$(1 - \alpha(\theta))(\theta - d(\theta)) - 1. \tag{21}$$

Again, if $\alpha(\theta) = 1$, then (21) is negative, so that $d(\theta) = 0$. Hence, $\alpha(\theta) = 1$ is admissible if and only if

$$\frac{\theta^2}{2} - (x + \lambda) + \theta + h(\theta) \geqslant 0, \ \theta \in \left[\underline{\theta}, \overline{\theta}\right], \tag{22}$$

which is true as long as $\underline{\theta} \geqslant \sqrt{2(x + \lambda) + 1} - 1$.

In what follows, we show that if $d(\theta) > 0$, then $\alpha(\theta) = 0$. Suppose, by contradiction, that this is not the case. Notice that for given $\alpha(\theta) < 1$, the objective function is strictly concave in $d(\theta)$. Hence, if $d(\theta) > 0$, strict concavity of the objective function implies that it must be equal to

$$\widehat{d}(\theta) \triangleq \theta - \frac{1}{1 - \alpha(\theta)}.$$

Substituting for $\widehat{d}(\theta)$ into the objective function, we have

$$\int_\theta \left\{\mathcal{V}(\alpha(\theta), \widehat{d}(\theta), \theta) - \alpha(\theta) x - (1 - \alpha(\theta))[\theta + h(\theta)]\right\} dF(\theta). \tag{23}$$

Next, we notice that if $d(\theta) > 0$, then $\alpha(\theta) > 0$ if and only if (20) is positive — i.e., if

$$\alpha(\theta) \leqslant 1 - \frac{1}{\sqrt{2(x + \lambda - (\theta + h(\theta)))}}, \quad \forall \theta \in \Theta. \tag{24}$$

Maximizing pointwisely (23) with respect to $\alpha(\theta)$ subject to (24), we have

$$\frac{1}{2(1 - \alpha(\theta))^2} - \lambda - x + \theta + h(\theta) - \eta(\theta) = 0. \tag{25}$$

where $\eta(\theta) \geqslant 0$ is the Lagrangian multiplier associated to (24). If $\eta(\theta) > 0$ then (24) is binding and (25) yields $\eta(\theta) = 0$: a contradiction. Moreover, if $\eta(\theta) = 0$ for some θ and $\alpha(\theta) > 0$, then the condition (19) implies that

$$\alpha(\theta) = 1 - \frac{1}{\sqrt{2(x + \lambda - (\theta + h(\theta)))}},$$

which is again a contradiction. Hence, $\alpha(\theta)=0$ whenever $d(\theta)>0$.

Finally, to show the result we need to sign the following difference

$$\mathcal{V}(\alpha(\theta)=1, d(\theta)=0, \theta)-\mathcal{V}(\alpha(\theta)=0, d(\theta)=\theta-1, \theta)=-(x+\lambda)-\left[\frac{1}{2}-(2\theta+h(\theta))\right],$$

which yields $\alpha(\theta)=1$ and $d(\theta)=0$ if and only if

$$2\theta+h(\theta)\geqslant\frac{1}{2}+x+\lambda. \tag{26}$$

Notice that as admits a solution $\theta^*\in\left(\underline{\theta},\overline{\theta}\right)$ if and only if

$$2\underline{\theta}<\frac{1}{2}+x+\lambda \quad\Rightarrow\quad x\geqslant\underline{x}\triangleq 2\underline{\theta}-\frac{1}{2}-\lambda,$$

which is always positive by Assumption **A3**; and

$$2\overline{\theta}+h\left(\overline{\theta}\right)\geqslant\frac{1}{2}+x+\lambda \quad\Rightarrow\quad x\geqslant\overline{x}\triangleq\underline{x}+2\Delta\theta+h\left(\overline{\theta}\right)>0.$$

Hence, the optimal policy is such that: *(i)* $\alpha(\theta)=1$ and $d(\theta)=0$ for every $\theta\in\Theta$ when $x\leqslant\underline{x}$; *(ii)* $\alpha(\theta)=1$ and $d(\theta)=0$ if and only if $\theta\geqslant\theta^*\in int\Theta$, while $\alpha(\theta)=0$ and $d(\theta)=\theta-1$ otherwise, for $x\in(\underline{x},\overline{x})$; *(iii)* $\alpha(\theta)=0$ and $d(\theta)=\theta-1$ for every $\theta\in\Theta$ when $x\geqslant\overline{x}$. ■

Proof of Corollary 4. The proof of this result is straightforward and will be omitted.

Proof of Proposition 5. To begin with notice that, with information acquisition, the country's expected payoff writes as

$$\mathcal{V}^I\triangleq-\int_{\underline{\theta}}^{\theta^*}\left\{\theta-\frac{1}{2}+\theta^*\right\}dF(\theta)-\int_{\theta^*}^{\overline{\theta}}[\lambda+x]\,dF(\theta).$$

Simple integration by parts imply

$$\int_{\underline{\theta}}^{\theta^*}\theta dF(\theta)=\theta^*F(\theta^*)-\int_{\underline{\theta}}^{\theta^*}F(\theta)\,d\theta. \tag{27}$$

Hence, using (27) together with the definition of θ^*, simple algebraic manipulations imply

$$\mathcal{V}^I=\int_{\underline{\theta}}^{\theta^*}F(\theta)\,d\theta+F(\theta^*)\,h(\theta^*)-(\lambda+x).$$

Differentiating with respect to θ^* we have

$$\frac{\partial\mathcal{V}^I}{\partial\theta^*}=2F(\theta^*)+F(\theta^*)\,\dot{h}(\theta^*),$$

which immediately proves the result since $\dot{h}(\theta^*) > 0$ by assumption **A2**. ■

Proof of Proposition 6. Suppose first that $\sigma^2 \geqslant \sigma_0^2$, so that the uninformed country always chooses to strike. Hence,

$$\mathcal{V}^I - \mathcal{V}^N = \int_{\underline{\theta}}^{\theta^*} F(\theta)\, d\theta + F(\theta^*)\, h(\theta^*) - x.$$

Notice that $\mathcal{V}^I - \mathcal{V}^N = -\underline{x}$ when $x \leqslant \underline{x}$ so that $\theta^* = \underline{\theta}$. Instead,

$$\mathcal{V}^I\big|_{\theta^*=\overline{\theta}} - \mathcal{V}^N = \int_{\underline{\theta}}^{\overline{\theta}} F(\theta)\, d\theta + h\left(\overline{\theta}\right) - \overline{x}.$$

for $x \geqslant \overline{x}$ since $\theta^* = \overline{\theta}$. Integrating by parts, this yields

$$\mathcal{V}^I\big|_{\theta^*=\overline{\theta}} - \mathcal{V}^N = \overline{\theta} F\left(\overline{\theta}\right) - \int_{\underline{\theta}}^{\overline{\theta}} \theta dF(\theta) + h\left(\overline{\theta}\right) - \overline{x} = -\overline{\theta} - \mu + \frac{1}{2} + \lambda,$$

which is positive if and only if $2\overline{\theta} \leqslant \sigma_0^2$. In addition, notice that $\mathcal{V}^I\big|_{\theta^*=\overline{\theta}} - \mathcal{V}^N$ is increasing in θ^*. Therefore, if $2\overline{\theta} \leqslant \sigma_0^2$, then there exists a unique x^* that solves

$$\int_{\underline{\theta}}^{\theta^*} F(\theta)\, d\theta + F(\theta^*)\, h(\theta^*) - x = 0,$$

and is such that $\mathcal{V}^I \geqslant \mathcal{V}^N$ as long as $x \geqslant x^*$. Clearly, if $2\overline{\theta} > \sigma_0^2$ then $\mathcal{V}^I\big|_{\theta^*=\overline{\theta}} < \mathcal{V}^N$ regardless of x.

Next, suppose that $\sigma^2 < \sigma_0^2$. In this case,

$$\mathcal{V}^I - \mathcal{V}^N = \int_{\underline{\theta}}^{\theta^*} F(\theta)\, d\theta + F(\theta^*)\, h(\theta^*) - (x + \lambda) - \frac{1}{2} + \mu + \frac{\sigma^2}{2},$$

rearranging we have

$$\mathcal{V}^I \geqslant \mathcal{V}^N \quad \Leftrightarrow \quad \sigma^2 \geqslant \sigma_0^2 - 2\left[\int_{\underline{\theta}}^{\theta^*} F(\theta)\, d\theta + F(\theta^*)\, h(\theta^*) - x\right]. \tag{28}$$

Hence, for $\sigma^2 < \sigma_0^2$, we have $\mathcal{V}^I \geqslant \mathcal{V}^N$ only if $2\overline{\theta} > \sigma_0^2$ and $x \geqslant x^*$, and if (28) holds; $\mathcal{V}^I < \mathcal{V}^N$ otherwise. ■

References

[1] Berman E., Shapiro J.N., Felter J.H. 2011, *Can Hearts and Minds Be Bought? The Economics of Counterinsurgency in Iraq*, Journal of Political Economy, **119**(4), 766-819.

[2] Cassidy, R. 2008, *Counterinsurgency and the Global War on Terror*, Westport, CT: Praeger.

[3] Enders W., Sandler T. 1995, *Terrorism: Theory and Applications*, Handbook of Defense Economics, **1**, 213-249.

[4] Fridovich, D.P., Krawchuk F. 2007, *Winning in the Pacific: The Special Operations Forces Indirect Approach,* Joint Force Quarterly, **44**, 24–27.

[5] Galula, D. 1964, *Counterinsurgency Warfare: Theory and Practice*, New York, Praeger.

[6] Hastings J.V., Chan R.J. 2013, *Target Hardening and Terrorist Signaling: The Case of Aviation Security*, Terrorism and Political Violence, **25**(5), 777-797.

[7] Kalyvas S. 2006, *The Logic of Violence in Civil War*, Cambridge University Press.

[8] Laffont J.J., Martimort D. 2002, *The Theory of Incentives*, Princeton University Press.

[9] Laffont J.J., Tirole J. 1993, *A Theory of Incentives in Procurement and Regulation*, MIT Press.

[10] McMaster H.R. 2008, *On War: Lessons to Be Learned*, Survival, **50**, 19-30.

[11] Nagl, J. 2002, *Counterinsurgency Lessons from Malaya and Vietnam: Learning to Eat Soup with a Knife*, Westport, CT: Praeger.

[12] Petersen, R. 2001, *Resistance and Rebellion: Lessons from Eastern Europe*, Studies in Rationality and Social Change, Cambridge, Cambridge Univ. Press.

[13] Petraeus, D. 2006, *Learning Counterinsurgency: Observations from Soldiering in Iraq*, Military Rev. (January–February), 45–55.

[14] Popkin S.L. 1979, *The Rational Peasant: The Political Economy of Rural Society in Vietnam*, Berkeley, University of California Press.

[15] Powell R. 2006, *War as a Commitment Problem*, International Organization, **60**, 169-203.

[16] Sandler T. 2015, *Terrorism and Counterterrorism: an Overview*, Oxford Economic Papers, **67**(1), 1-20.

[17] Sepp K. 2005, *Best Practices in Counterinsurgency*, Military Review, 8-12.

[18] SCHNEIDER F., BRÜCK T., MEIERRIEKS D. 2015, *The Economics of Counterterrorism: A Survey*, Journal of Economic Surveys, **29**(1), 131-157.

[19] WANG C., BIER V.M. 2011, *Target-hardening Decisions Based on Uncertain Multiattribute Terrorist Utility*, Decision Analysis, **8**(4), 286-302.