DECISION ANALYSIS
Vol. 11, No. 1, March 2014, pp. 43–62
ISSN 1545-8490 (print) | ISSN 1545-8504 (online)
http://dx.doi.org/10.1287/deca.2013.0288

Modeling Credible Retaliation Threats in Deterring the Smuggling of Nuclear Weapons Using Partial Inspection—A Three-Stage Game

Xiaojun (Gene) Shan, Jun Zhuang
Department of Industrial and Systems Engineering, University at Buffalo, The State University of New York, Buffalo, New York 14260 {xshan@buffalo.edu, jzhuang@buffalo.edu}

Deterring the smuggling of nuclear weapons in container freight is critical. Previous work has suggested that such deterrence could be achieved by retaliation threats and partial inspection. However, pre-event declared retaliation threats may not be credible, causing the desired deterrence not to be achieved. In this paper, we extend and complement the work of Haphuriwat et al. (2011) to model credible retaliation threats in a three-stage game, by introducing two additional decision variables and five additional parameters. Our results suggest that noncredible retaliation could be at equilibrium when the reputation loss is low, the reward from the public for retaliation is low, or the costs of retaliation are high. When the declared retaliations are noncredible, we quantitatively show that a higher inspection level would be required to deter nuclear smuggling than would be needed if retaliation threats are always credible. This paper provides additional quantitative insights on the decision-making process for container screening to deter nuclear smuggling.

Keywords: noncredible retaliation threats; credible retaliation threats; deterrence; game theory; smuggling; terrorism
History: Received on January 5, 2012. Accepted by former Editor-in-Chief L. Robin Keller on October 6, 2013, after 3 revisions.

1. Introduction

Recently, Haphuriwat et al. (2011) studied deterrence of nuclear smuggling through both retaliation threats and partial inspection. This topic is appealing since retaliation has not been extensively studied in the literature of defender–attacker games. However, Haphuriwat et al. (2011) did not consider the costs of retaliation and assume that the declared retaliation is credible (in stage 1 of their two-stage game; see Figure 1 for the detailed game tree) and committed after a smuggling attempt (in stage 2 of their game). This assumption might not always hold, especially when the (material and political) costs of retaliation are extremely high. The problem of making credible retaliation threats in deterring nuclear smugglers is important because if the declared threat is not credible, partial inspection without considering retaliation credibility may not achieve deterrent effects against nuclear smuggling despite declared threats; that is, nuclear smugglers who are threatened with retaliation from a declared threat may not be deterred by partial inspection if they think the threat is not credible. Therefore, it is critical to evaluate the circumstances under which the defender's retaliation threat is credible and how the associated players' optimal decisions might change as a result.

To our knowledge, few researchers have studied the credibility of retaliation threats in smuggling games involving nuclear weapons. One notable exception is Schelling (1984, Chapter 14), which studies nuclear terrorism and deterrence. This paper contributes to the literature of deterrence against nuclear smuggling by studying the required conditions for both credible and noncredible retaliation threats, and quantifying the associated optimal levels of container inspection. In particular, we extend and complement the work of Haphuriwat et al. (2011) by modeling credible retaliation threats in a three-stage game (where the defender retaliate in stage 3), by introducing two additional decision variables (whether to retaliate in stage 3 after either a successful or foiled smuggling attempt) and five additional parameters (costs and rewards of retaliation and declaration cost). They are introduced and discussed in detail in §2 along with detailed game trees from both Haphuriwat et al. (2011) and the current paper.

Research on deterrence can be traced back to the Cold War. The credibility of retaliation for the period of the Cold War has been extensively examined in the game-theoretic literature (Schelling 1984, 2008). The classic entrant game for a monopoly market is a typical example of noncredible threats (Kreps 1990). Dixit and Nalebuff (1991) discuss credible threats as containing two components: a plan of sequenced actions and the commitment to make the threat credible. These authors use game-theoretic reasoning to reach important insights without formal mathematical analysis. However, mathematical formalism could potentially lead to more quantitative insights.

The effectiveness of the post-Cold War deterrence policy against nuclear threats is under debate. Goldstein (2000) focuses on international relationships and argues that nuclear deterrence is still a key component of security policies of global superpowers and an alternative for less powerful countries who have concerns about more capable adversaries. However, unlike the Cold War, when most critical infrastructures in both the United States and the Soviet Union had nowhere to hide after launching a nuclear attack, terrorist groups such as Al Qaeda have no such concerns (Allison 2005). By contrast, a number of researchers argue that many terrorist groups might be deterred by properly designed deterrence policies (Trager and Zagorcheva 2006) due to the terrorists' strategic nature. (Marshall 2012, p. 821) states: "It's now clear that many terrorists are well-educated and seemingly rational." National Commission on Terrorist Attacks upon the United States (2004, p. 383) concludes: "Our report shows that the terrorists analyze defenses. They plan accordingly." Moreover, Haller (2013, p. vii) states: "…current adversaries may be deterred from the use of nuclear weapons differently than were Cold War adversaries." Therefore, those authors suggest that at least some terrorists seem strategic and might be deterred. Note that nuclear smuggling and attack need special attention given their potentially severe consequences.

Although some terrorists, such as suicide bombers, appear irrational and willing to die, they are probably willing to sacrifice their lives for a greater cause and thus still make strategic decisions. For example, (Knopf 2010, p. 13) discusses: "Colin Gray points out that, despite the grandiosity of its objectives, Al Qaeda 'functions strategically' in trying to use its suicide attacks to advance those objectives. As a result, 'It can be deterred by the fact and expectation of strategic failure.' " Moreover, even though operating personnel of terrorist activities are willing to die, terrorist group leaders may be deterred by strategically designed retaliation policies. When deciding whether to attempt to smuggle nuclear weapons into the United States, the smuggler would probably consider the possibility of failure as well as the potentially high penalty cost.

On the other side, the United States does retaliate against terrorist attacks. For example, on October 7, 2001 (about four weeks after the September 11 terrorist attacks), the United States sent troops to Afghanistan to attack Al Qaeda camps supported by the Taliban regime of Afghanistan. "Bush said the action was taken after the Taliban refused to meet several non-negotiable American demands. . . . 'None of these demands was met, and now, the Taliban will pay a price,' he said" (CNN 2001). Moreover, the constant hunting and final killing of Al Qaeda's leader Osama Bin Laden on May 1, 2011, is another example of retaliation (Miller 2011, Helene 2011). We envision that any nuclear terrorist attacks on U.S. soil would lead to severe retaliation efforts by the United States.

A properly designed and declared retaliation policy could inflict a potentially high penalty cost on a terrorist group that is seeking a rogue state sponsor. There are two main ways for terrorists to obtain nuclear materials or weapons from a state sponsor: (a) a rogue state willingly gives them to the terrorist group or (b) the terrorist group steals them from a nation that lacks secure methods protecting nuclear materials or weapons. Levi (2004) suggests a declared retaliation policy against the nation where the nuclear weapons originate. With the rapid advance in technology, traces of any debris from a nuclear attack could potentially reveal its origin. Therefore, out of the fear of potentially severe retaliation, the rogue state would make the cost of a terrorist group seeking nuclear sponsorship extremely high (such as increasing terrorist operating costs by interdicting terrorist activities).

Knopf (2010) discusses three categories of deterrent approaches against terrorists, which are related to the assumption about nuclear smugglers' behavior in the current paper. The first category is referred to as "indirect deterrence," where the retaliation targets are not terrorists per se, but the terrorists' state sponsor or financier. The second category is "deterrence by denial," which translates to inspecting every incoming container in border security. The full inspection level is neither necessary nor economical, as discussed in Wang and Zhuang (2011). That serves as one of the motivations of this paper, that is, to study how to deter the smuggling of nuclear weapons by partial inspection and credible retaliation threats and what would happen when credibility fails. In §3, we show that in some cases, inspecting all containers is a requirement to deter smuggling attempts, which is similar to "deterrence by denial." The third category is "punishing terrorists, by threatening societal targets." The key with the third category is to identify appropriate societal targets to effectively deter terrorists (Knopf 2010).

Decision analytical models have made significant contributions to homeland security research. Within the context of border security (e.g., container inspection), Morton et al. (2007) study where to locate radiation sensors to interdict the smuggling of nuclear materials. Bakır (2008) uses decision trees to evaluate countermeasures in cargo security. Merrick and McLay (2010) study screening as a deterrence method using a non-game-theoretic model. Bier and Haphuriwat (2011) and Wang and Zhuang (2011) suggest that deterrence could be achieved by partial inspection. Besides border security, a portion of the limited amount of defensive resources against an adaptive adversary is dedicated to target hardening (Zhuang and Bier 2007). Choosing optimal levels of defense for a given target requires the understanding of the adversary's objective, which could contain multiple attributes (Keeney 2007, Keeney and von Winterfeldt 2011, Wang and Bier 2011). Robust (e.g., Nikoofal and Zhuang 2012), multiple-stage (where only one player moves in each stage; e.g., Brown et al. 2006, Hausken and Zhuang 2011b), and multiple-period (where both players move in each period; e.g., Hausken and Zhuang 2011a, Zhuang et al. 2010, Jose and Zhuang 2013) games between attackers and defenders have been studied.

The remainder of this paper is structured as follows: §2 extends the game-theoretic model in Haphuriwat et al. (2011) and compares the equilibrium strategies in the two models. Section 3 illustrates the model with the real data that are used in Haphuriwat et al. (2011). Section 4 concludes and provides some future research directions. The appendix provides the solution procedure to the model, additional solution explanations, the proofs for the propositions, and additional numerical illustrations.

2. Extending the Model in Haphuriwat et al. (2011)

2.1. Notation, Definition of Credible Retaliation, and Assumptions

Haphuriwat et al. (2011) develop a two-stage nuclear smuggling game, where the defender decides on whether to declare retaliation and how many containers to inspect in stage 1, and the smuggler decides on whether to smuggle a nuclear weapon in stage 2 (see Figure 1(a) for a detailed game tree). We extend the model in Haphuriwat et al. (2011) to a three-stage game by introducing two additional defender's decision variables (whether to retaliate in stage 3, $D'(s)$, where s is an indicator variable for the outcome of a smuggling attempt and equals 1 if successful or 0 if foiled) and five additional parameters (reward/cost of retaliation, $c(D_1, s, D_3)$, where $s = 1$ or 0 and D_1 and D_3 are the defender's stage 1 and 3 declaration or retaliation decisions, respectively, and retaliation declaration cost e).

In particular, the defender's stage 3 decision is denoted by $D'(s)$ (where $s = 1$ or 0) when the defender decides to retaliate after either a successful ($s = 1$) or foiled ($s = 0$) attempt (see Figure 1(b) for a detailed game tree and see the three-stage game described in §2.3). Reward/cost of retaliation is represented by $c(1, 1, 1)$, $c(1, 1, 0)$, $c(1, 0, 1)$, and $c(1, 0, 0)$. For example, we have $c(1, 1, 0)$ as the cost of not retaliating in stage 3 after the defender declares retaliation in stage 1 and the attempted smuggling was successful. All notation (decision variables, parameters, variables, thresholds, conditions, and required inspection levels) are listed and explained in Table 1, including those that are used

Table 1 Notation (Used in Haphuriwat et al. 2011 or Introduced in This Paper)

	New variable and parameters
s	Indicator variable: 1 if the smuggling attempt is successful, 0 otherwise
$c(D_1, s, D_3) < 0$	Reward/cost to the defender if D_1 and D_3 is chosen for s
$e > 0$	Cost to the defender for declaring retaliation (including costs associated with deterrence research on the design and declaration of an effective retaliation policy)
	Decision variables in Haphuriwat et al. (2011)
n	Number of containers to be inspected by the defender
D_1	$\equiv D(s)$, defender's stage 1 decision: 1 if the defender declares to retaliate after s, 0 otherwise
$A(n, D_1, D_3)$	1 if the smuggler decides to smuggle nuclear bombs into the U.S. in the face of the defensive policy (n, D_1), reasoning about the threat credibility, 0 otherwise
	New decision variables
D_3	$\equiv D'(s) = D_3(D_1, s)$, defender's stage 3 decision: 1 if the defender retaliates after s, 0 otherwise. Note that the reward/cost of retaliation is a function of D_1.
	Parameters in Haphuriwat et al. (2011)
$N > 0$	Total number of containers
$m > 0$	Number of nuclear bombs involved in a single smuggling attempt; we assume that $m = 1$ throughout the paper except in §2.5, and §A.7 of the appendix, where we compare the solution with that in Haphuriwat et al. (2011)
$v > 0$	Expected damage if at least one nuclear weapon is successfully smuggled into the U.S.
$d > 0$	Cost to the defender of inspecting a container
$a > 0$	Cost of acquiring and smuggling a nuclear weapon
$\alpha > 0$	Parameter reflecting economies of scale in acquiring multiple nuclear weapons; we assume that $\alpha = 1$ throughout the paper except in §2.5, where we compare the solution to the extended model with that in Haphuriwat et al. (2011)
$p \in [0, 1]$	Conditional probability of successfully detecting a nuclear bomb, given inspection of a container that contains a bomb
$r(s) > 0$	Cost to the defender of retaliation after s
$k_A(s) \geq 0$	Cost to the smuggler of retaliation after s
	New thresholds and conditions
$T_1 > 0$	$\equiv [(1 - p^m)(v - k_A(1)) - p^m k_A(0)]/m$; threshold 1 for a in Figures 2–3
$T_2 > 0$	$\equiv (v - k_A(1))/m$; threshold 2 for a in Figures 2–3
$T_3 > 0$	$\equiv v/m$; threshold 3 for a in Figure 2
$C_1 > 0$	$\equiv \left\{d > p^m \frac{v}{N}\right\}$; condition 1 in Figure 2
$C_2 > 0$	$\equiv \left\{d < \frac{p(v-e)}{N}\left(1 - \frac{ma + k_A(0)}{v - k_A(1) + k_A(0)}\right)^{-1/m}\right\}$; condition 2 in Figure 2
$C_3 > 0$	$\equiv \left\{d > \frac{pe}{N}\left(1 - \frac{ma}{v}\right)^{-1/m}\right\}$; condition 3 in Figure 2
$C_4 > 0$	$\equiv \left\{\frac{dN}{p}\sqrt[m]{1 - \frac{ma}{v}} < \min\left\{\frac{dN}{p}\sqrt[m]{1 - \frac{ma}{v - k_A(1)}} + e, v\right\}\right\}$; condition 4 in Figure 2
$C_5 > 0$	$\equiv \left\{\frac{dN}{p}\sqrt[m]{1 - \frac{ma}{v - k_A(1)}} + e < \min\left\{\frac{dN}{p}\sqrt[m]{1 - \frac{ma}{v}}, v\right\}\right\}$; condition 5 in Figure 2
$C_6 > 0$	$\equiv \left\{\frac{dN}{p}\sqrt[m]{1 - \frac{ma + k_A(0)}{v + k_A(0)}} + e < \frac{dN}{p}\sqrt[m]{1 - \frac{ma}{v}}\right\}$; condition 6 in Figure 2
$C_7 > 0$	$\equiv \left\{d > \frac{pv}{N}\left(1 - \frac{ma}{v}\right)^{-1/m}\right\}$; condition 7 in Figure 2

Table 1 (Continued)

	New required inspection level
$n_1 > 0$	$\equiv \frac{N}{p}\sqrt[m]{1-\frac{ma+k_A(0)}{v-k_A(1)+k_A(0)}}$; threshold 1 for n in Figures 2–3
$n_2 > 0$	$\equiv \frac{N}{p}\sqrt[m]{1-\frac{ma}{v}}$; threshold 2 for n in Figure 2
$n_3 > 0$	$\equiv \frac{N}{p}\sqrt[m]{1-\frac{ma}{v-k_A(1)}}$; threshold 3 for n in Figure 2
$n_4 > 0$	$\equiv \frac{N}{p}\sqrt[m]{1-\frac{ma+k_A(0)}{v+k_A(0)}}$; threshold 4 for n in Figure 2

in Haphuriwat et al. (2011)[1] and those that are newly introduced in this paper.

We acknowledge that the effectiveness of retaliation threats in deterring the smuggling of nuclear weapons is controversial, as discussed in §1. In this paper, we assume that retaliation against the smuggler itself or the state supplier of the nuclear weapons would cause a significantly high penalty cost in the smuggler's utility function (assuming retaliation against the state supplier of the nuclear weapons would indirectly pose a heavy toll on the smuggler, such as difficulty in terrorist operations causing a significantly increased operating cost). Furthermore, retaliation against terrorist group leaders might have such an effect as well.

An alternative to a three-stage game, as studied in this paper, is that when declaring retaliation threats, the defender binds herself to retaliate using mechanisms such as a trip wire in stage 1, and thus does not have the option not to retaliate, which eliminates stage 3 of the game. Note that the modeling framework developed in this paper is general and is able to account for the possibility of binding oneself to respond. Mechanisms such as a trip wire could still be compatible with this paper in the sense of being modeled as sufficiently high cost of breaking the promise to retaliate (i.e., high levels of $c(1,1,0)$ and $c(1,0,0)$) so that making noncredible threats is never part of the defender's equilibrium strategy.

[1] The notation used in Haphuriwat et al. (2011) has been modified. In particular, we have $D(1)$, $D(0)$, $D'(1)$, and $D'(0)$ instead of D_S, D_F, D'_S, and D'_F; $r(1)$, $r(0)$, $k_A(1)$, and $k_A(0)$ instead of S_d, F_d, S_a, and F_a; and m, v, d, and a instead of M, V, C_d, and C_a, respectively.

Definition 1. A retaliation threat is credible if one of the following three conditions is satisfied at equilibrium:

(a) $D(1)=D'(1)=D(0)=D'(0)=1$;
(b) $D(1)=D'(1)=1$ and $D(0)=D'(0)=0$;
(c) $D(1)=D'(1)=0$ and $D(0)=D'(0)=1$.

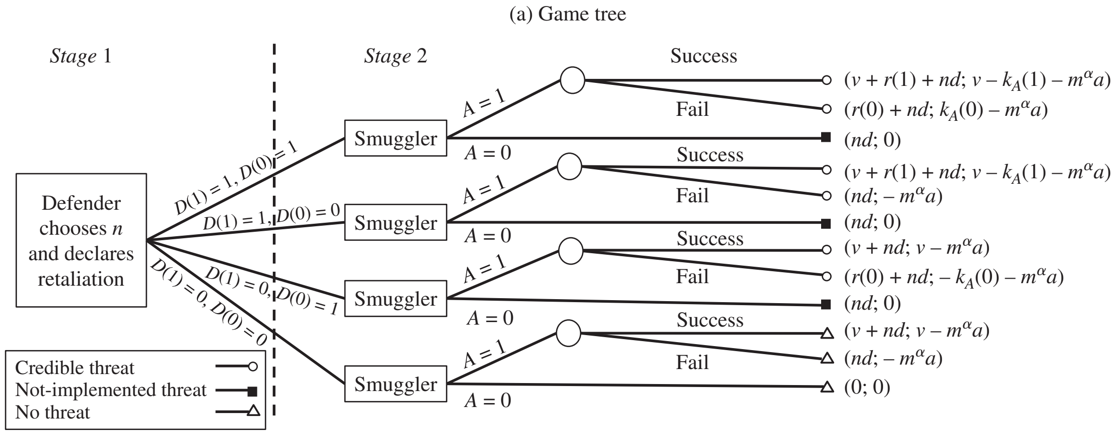
Remark 1. Case (a) corresponds to credible retaliation threats against both a successful and a foiled smuggling attempt. Case (b) corresponds to a credible threat against a successful smuggling attempt, but no threat against a foiled smuggling attempt. Case (c) corresponds to a credible threat against a foiled smuggling attempt, but no threat against a successful smuggling attempt. If $D(1)=D'(1)=D(0)=D'(0)=0$, the defender does not declare retaliation and does not retaliate against a successful or foiled smuggling attempt. If $D(1)\neq D'(1)$ or $D(0)\neq D'(0)$, the retaliation threat is not credible.

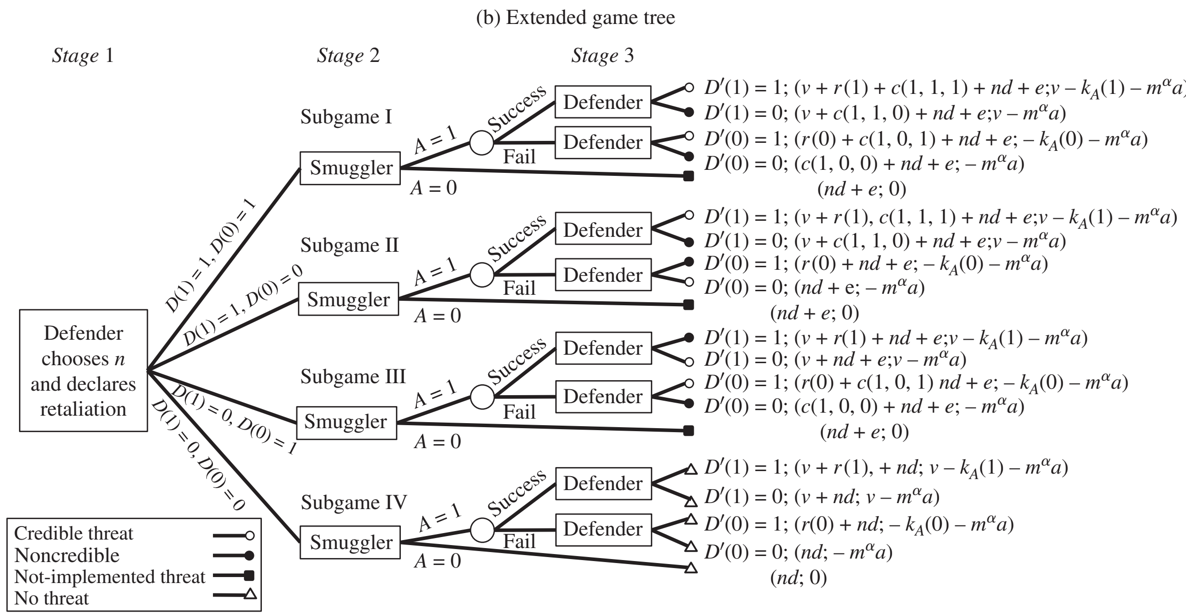
2.2. Comparing the Game Trees

Figure 1(a) depicts the game tree that is implicitly used in Haphuriwat et al. (2011), where declared retaliation is assumed to be carried out. Figure 1(b) presents an extended game tree modeling credible threats, where the defender first declares whether to retaliate and chooses an inspection level, the smuggler then decides whether to smuggle nuclear weapons, and, finally, the defender decides whether to actually retaliate.

There are two main differences between the two trees: First, the tree in Figure 1(a) depicts a two-stage game with 12 terminal nodes where the defender decides in stage 1 only, whereas the tree in Figure 1(b) depicts a three-stage game with 20 terminal nodes

Figure 1 **Comparison Between (a) The Game Tree Implicitly Used in Haphuriwat et al. (2011), and (b) The Extended Game Tree Modeling Credible Threats Introduced in This Paper**





Note. The expected costs for the defender and expected payoffs for the smuggler are listed in the order of (Defender; Smuggler).

where the defender decides in both stages 1 and 3. Second, retaliation threats are assumed to be credible in the game tree from Figure 1(a), and thus there are six terminal nodes of credible threat ("∘"), three terminal nodes of not-implemented threat ("■"; no smuggling attempt follows the declared retaliation, and therefore the declared retaliation is not implemented), and three terminal nodes of no threat ("△"). In contrast, in Figure 1(b), there are six terminal nodes of credible threat ("∘"), six terminal nodes of noncredible threat ("•"), three terminal nodes of not-implemented threats ("■"), and five terminal nodes of no threat ("△").

2.3. The Extended Model

In this section, we describe the extended model as illustrated in Figure 1(b). There are three stages of decision making. In stage 1, the defender decides whether to declare retaliation against a successful and/or foiled smuggling attempt and how many containers to inspect. In stage 2, the smuggler decides whether to attempt smuggling nuclear weapons. Finally, in stage 3, the defender decides whether to launch actual retaliation if either a successful or a foiled smuggling attempt happens. Note that we focus on the case where $m=1$ throughout this paper, except in §2.5 and §A.7 of the appendix, where we compare the solution to the extended model with that in Haphuriwat et al. (2011).

We assume that the defender's objective in stage 1 is to minimize the total expected losses, including loss from a successful smuggling attempt, the retaliation cost (either after a successful smuggling attempt or a foiled smuggling attempt), the reward or reputation loss (cost) from keeping or breaking the promise about retaliation, the inspection cost, and the cost of retaliation declaration:

$$\text{Stage 1}\quad \min_{D(1)\in\{0,1\},\,D(0)\in\{0,1\},\,n\in\{1,\dots,N\}}\Bigg\{\Bigg[\bigg(1-\Big(\frac{np}{N}\Big)^{m}\bigg)\cdot\underbrace{(v+r(1)\hat{D}'(1)+c(D(1),1,\hat{D}'(1)))}_{\text{[loss+retaliation cost+reward or reputation loss]}}+\Big(\frac{np}{N}\Big)^{m}\underbrace{(r(0)\hat{D}'(0)+c(D(0),0,\hat{D}'(0)))}_{\text{[retaliation cost+reward or reputation loss]}}\Bigg]\hat{A}(n,D_1,D_3)+\underbrace{nd}_{\text{[inspection cost]}}+\underbrace{e[1-(1-D(1))(1-D(0))]}_{\text{[declaration cost]}}\Bigg\}.$$

In stage 2, the smuggler chooses whether to smuggle (using decision variable $A\in\{0,1\}$). This decision is contingent upon the smuggler's reasoning about the defender's decision in stage 3, to maximize the expected loss minus possible retaliation cost and smuggling cost; that is, if $(1-((np)/N)^m)(v-k_A(1)D_3(D_1,1))-((np)/N)^m k_A(0)D_3(D_1,0)-ma>0$, the smuggler will smuggle; otherwise, the smuggler will not smuggle. Mathematically, we have:

$$\text{Stage 2}\quad \max_{A\in\{0,1\}}\Bigg[\bigg(1-\Big(\frac{np}{N}\Big)^{m}\bigg)\underbrace{(v-k_A(1)D_3(D_1,1))}_{\text{[loss+retaliation cost]}}-\Big(\frac{np}{N}\Big)^{m}\underbrace{k_A(0)D_3(D_1,0)}_{\text{[retaliation cost]}}-\underbrace{ma}_{\text{[smuggling cost]}}\Bigg]A.$$

Note that the smuggler's decision (stage 2) is a function of both what the defender will do (D') and what the defender said that he would do (D), since the optimal defender's response (stage 3) depends upon the cost of responding (retaliation), which then depends upon the defender's previous public declarations (stage 1). Mathematically, we have $A(n,D_1,D_3)=A(n,D_1,D_3(D_1,s))=A(n,D_1)$.

In Haphuriwat et al. (2011), the planned retaliation is fully committed after a smuggling attempt, and thus the smuggler's decision in stage 2 is only a function of the defender's decision in stage 1. In this paper, when the smuggler makes a decision of whether to attempt to smuggle a nuclear weapon, the smuggler takes into consideration the credibility of retaliation threats.

In stage 3, the defender decides whether to retaliate to minimize the retaliation cost minus reward (or plus reputation loss) from keeping (or breaking) the promise about retaliation:

$$\text{Stage 3}\quad \min_{D_3\in\{0,1\}}\Big\{\underbrace{r(s)D_3}_{\text{[retaliation cost]}}+\underbrace{c(D_1,s,D_3)}_{\text{[reward or reputation loss]}}\Big\},$$

where $s=1$ or 0.

2.4. Solution to the Extended Model

We first solve the four subgames in Figure 1(b) separately as documented in §A.1 of the appendix and then solve the entire game using backward induction to get the equilibrium solution, which prescribes the optimal strategies and payoffs for all players (Mas-Colell et al. 1995). In particular, the optimal response of the defender in s3 is as follows: $D_3^*(D_1)=\arg\max_{D_3\in\{0,1\}}[r(s)D_3+c(D_1,s,D_3)]$, where $s=1$ or 0. Recall that when the credibility of the defender's declared retaliation policy (stage 1 decision) is considered, what the smuggler *thinks the response will actually be* (D_3) is also a function of what the defender *said it would be* (D_1).

With backward induction, we use the optimal response of the defender in stage 3 to solve for the optimal decision for the smuggler in stage 2. In particular, we have:

$$\hat{A}(n,D_1)=\arg\max_{A\in\{0,1\}}\Bigg[\bigg(1-\Big(\frac{np}{N}\Big)^{m}\bigg)\underbrace{(v-k_A(1)D_3^*(D_1,1))}_{\text{[loss + retaliation cost]}}-\Big(\frac{np}{N}\Big)^{m}\cdot\underbrace{k_A(0)D_3^*(D_1,0)}_{\text{[retaliation cost]}}-\underbrace{ma}_{\text{[smuggling cost]}}\Bigg]A.$$

According to the values of $r(1)$ and $r(0)$, we divide the parameter space into four regions as shown in

Table 2 Conditions for Credible Threats

Conditions	Region in Figure 2	Types of credible threat
$r(1) < c(1,1,0) - c(1,1,1)$, $r(0) < c(1,0,0) - c(1,0,1)$	Region I	(a), (b), and (c) in Definition 1
$r(1) < c(1,1,0) - c(1,1,1)$, $r(0) \geq c(1,0,0) - c(1,0,1)$	Region II	(b) in Definition 1
$r(1) \geq c(1,1,0) - c(1,1,1)$, $r(0) < c(1,0,0) - c(1,0,1)$	Region III	(c) in Definition 1

Figure 2, which presents the solution to the game tree in Figure 1(b) using thresholds T_1, T_2, T_3, n_1, n_2, n_3, and n_4 and conditions C_1, C_2, C_3, C_4, C_5, C_6, and C_7 as defined in Table 1. In particular, Regions I–IV of Figure 2 correspond to four events: (I) the costs of retaliation against both successful and foiled attempts are less than the lost reputation of not retaliating; (II) the cost of retaliation against a successful attempt is less than the lost reputation of not retaliating, whereas the cost of retaliation against a foiled attempt is equal to or more than the lost reputation of not retaliating; (III) the cost of retaliation against a successful attempt is equal to or more than the lost reputation of not retaliating, whereas the cost of retaliation against a foiled attempt is less than the lost reputation of not retaliating; and (IV) the costs of retaliation against both successful and foiled attempts are equal to or more than the lost reputation of not retaliating. That is, Regions I–IV correspond to (I) $\{r(1) < c(1,1,0) - c(1,1,1)$ and $r(0) < c(1,0,0) - c(1,0,1)\}$, (II) $\{r(1) < c(1,1,0) - c(1,1,1)$ and $r(0) \geq c(1,0,0) - c(1,0,1)\}$, (III) $\{r(1) \geq c(1,1,0) - c(1,1,1)$ and $r(0) < c(1,0,0) - c(1,0,1)\}$, and (IV) $\{r(1) \geq c(1,1,0) - c(1,1,1),\ r(0) \geq c(1,0,0) - c(1,0,1)\}$, respectively.

Proposition 1. *At equilibrium, we have credible threats if one of the conditions in Table* 2 *holds.*

Remark 2. Proposition 1 implies that when the costs of retaliation ($r(s)$) are low, the reputation loss ($c(1,s,0)$) is high, or the reward from the public for retaliation against a smuggling attempt ($c(1,s,1)$) is high (more negative), we have all three types of credible retaliation threats as shown in Region I.2–I.3 of Figure 2 (type (a) in I.2 and I.3; types (a), (b), and (c) in I.3). Since the costs of retaliation to the defender are lower than the reputation loss if the defender breaks the promise to retaliate plus the reward from the public of keeping the promise to retaliate, the defender will always commit to retaliation after either a successful or foiled smuggling attempt. As a result, we have all three types of credible retaliation threats, depending on the defender's decision at s 1 (declaring retaliations against a successful and/or foiled smuggling attempt). When $r(s) < c(1,s,0) - c(1,s,1)$ for either $s = 1$ or 0 is not satisfied, we could have both credible and noncredible retaliation threats at equilibrium, as shown in Regions II.2–II.3 and III.2 of Figure 2. When $r(s) < c(1,s,0) - c(1,s,1)$ for both $s = 1$ and 0 is not satisfied (retaliation is too costly or not worthy as compared to reputation loss from breaking the promise to retaliate or reward from keeping the promise to retaliate), there does not exist any retaliation threat at equilibrium as shown in Region IV.2 of Figure 2.

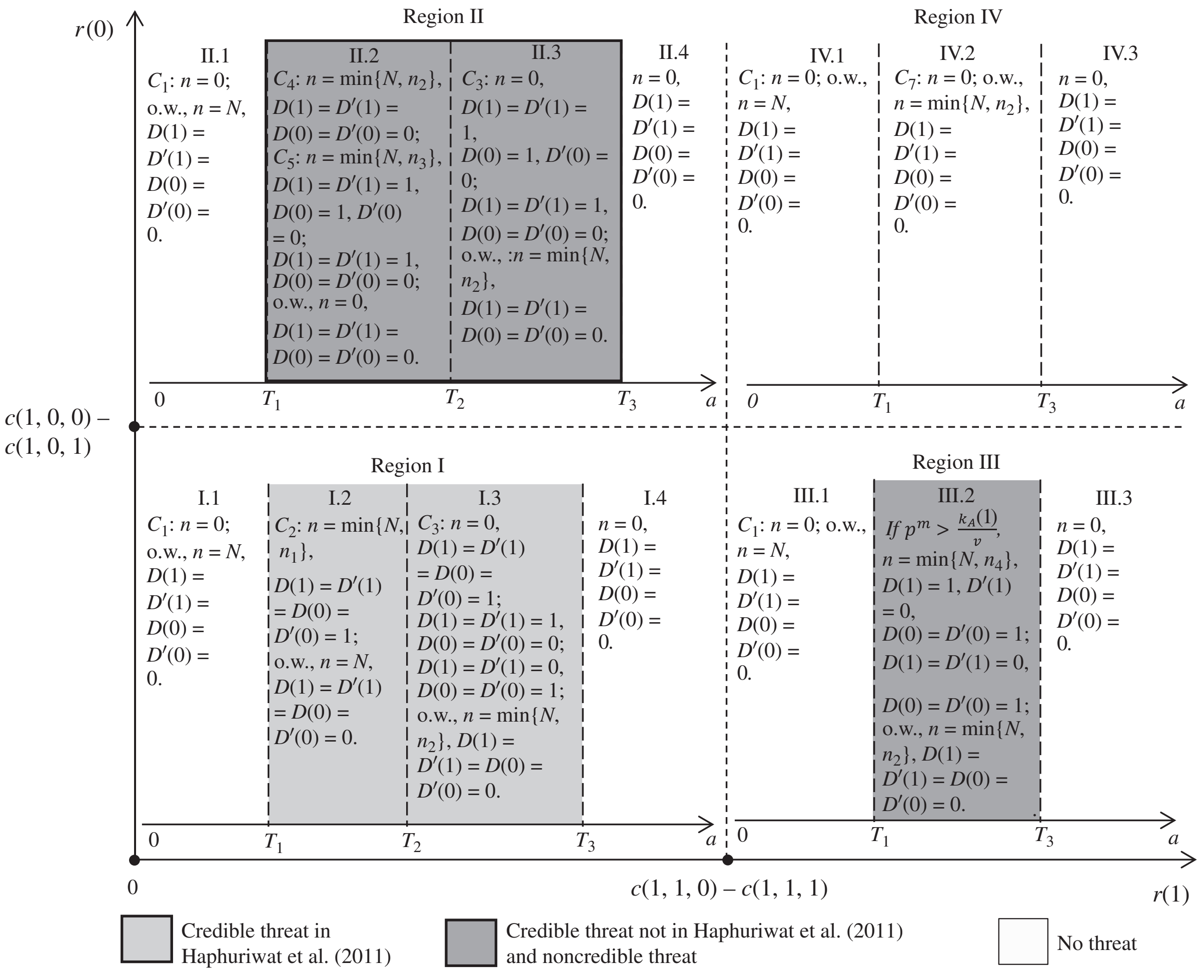
In addition, when the smuggling cost is moderate ($T_1 < a \leq T_2$), a set of strategies of retaliation declaration and inspection would effectively deter the smuggler, as shown in Regions I.2, II.2, and III.2 of Figure 2. In contrast, when the smuggling cost is relatively high ($T_2 < a \leq T_3$), declared retaliation would be sufficiently effective to deter the smuggling attempts, as shown in Regions I.3 and II.3 of Figure 2. Section A.5 of the appendix provides additional explanations to solutions to Regions II–IV in Figure 2.

2.5. Comparing the Solution in Haphuriwat et al. (2011) and That in This Paper

In this section, we compare the equilibrium solution provided in Figure 2 with the solution provided in §3 of Haphuriwat et al. (2011), whose Figure 1 is adapted in Figure 3.[2] In Haphuriwat et al. (2011), there is one condition ($r(s) < c(1,s,0) - c(1,s,1)$ where $s = 1$ or 0) where retaliation threats are credible, corresponding to the left 75% of Region I of Figure 2 and illustrated in Figure 3. In contrast, we have three additional conditions corresponding to Regions II–IV of Figure 2, respectively. In Regions II–IV of Figure 2, we have both credible and noncredible threats at equilibrium, and higher threshold inspection levels n would be employed to deter the smuggler.

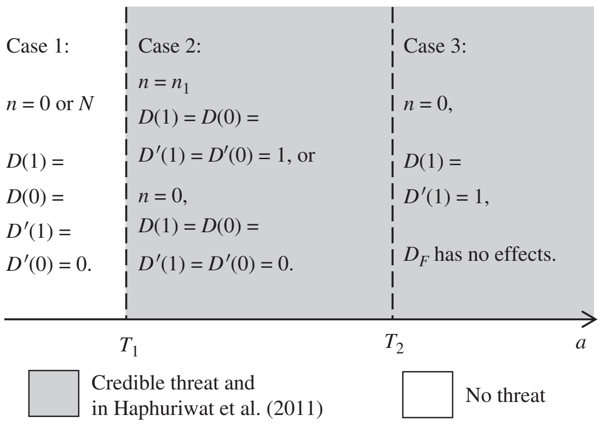
[2] Recall that notation used in Figure 1 has been modified. In particular, we have $D(1)$, $D(0)$, $D'(1)$, and $D'(0)$ instead of D_S, D_F, D'_S, and D'_F; $r(1)$, $r(0)$, $k_A(1)$, and $k_A(0)$ instead of S_d, F_d, S_a, and F_a; and m, v, d, and a instead of M, V, C_d, and C_a, respectively.

Figure 2 Solution to the Game Tree in Figure 1(b) Using Thresholds T_1, T_2, T_3, and n_1; Required Inspection Levels n_2, n_3, and n_4; and Conditions C_1, C_2, C_3, C_4, C_5, C_6, and C_7 as Defined in Table 1



In particular, although Regions I.1–I.3 in Figure 2 are similar to Cases 1–3 in Figure 3, there are two main differences between Region I in Figure 2 and Figure 3. First, there is an additional case (Region I.4 in Figure 2), when the smuggling cost is high ($a > T_3$; this refers to the condition that the cost of smuggling is higher than level T_3, which is specified in Table 1). In Region I.4, the defender does not desire to declare retaliation or inspect any cargo since the high smuggling cost alone will prevent the smuggler from smuggling nuclear weapons, whereas in Figure 3 it is assumed that retaliation declaration is free and therefore the defender would declare retaliation at equilibrium. Second, if the inspection cost is low (condition C_3 does not hold), in Region I.3 of Figure 2, a certain inspection level ($n = \min\{N, n_2\}$) in this model can deter the smuggler, instead of declaring any retaliation as in Case 2 of Figures 3. In contrast, in Case 2 of Figure 3 ($T_1 < a \le T_2$), a paired strategy of retaliation threats and inspection ($n = \min\{N, n_1\}$) can deter the smuggler only if the threat is credible. If the threat is noncredible ($r(s) > c(1, s, 0) - c(1, s, 1)$ for $s = 1$ or 0), the smuggler will not be deterred even if the defender inspects $n = \min\{N, n_1\}$ containers. In this case, the payoff of smuggling for the smuggler is $v((m^\alpha a + k_A(0))/(v - k_A(1) + k_A(0))) - k_A(0)((v - k_A(1) - m^\alpha a)/(v - k_A(1) + k_A(0))) - m^\alpha a$ if $r(1) > c(1, 1, 0) - c(1, 1, 1)$, or $(v - k_A(1))((m^\alpha a + k_A(0))/(v - k_A(1) + k_A(0))) - m^\alpha a$ if $(r(0) > c(1, 0, 0) - c(1, 0, 1))$, and both terms are greater than zero (the smuggler's payoff of not smuggling,

Figure 3 **Optimal Defender Strategies Adapted from Figure 1 of Haphuriwat et al. (2011) Using Threshold T_1, Required Inspection Level n_1, and Conditions C_1, C_2, and C_3 as Defined in Table 1**



Note. This is similar to the left 75% of Region I in Figure 2, except Region I.4, where high cost of smuggling alone is sufficient in deterring the smuggler.

see Lemma 1 in §A.6 of the appendix). Therefore, a smuggling attempt will happen since the defender declares a noncredible retaliation.

Similarly, in Case 3 of Figure 3, a threat of retaliation alone is sufficient in deterring smuggling assuming the threat is credible. However, if the threat is noncredible ($r(1) > c(1,1,0) - c(1,1,1)$), the payoff of successfully smuggling nuclear weapons is $v - m^{\alpha}a$, which could be greater than 0 as long as a is not too large. Therefore, a smuggling attempt could happen when the defender declares a noncredible retaliation and does not inspect any containers.

Proposition 2. *At equilibrium, if the threats are noncredible, the threshold inspection level to deter smuggling (n^*) is higher than or equal to that if the threats are assumed to be credible. In particular:*

- *When $a < T_1$, there is no retaliation threat, and the defender either inspects all the containers or none, depending on inspection cost d.*
- *When $T_1 \le a < T_2$,*
  —threshold n_2 (Region II.2 in Figure 2) $\ge$ threshold n_1 (Region I.2 in Figure 2);
  —threshold n_3 (Region II.2 in Figure 2) $\ge$ threshold n_1 (Region I.2 in Figure 2);
  —threshold n_4 (Region III.2 in Figure 2) $\ge$ threshold n_1 (Region I.2 in Figure 2).
- *When $T_2 \le a < T_3$,*
  —threshold n_4 (Regions III.2 in Figure 2) > 0 (Region I.3 in Figure 2).
- *When $a \ge T_3$, there is no retaliation and the defender does not inspect any containers.*

Remark 3. Proposition 2 implies that if the threats to retaliate are noncredible, the defender must inspect more or the same number of containers to deter the smuggler. Proposition 2 is illustrated by the numerical examples in §3.

There are three regions where both credible and noncredible threats are at equilibrium: Regions II.2, II.3, and III.2 in Figure 2. We use Region II.2 as an example. The reason for noncredible threats at equilibrium is that for both credible ($D(1) = D'(1) = 1$, $D(0) = D'(0) = 0$) and noncredible ($D(1) = D'(1) = 1$, $D(0) = 1$, $D'(0) = 0$) threats, when the defender declares retaliation threats, which deter smuggling attempts together with inspecting n_3 containers, the payoff to the defender is the summation of the inspection cost and the declaration cost (including research cost on how to develop an effective retaliatory policy) and does not differ between credible and noncredible threats. Note that we do not differentiate the declaration costs between successful and foiled smuggling attempts. Therefore, as long as the defender declares the retaliatory policy and inspects the $\min\{N, n_3\}$ containers, the smuggler is deterred, and noncredible threats can be achieved at equilibrium (the defender's stage 3 decisions were not actually made since the smuggling attempts would have been deterred). This implies that with partial inspection and retaliation threats, the smuggler would be deterred even if the retaliation could be noncredible (and thus never carried out).

2.6. Discussion

Game-theoretic models advise the defender regarding the optimal inspection level together with appropriately declared retaliatory policy to deter any smuggling attempts. There is some criticism regarding the assumptions and applications of game theory. For example, Axelrod (1997) states that the unrealistic assumptions of game theory limit its applications, and Poundstone (1992) criticizes game theory as a tool for predicting human behavior. There are some trade-offs between reality and solvability. Although we acknowledge the limitation of game-theoretic reasoning about the strategic interactions between attackers and defenders, game-theoretic analyses of terrorism have yielded many important insights that cannot be obtained by nonstrategic analyses (see Sandler and Arce 2003, Cox 2009 and Hall 2009). Azad (2011) also comments: "Game theory has found its best application in the development of models of deterrence." To derive meaningful results, certain simplifying assumptions have to be made. In addition, research on behavioral game theory has been developed to address such comments by combining game-theoretic modeling and experimental approaches to make better predictions about human behavior (Camerer 2003).

Furthermore, we acknowledge that the validity of defender–attacker game-theoretic models relies on their assumptions about the attacker's behavior. Traditional game-theoretic models based on the strong assumption that players are perfectly rational do not give a perfect description of the attacker's behavior and thus would be limited in their predictive power. However, defender–attacker games provide useful models to study the strategic interactions between parties with conflicting interests (e.g., Cavusoglu and Raghunathan 2004; Bier et al. 2008; Barrett 2010; Haphuriwat et al. 2011; Wang and Bier 2011; Sevillano et al. 2012; Shan and Zhuang 2013a, c). Moreover, Yang et al. (2013) recently developed a new set of Stackelberg security games (which are successfully deployed in several important real-world security domains) coping with human attackers who are boundedly rational by incorporating prospect theory (Kahneman and Tversky 1979) and a stochastic discrete choice model (Train 2003). Similarly, McLay et al. (2012) established a level-k game-theoretic model accounting for bounded rationality using robust optimization methodologies. In addition, Powell (2007), Hao et al. (2009), Shan and Zhuang (2013b) examined the nonstrategic behavior of the attacker with mathematical modeling. Using game theory, this current paper still yields some insights such as the finding that noncredible threats could be found at equilibrium.

3. Illustration of the Extended Model

To illustrate our extended model, we use the parameter estimates that are used in Haphuriwat et al. (2011), as shown in Table 3.

Table 3 Parameter Estimates Used in Haphuriwat et al. (2011)

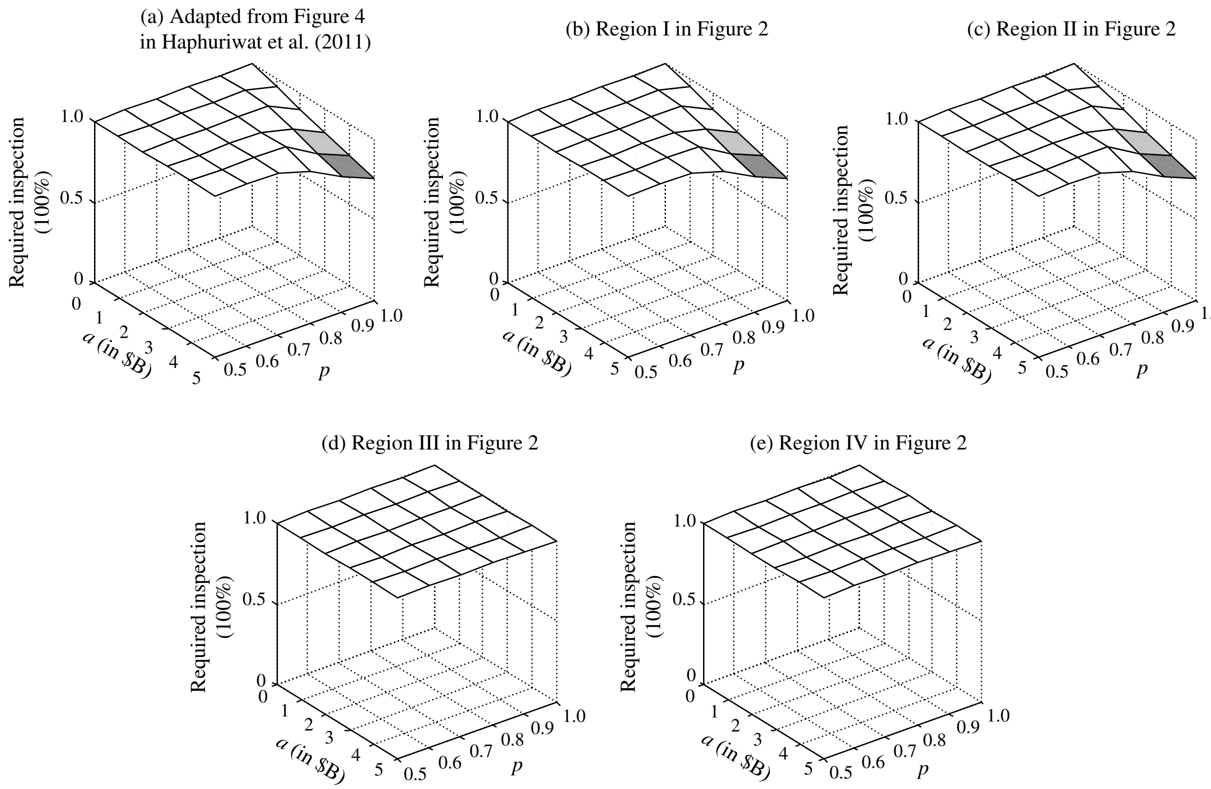
Parameter	Estimated range	Best estimate
N	10–16 million	12 million
m	1–5	1–5
v	\$500 billion–\$10 trillion	\$3 trillion
d	\$20–\$50	\$30
a	\$2 million–\$5 billion	\$500 million–\$1 billion
$k_A(1)$	\$500 billion–\$10 trillion	\$2.98 trillion
$k_A(0)$	\$0–\$10 trillion	\$0 or \$50 billion
p	0–1	0.5–1

Proposition 3. *All five additional parameters that are introduced in this paper (i.e.,* $c(1,1,1)$, $c(1,1,0)$, $c(1,0,1)$, $c(1,0,0)$, *and* e*) will only have an impact on whether retaliation against successful or foiled smuggling attempts is credible, and will not influence the exact number of containers to inspect.*

Remark 4. Proposition 3 implies that there is no need to estimate more parameters than Haphuriwat et al. (2011) when deciding how many containers to inspect. Recall that the definitions of required inspection levels do not contain the values of all five additional parameters that are introduced in this paper (i.e., $c(1,1,1)$, $c(1,1,0)$, $c(1,0,1)$, $c(1,0,0)$, and e). When the smuggler decides whether to smuggle, he compares his payoffs of smuggling and not smuggling (i.e., 0), which are functions of the inspection level (n) and the defender's decision ($D'(1)$ or $D'(0)$) in stage 3. Whereas the binary decision variables $D'(1)$ or $D'(0)$ affect the payoff and decision of the smuggler, the five additional parameters will only affect the defender's decision in stage 3. Therefore, those five additional parameters will only have an impact on whether there will be credible threats at equilibrium (as shown in Proposition 1), but will not influence the exact number of containers to inspect. Note that although the exact number of containers to inspect for deterrence differs depending on whether the retaliation threat is credible, the formulas used to calculate the exact number do not involve the five additional parameters. The insights from the model could not be reached by intuition; for example, the exact quantitative insights regarding the required inspection level for deterrence if the retaliation is not credible requires calculation.

Figures 4(a) and 5(a) are adapted from Figures 4 and 6 in Haphuriwat et al. (2011), using the same baseline

Figure 4 **Comparing Required Inspection Levels** (100%) **Adapted from Figure 4 of Haphuriwat et al. (2011) and Four Regions in Figure 2, Using the Same Baseline Values with** $k_A(0) = 0$



values. Figures 4(b)–4(e) and 5(b)–5(e) illustrate the required inspection levels for Regions I–IV in Figure 2, respectively. In particular, Figure 4 shows the required inspection levels when facing a smuggler attempting to smuggle in a single weapon with $k_A(0) = 0$; Figure 5 shows the required inspection levels facing a smuggler attempting to smuggle in more than one weapon with $k_A(0) = \$50$ billion.

If retaliation after a successful smuggling attempt is credible, the required inspection levels are the same among panels (a)–(c) of Figure 4, respectively. In these cases, partial inspection could successfully deter the smuggling attempts under many circumstances. In contrast, when retaliation against a successful smuggling attempt is not credible, a 100% level of inspection is always required in Figure 4(d) and 4(e).

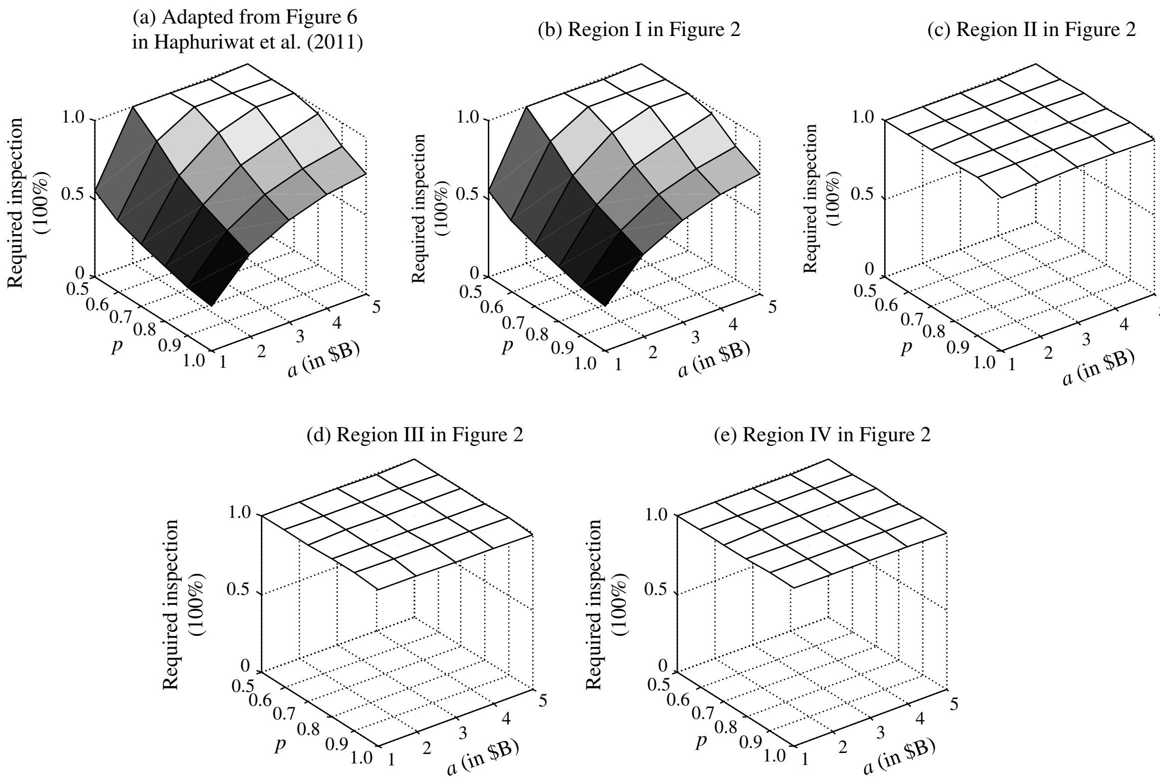
Similarly, if both types of retaliation are credible, the required inspection levels are the same between panels (a) and (b) of Figure 5. In particular, retaliation after a foiled smuggling attempt significantly helps to achieve deterrence with partial inspection. However, if either of the two types of retaliation threats is not credible, a level of 100% inspection is required in all parameter ranges in Figure 5(c)–5(e). Section A.7 of the appendix provides results for the case where m is greater than or equal to 1.

4. Conclusion and Future Research Directions

4.1. Conclusion

In this paper, we model the role of credible or noncredible retaliation threats in deterring the smuggling of nuclear weapons. Previous research suggests deterring the smuggling of nuclear weapons by retaliation threats and partial inspection. However, the credibility of such threats might influence the deterrence effects, which has not been extensively studied in the literature of deterrence against nuclear smuggling. In this paper,

Figure 5 **Comparing Required Inspection Levels** (100%) **Adapted from Figure 6 of Haphuriwat et al. (2011) and Four Regions in Figure 2, Using the Same Baseline Values with** $k_A(0) = \$50$ **Billion**



we contribute to the literature by extending Haphuriwat et al. (2011) and modeling a three-stage game to study the credibility of retaliation threats and identify the conditions required for the threats to be credible. Specifically, our results show that credible retaliation could be an equilibrium strategy when the penalty from breaking the promise of retaliation (i.e., reputation loss) is high, the reward from retaliation is high, or the costs of retaliation are low. By contrast, noncredible retaliation could be an equilibrium strategy when the reputation loss is low, the reward from the public for retaliation is low, or the costs of retaliation are high.

When the cost of retaliation to the smuggler against a foiled smuggling attempt is low, deterrence with partial inspection is difficult to achieve, even if the threat to retaliate against a successful smuggling attempt is credible. If the threat to retaliate against a successful smuggling attempt is not credible, a higher inspection level (e.g., 100%) would be required. In contrast, when the cost of retaliation to the smuggler against a foiled smuggling attempt is medium, deterrence through partial inspection becomes much more feasible, especially if the smuggler only attempts to smuggle in one weapon and the required inspection level becomes significantly lower. Therefore, it would be beneficial for the defender to declare retaliation against both a successful and a foiled smuggling attempt. If retaliation against either successful or foiled smuggling attempts is not credible, a higher inspection level (e.g., 100%) would be required.

4.2. Future Research Directions

In terms of future research directions, the strategic interactions of inspection and smuggling are likely to happen in a repeated game between the defender and the smuggler, as discussed in §6.5 of Haphuriwat et al. (2011). Both players can learn from the outcomes of the previous periods and update their beliefs about some uncertainties such as costs and benefits of smuggling and/or retaliation. Retaliation benefits could come from the reduced attacking capacity in the future periods (increased attack unit costs or decreased attack budget). Considering discount factors in such multiperiod games might also yield some additional insights about myopic players. This paper focuses on how to deter smuggling attempts of nuclear weapons with retaliation threats and partial inspection in a one-period game, which incorporates the temporal effects of the reputation loss into the current payoff. The effects of retaliation after a smuggling attempt could include deterring future attacks, which could be studied with multiperiod games.

Games of incomplete information can also be investigated in the context of container inspection. One concern about deterring the smuggling of nuclear weapons is that the cost of retaliation to the smuggler may not be sufficiently high to achieve deterrence effects. A Bayesian game, where the smuggler does not know the payoff of the defender, can be studied. In reality, as long as the terrorist believes that such cost could be sufficiently high with some high probability, such threats would be credible and thus could deter the smuggler.

Signaling games could also be incorporated into modeling credible retaliation in deterring smugglers. As studied by Zhuang et al. (2010), deception could be an equilibrium strategy for the defender over truthful disclosure, especially when the smuggler is uncertain about some of the defender's attributes. In such cases, we expect that signals of possible retaliation may not always be credible, but could be part of equilibrium strategies involving deception.

We acknowledge that there are some terrorist groups who cannot be deterred by credible retaliation threats. There might be two main reasons: (1) the severity of retaliation is ignored by them, and (2) they intend to give the government the impression of being undeterrable to get the government to give up any deterrence efforts. Those two possibilities could be modeled by modifying the payoff to the smuggler. For instance, to model scenario (2), there would be a positive payoff to the smuggler when the smuggling attempt occurs given the defender declares a retaliation threat.

Another interesting extension is to consider the damage caused by smuggled nuclear weapons to be a function of the number of smuggled nuclear weapons. In particular, if a couple of nuclear weapons are detonated simultaneously in different cities, the potential damage could be much more than separating two nuclear attacks by time.

Acknowledgments

The authors thank the former editor-in-chief, associate editor, and two anonymous referees of *Decision Analysis* for constructive comments. This research was partially supported by the U.S. Department of Homeland Security (DHS) through the National Center for Risk and Economic Analysis of Terrorism Events (CREATE) [Grant 2010-ST-061-RE0001]. This research was also partially supported by the United States National Science Foundation (NSF) [Award 1200899]. However, any opinions, findings, and conclusions or recommendations in this document are those of the authors and do not necessarily reflect views of the DHS, CREATE, or NSF.

Appendix

A.1. Solution to the Model as Specified in Figure 1(b)

In this section, we introduce two additional notation: U_A (the utility of the smuggler) and U_D (the disutility of the defender), which represent the total (expected) costs and payoffs in the right part of the game tree in Figure 1(b). We show the solution processes for Regions I.1, I.2, I.3, I.4, II.1, II.2, II.3, III.1, III.2, III.3, III.4, IV.1, IV.2, and IV.3, all in Figure 2.

- *Region* I.1 in *Figure* 2, *where* $a \le T_1$.

I.1.1. In Subgame I in Figure 1(b), $D(1)=1$ and $D(0)=1$. Solving the Stage 3 game, we obtain that $D'(1)=1$, $D'(0)=1$. In Stage 2, we obtain that $U_A=[1-((np)/N)^m](v-k_A(1))-((np)/N)^m k_A(0)-m^\alpha a$. In Stage 1, if $n=N$, $U_A>0$ and $U_D=(1-p^m)(v+r(1)+c(1,1,1))-p^m(r(0)+c(1,0,1))+Nd+e$. If $n=0$, $U_A>0$ and $U_D=v+r(1)+c(1,1,1)+e$.

I.1.2. In Subgame II, $D(1)=1$, $D(0)=0$, $D'(1)=1$, $D'(0)=0$, and $U_A=[1-((np)/N)^m](v-k_A(1))-m^\alpha a$. If $n=N$, $U_A>0$ and $U_D=(1-p^m)(v+r(1)+c(1,1,1))+Nd+e$. If $n=0$, $U_A>0$ and $U_D=v+r(1)+c(1,1,1)+e$.

I.1.3. In Subgame III, $D(1)=0$, $D(0)=1$, $D'(1)=0$, $D'(0)=1$, and $U_A=[1-((np)/N)^m]v-((np)/N)^m k_A(0)-m^\alpha a$. If $n=N$, $U_A>0$ and $U_D=(1-p^m)v+p^m(r(0)+c(1,0,1))+Nd+e$. If $n=0$, $U_A>0$ and $U_D=v+e$.

I.1.4. In Subgame IV, $D(1)=0$, $D(0)=0$, $D'(1)=0$, $D'(0)=0$, and $U_A=[1-((np)/N)^m]v-m^\alpha a$. If $n=N$, $U_A>0$ and $U_D=(1-p^m)v+Nd$. If $n=0$, $U_A>0$ and $U_D=v$.

Comparing U_D for $n=N$ or 0 in all four subgames, Subgame IV minimizes U_D. Comparing U_D for $n=N$ or 0 in Subgame IV, if $(1-p^m)v+Nd>v$ or, equivalently, $d>(v/N)p^m$, the defender's equilibrium strategy is $n=0$, $D(1)=D(0)=D'(1)=D'(0)=0$, and payoff is v. Otherwise, the equilibrium strategy is $n=N$, $D(1)=D(0)=D'(1)=D'(0)=0$, and payoff is $(1-p^m)v+Nd$.

- *Region* I.2 *in Figure* 2, *where* $T_1<a\le T_2$.

I.2.1. In Subgame I, $D'(1)=1$, $D'(0)=1$, and $U_A=[1-((np)/N)^m](v-k_A(1))-((np)/N)^m k_A(0)-m^\alpha a$. To deter the smuggler, the defender can choose n to make the smuggler indifferent between attacking and not attacking, or, equivalently, $U_A = 0$. So, if $n = (N/p)[1 - (m^\alpha a + k_A(0))/(v - k_A(1) + k_A(0))]^{1/m}$, $U_A = 0$ and $U_D = (N/p)[1 - ((m^\alpha a + k_A(0))/(v - k_A(1) + k_A(0)))]^{1/m} d + e$. If $n = 0$, $U_A > 0$ and $U_D = v + r(1) + c(1,1,1) + e$.

I.2.2. In Subgame II, $D'(1) = 1$, $D'(0) = 0$, and $U_A = [1 - ((np)/N)^m](v - k_A(1)) - m^\alpha a$. Similarly, to find optimal n, we need $U_A = 0$. If $n = (N/p)[1 - (m^\alpha a)/(v - k_A(1))]^{1/m}$, $U_A = 0$ and $U_D = (N/p)[1 - ((m^\alpha a)/(v - k_A(1)))]^{1/m} d + e$. If $n = 0$, $U_A > 0$ and $U_D = v + r(1) + c(1,1,1) + e$.

I.2.3. In Subgame III, $D'(1) = 0$, $D'(0) = 1$, and $U_A = [1 - ((np)/N)^m]v - ((np)/N)^m k_A(0) - m^\alpha a$. Similarly, to find optimal n, we need $U_A = 0$. If $n = (N/p)[1 - (m^\alpha a + k_A(0))/(v + k_A(0))]^{1/m}$, $U_A = 0$ and $U_D = (N/p)[1 - ((m^\alpha a + k_A(0))/(v + k_A(0)))]^{1/m} d + e$. If $n = 0$, $U_A > 0$ and $U_D = v + r(1) + e$.

I.2.4. In Subgame IV, $D'(1) = 0$, $D'(0) = 0$, and $U_A = [1 - ((np)/N)^m]v - m^\alpha a$. Similarly, to find the optimal n, we need $U_A = 0$. If $n = (N/p)[1 - (m^\alpha a)/v]^{1/m}$, $U_A = 0$ and $U_D = (N/p)[1 - ((m^\alpha a)/v)]^{1/m} d$. If $n = 0$, $U_A > 0$ and $U_D = v$.

Given $n > 0$, Subgame I gives the minimal $U_D = (N/p)[1 - ((m^\alpha a + k_A(0))/(v - k_A(1) + k_A(0)))]^{1/m} d + e$. Given $n = 0$, Subgame IV gives the minimal $U_D = v$. Comparing these two optimal payoffs, if $(N/p)[1 - ((m^\alpha a + k_A(0))/(v - k_A(1) + k_A(0)))]^{1/m} d + e < v$ or, equivalently, $d < ((p(v - e))/N)(1 - (m^\alpha a + k_A(0))/(v - k_A(1) + k_A(0)))^{-1/m}$, the equilibrium strategy for the defender is to choose $n = (N/p)[1 - (m^\alpha a + k_A(0))/(v - k_A(1) + k_A(0))]^{1/m}$, $D(1) = D(0) = D'(1) = D'(0) = 1$, and the equilibrium payoff is $(N/p)[1 - ((m^\alpha a + k_A(0))/(v - k_A(1) + k_A(0)))]^{1/m} d + e$. Otherwise, the equilibrium strategy for the defender is to choose $n = 0$, $D(1) = D(0) = D'(1) = D'(0) = 0$, and the equilibrium payoff is v.

- *Region* I.3 *in Figure* 2, *where* $T_2 < a \le T_3$.

I.3.1. In Subgame I, $D'(1) = 1$, $D'(0) = 1$, and $U_A = [1 - ((np)/N)^m](v - k_A(1)) - ((np)/N)^m k_A(0) - m^\alpha a < 0$. Therefore, the optimal $n = 0$, $U_A < 0$, and $U_D = e$.

I.3.2. In Subgame II, $D'(1) = 1$, $D'(0) = 0$, and $U_A = [1 - ((np)/N)^m](v - k_A(1)) - m^\alpha a < 0$. Therefore, the optimal $n = 0$, $U_A < 0$, and $U_D = e$.

I.3.3. In Subgame III, $D'(1) = 0$, $D'(0) = 1$, and $U_A = [1 - ((np)/N)^m]v - ((np)/N)^m k_A(0) - m^\alpha a$. Since we assume that $k_A(0) < k_A(1)$, the optimal $n = 0$, $U_A < 0$, and $U_D = e$.

I.3.4. In Subgame IV, $D'(1) = 0$, $D'(0) = 0$, and $U_A = [1 - ((np)/N)^m]v - m^\alpha a$. To find the optimal n, we need $U_A = 0$. If $n = (N/p)[1 - (m^\alpha a)/v]^{1/m}$, $U_A = 0$ and $U_D = (N/p)[1 - ((m^\alpha a)/v)]^{1/m} d$. If $n = 0$, $U_A > 0$ and $U_D = v$.

We observe that given $n = 0$, Subgames I–III give the minimal $U_D = e$. Given $n = (N/p)[1 - (m^\alpha a)/v]^{1/m}$, Subgame IV gives the minimal $U_D = (N/p)[1 - ((m^\alpha a)/v)]^{1/m} d$. Comparing these two optimal payoffs, we note that if $(N/p)[1 - ((m^\alpha a)/v)]^{1/m} d < e$ or, equivalently, $d < ((p(v - e))/N)(1 - (m^\alpha a)/v)^{-1/m}$, the equilibrium strategy for the defender is to choose $n = (N/p)[1 - ((m^\alpha a)/v)]^{1/m} d$, $D(1) = D(0) = D'(1) = D'(0) = 0$, and the equilibrium payoff is $(N/p)[1 - ((m^\alpha a)/v)]^{1/m} d$. Otherwise, the equilibrium strategy for the defender is to choose $n = 0$, $D(1) = D(0) = D'(1) = D'(0) = 1$ or $D(1) = D(0) = D'(1) = 1$, $D'(0) = 0$ or $D(1) = D'(1) = 0$, $D(0) = D'(0) = 1$, and the equilibrium payoff is e.

- *Region* I.4 *in Figure* 2, *where* $a > T_3$.

I.4.1. In Subgame I, $D'(1) = 1$, $D'(0) = 1$, and $U_A = [1 - ((np)/N)^m](v - k_A(1)) - ((np)/N)^m k_A(0) - m^\alpha a < 0$. Therefore, the optimal $n = 0$, $U_A < 0$, and $U_D = e$.

I.4.2. In Subgame II, $D'(1) = 1$, $D'(0) = 0$, and $U_A = [1 - ((np)/N)^m](v - k_A(1)) - m^\alpha a < 0$. Therefore, the optimal $n = 0$, $U_A < 0$, and $U_D = e$.

I.4.3. In Subgame III, $D'(1) = 0$, $D'(0) = 1$, and $U_A = [1 - ((np)/N)^m]v - ((np)/N)^m k_A(0) - m^\alpha a$. Since we assume that $k_A(0) < k_A(1)$, the optimal $n = 0$, $U_A < 0$, and $U_D = e$.

I.4.4. In Subgame IV, $D'(1) = 0$, $D'(0) = 0$, and $U_A = [1 - ((np)/N)^m]v - m^\alpha a < 0$. Therefore, the optimal $n = 0$, $U_A < 0$, and $U_D = e$.

We observe that the equilibrium strategy for the defender is to choose $n = 0$, $D(1) = D'(1) = D(0) = D'(0) = 0$, $U_A < 0$, and $U_D = 0$.

- *Region* II.1. This region is identical to Region I.1 of Figure 2.
- *Region II.2 in Figure* 2, *where* $T_1 < a \le T_2$.

II.2.1. In Subgame I, $D'(1) = 1$, $D'(0) = 0$, and $U_A = [1 - ((np)/N)^m](v - k_A(1)) - m^\alpha a$. To find the optimal n, we need $U_A = 0$. If $n = (N/p)[1 - (m^\alpha a)/(v - k_A(1))]^{1/m}$, $U_A = 0$ and $U_D = (N/p)[1 - ((m^\alpha a)/(v - k_A(1)))]^{1/m} d + e$. If $n = 0$, $U_A > 0$ and $U_D = v + r(1) + c(1,1,1) + e$.

II.2.2. In Subgame II, $D'(1) = 1$, $D'(0) = 0$, and $U_A = [1 - ((np)/N)^m](v - k_A(1)) - m^\alpha a$. Similarly, to find the optimal n, we need $U_A = 0$. If $n = (N/p)[1 - (m^\alpha a)/(v - k_A(1))]^{1/m}$, $U_A = 0$ and $U_D = (N/p)[1 - ((m^\alpha a)/(v - k_A(1)))]^{1/m} d + e$. If $n = 0$, $U_A > 0$ and $U_D = v + r(1) + c(1,1,1) + e$.

II.2.3. In Subgame III, $D'(1) = 0$, $D'(0) = 0$, and $U_A = [1 - ((np)/N)^m]v - m^\alpha a$. Similarly, to find the optimal n, we need $U_A = 0$. If $n = (N/p)[1 - (m^\alpha a)/v]^{1/m}$, $U_A = 0$ and $U_D = (N/p)[1 - ((m^\alpha a)/v)]^{1/m} d + e$. If $n = 0$, $U_A > 0$ and $U_D = v + r(1) + e$.

II.2.4. In Subgame IV, $D'(1) = 0$, $D'(0) = 0$, and $U_A = [1 - ((np)/N)^m]v - m^\alpha a$. Similarly, to find the optimal n, we need $U_A = 0$. If $n = (N/p)[1 - (m^\alpha a)/v]^{1/m}$, $U_A = 0$ and $U_D = (N/p)[1 - ((m^\alpha a)/v)]^{1/m} d$. If $n = 0$, $U_A > 0$ and $U_D = v$.

We observe that if

$$U_D = (N/p)[1 - ((m^\alpha a)/(v - k_A(1)))]^{1/m} d + e < \min\{(N/p)[1 - ((m^\alpha a)/v)]^{1/m} d, v\},$$

the optimal $n > 0$, and Subgames I and II give the minimal $U_D = (N/p)[1 - ((m^\alpha a)/(v - k_A(1)))]^{1/m} d + e$. Therefore, the equilibrium strategy for the defender is to choose $n = (N/p)[1 - (m^\alpha a)/(v - k_A(1))]^{1/m}$, $D(1) = D(0) = D'(1) = D'(0) = 1$ or $D(1) = D'(1) = 1$, $D(0) = D'(0) = 0$, and the equilibrium payoff is $U_D = (N/p)[1 - ((m^\alpha a)/(v - k_A(1)))]^{1/m} d + e$. If $U_D = (N/p)[1 - ((m^\alpha a)/v)]^{1/m} d + e < \min\{(N/p)[1 - ((m^\alpha a)/(v - k_A(1)))]^{1/m} d, v\}$, the optimal $n > 0$, and Subgame IV gives the minimal $U_D = (N/p)[1 - ((m^\alpha a)/v)]^{1/m} d$. Therefore, the equilibrium strategy for the defender is to choose $n =$

$(N/p)[1-(m^\alpha a)/v]^{1/m}$, $D(1)=D(0)=D'(1)=D'(0)=0$, and the equilibrium payoff is $U_D=(N/p)[1-((m^\alpha a)/v)]^{1/m}d+e$. Otherwise, the optimal $n=0$, and Subgame IV gives the minimal $U_D=v$. Therefore, the equilibrium strategy for the defender is to choose $n=0$, and the equilibrium payoff is $U_D=v$.

- *Region* II.3 *in Figure* 2, *where* $T_2<a\le T_3$.

II.3.1. In Subgame I, $D'(1)=1$, $D'(0)=0$, and $U_A=[1-((np)/N)^m](v-k_A(1))-m^\alpha a<0$. Therefore, the optimal $n=0$, $U_A<0$, and $U_D=e$.

II.3.2. In Subgame II, $D'(1)=1$, $D'(0)=0$, and $U_A=[1-((np)/N)^m](v-k_A(1))-m^\alpha a<0$. Therefore, the optimal $n=0$, $U_A<0$, and $U_D=e$.

II.3.3. In Subgame III, $D'(1)=0$, $D'(0)=0$, and $U_A=[1-((np)/N)^m]v-m^\alpha a>0$. To find the optimal n, we need $U_A=0$. If $n=(N/p)[1-(m^\alpha a)/v]^{1/m}$, $U_A=0$ and $U_D=(N/p)[1-((m^\alpha a)/v)]^{1/m}d$. If $n=0$, $U_A>0$ and $U_D=v$.

II.3.4. In Subgame IV, $D'(1)=0$, $D'(0)=0$, and $U_A=[1-((np)/N)^m]v-m^\alpha a$. To find the optimal n, we need $U_A=0$. If $n=(N/p)[1-(m^\alpha a)/v]^{1/m}$, $U_A=0$ and $U_D=(N/p)[1-((m^\alpha a)/v)]^{1/m}d$. If $n=0$, $U_A>0$ and $U_D=v$.

We observe that if $(N/p)[1-((m^\alpha a)/v)]^{1/m}d<e$, the optimal $n=(N/p)[1-(m^\alpha a)/v]^{1/m}$, and Subgames III and IV give the minimal $U_D=(N/p)[1-((m^\alpha a)/v)]^{1/m}d$. Therefore, the equilibrium strategy for the defender is $D(1)=1$, $D'(1)=D(0)=D'(0)=0$. Otherwise, the optimal $n=0$, and Subgames I and II give the minimal $U_D=e$. Therefore, the equilibrium strategy for the defender is $D(1)=D'(1)=D(0)=D'(0)=1$ or $D(1)=D'(1)=1$, $D(0)=D'(0)=0$.

- *Region* II.4 *in Figure* 2, *where* $a>T_3$.

II.4.1. In Subgame I, $D'(1)=1$, $D'(0)=0$, and $U_A=[1-((np)/N)^m](v-k_A(1))-m^\alpha a<0$. Therefore, the optimal $n=0$, $U_A<0$, and $U_D=e$.

II.4.2. In Subgame II, $D'(1)=1$, $D'(0)=0$, and $U_A=[1-((np)/N)^m](v-k_A(1))-m^\alpha a<0$. Therefore, the optimal $n=0$, $U_A<0$, and $U_D=e$.

II.4.3. In Subgame III, $D'(1)=0$, $D'(0)=0$, and $U_A=[1-((np)/N)^m]v-m^\alpha a<0$. Therefore, the optimal $n=0$, $U_A<0$, and $U_D=e$.

II.4.4. In Subgame IV, $D'(1)=0$, $D'(0)=0$, and $U_A=[1-((np)/N)^m]v-m^\alpha a<0$. Therefore, the optimal $n=0$, $U_A<0$, and $U_D=0$.

We observe that the equilibrium strategy for the defender is to choose $n=0$, $D(1)=D'(1)=D(0)=D'(0)=0$, $U_A<0$, and $U_D=0$.

- *Region* III.1. This region is identical to Region I.1 of Figure 2.
- *Region* III.2 *in Figure* 2, *where* $T_1<a\le T_3$.

III.2.1. In Subgame I, $D'(1)=0$, $D'(0)=1$, and $U_A=[1-((np)/N)^m]v-((np)/N)^M k_A(0)-m^\alpha a$. To find the optimal n, we need $U_A=0$. If $n=(N/p)[1-(m^\alpha a+k_A(0))/(v+k_A(0))]^{1/m}$, $U_A=0$ and $U_D=(N/p)[1-((m^\alpha a+k_A(0))/(v+k_A(0)))]^{1/m}d+e$. If $n=0$, $U_A>0$ and $U_D=v+c(1,1,0)+e$.

III.2.2. In Subgame II, $D'(1)=0$, $D'(0)=0$, and $U_A=[1-((np)/N)^m]v-m^\alpha a$. Similarly, to find the optimal n, we need $U_A=0$. If $n=(N/p)[1-(m^\alpha a)/v]^{1/m}$, $U_A=0$ and $U_D=(N/p)[1-((m^\alpha a)/v)]^{1/m}d+e$. If $n=0$, $U_A>0$ and $U_D=v+c(1,1,1)+e$.

III.2.3. In Subgame III, $D'(1)=0$, $D'(0)=0$, and $U_A=[1-((np)/N)^m]v-m^\alpha a$. Similarly, to find the optimal n, we need $U_A=0$. If $n=(N/p)[1-(m^\alpha a)/v]^{1/m}$, $U_A=0$ and $U_D=(N/p)[1-((m^\alpha a)/v)]^{1/m}d+e$. If $n=0$, $U_A>0$ and $U_D=v+e$.

III.2.4. In Subgame IV, $D'(1)=0$, $D'(0)=0$, and $U_A=[1-((np)/N)^m]v-m^\alpha a$. Similarly, to find the optimal n, we need $U_A=0$. If $n=(N/p)[1-(m^\alpha a)/v]^{1/m}$, $U_A=0$ and $U_D=(N/p)[1-((m^\alpha a)/v)]^{1/m}d$. If $n=0$, $U_A>0$ and $U_D=v$.

We observe that if $(N/p)[1-(m^\alpha a+k_A(0))/(v+k_A(0))]^{1/m}\cdot d+e<\min\{(N/p)[1-((m^\alpha a)/v)]^{1/m}d,v\}$, the optimal $n=(N/p)[1-(m^\alpha a)/(v-k_A(1))]^{1/m}$, and Subgame I ($D(1)=D(0)=D'(1)=D'(0)=1$ or $D(1)=D'(1)=1$) gives the minimal $U_D=(N/p)[1-((m^\alpha a+k_A(0))/(v+k_A(0)))]^{1/m}d+e$. If

$$(N/p)[1-((m^\alpha a)/v)]^{1/m}d<\min\{(N/p)[1-((m^\alpha a+k_A(0))/(v+k_A(0)))]^{1/m}d+e,v\},$$

the optimal $n=(N/p)[1-(m^\alpha a)/v]^{1/m}$, and Subgame IV ($D(1)=D(0)=D'(1)=D'(0)=0$) gives the minimal $U_D=(N/p)[1-((m^\alpha a+k_A(0))/(v+k_A(0)))]^{1/m}d$. Otherwise, the optimal $n=0$, and Subgame IV ($D(1)=D(0)=D'(1)=D'(0)=0$) gives the minimal $U_D=v$.

- *Region* III.3 *in Figure* 2, *where* $a\ge T_3$.

III.3.1. In Subgame I, $D'(1)=0$, $D'(0)=1$, and $U_A=[1-((np)/N)^m]v-((np)/N)^M k_A(0)-m^\alpha a<0$. Therefore, the optimal $n=0$, $U_A<0$, and $U_D=e$.

III.3.2. In Subgame II, $D'(1)=0$, $D'(0)=0$, and $U_A=[1-((np)/N)^m]v-m^\alpha a<0$. Therefore, the optimal $n=0$, $U_A<0$, and $U_D=e$.

III.3.3. In Subgame III, $D'(1)=0$, $D'(0)=1$, and $U_A=[1-((np)/N)^m]v-((np)/N)^M k_A(0)-m^\alpha a<0$. Therefore, the optimal $n=0$, $U_A<0$, and $U_D=e$.

III.3.4. In Subgame IV, $D'(1)=0$, $D'(0)=0$, and $U_A=[1-((np)/N)^m]v-m^\alpha a<0$. Therefore, the optimal $n=0$, $U_A<0$, and $U_D=0$.

We observe that the equilibrium strategy for the defender is to choose $n=0$, $D(1)=D'(1)=D(0)=D'(0)=0$, $U_A<0$, and $U_D=0$.

- *Region* IV.1. This region is identical to Region I.1 of Figure 2.
- *Region* IV.2 *in Figure* 2, *where* $T_1<a\le T_3$.

IV.2.1. In Subgame I, $D'(1)=0$, $D'(0)=0$, and $U_A=[1-((np)/N)^m]v-m^\alpha a$. To find the optimal n, we need $U_A=0$. If $n=(N/p)[1-(m^\alpha a)/v]^{1/m}$, $U_A=0$ and $U_D=(N/p)[1-((m^\alpha a)/v)]^{1/m}d+e$. If $n=0$, $U_A>0$ and $U_D=v+c(1,1,1)+e$.

IV.2.2. In Subgame II, $D'(1)=0$, $D'(0)=0$, and $U_A=[1-((np)/N)^m]v-m^\alpha a$. Similarly, to find the optimal n, we need $U_A=0$. If $n=(N/p)[1-(m^\alpha a)/v]^{1/m}$, $U_A=0$ and $U_D=(N/p)[1-((m^\alpha a)/v)]^{1/m}d+e$. If $n=0$, $U_A>0$ and $U_D=v+c(1,1,0)+e$.

IV.2.3. In Subgame III, $D'(1)=0$, $D'(0)=0$, and $U_A=[1-((np)/N)^m]v-m^\alpha a$. Similarly, to find the optimal n, we need $U_A=0$. If $n=(N/p)[1-(m^\alpha a)/v]^{1/m}$, $U_A=0$ and $U_D=(N/p)[1-((m^\alpha a)/v)]^{1/m}d+e$. If $n=0$, $U_A>0$ and $U_D=v+e$.

IV.2.4. In Subgame IV, $D'(1)=0$, $D'(0)=0$, and $U_A=[1-((np)/N)^m]v-m^\alpha a$. Similarly, to find the optimal n, we need $U_A=0$. If $n=(N/p)[1-(m^\alpha a)/v]^{1/m}$, $U_A=0$ and $U_D=(N/p)[1-((m^\alpha a)/v)]^{1/m}d$. If $n=0$, $U_A>0$ and $U_D=v$.

We observe that if $(N/p)[1-((m^\alpha a)/v)]^{1/m}d+e<v$, the optimal $n=(N/p)[1-(m^\alpha a)/v]^{1/m}$, and Subgame IV ($D(1)=D(0)=D'(1)=D'(0)=0$) gives the minimal $U_D=(N/p)[1-((m^\alpha a)/v)]^{1/m}d+e$. Otherwise, the optimal $n=0$, and Subgame IV ($D(1)=D(0)=D'(1)=D'(0)=0$) gives the minimal $U_D=v$.

- *Region* IV.3 *in Figure* 2, *where* $a\ge T_3$.

IV.3.1. In Subgame I, $D'(1)=0$, $D'(0)=0$, and $U_A=[1-((np)/N)^m]v-m^\alpha a<0$. Therefore, the optimal $n=0$, $U_A<0$, and $U_D=e$.

IV.3.2. In Subgame II, $D'(1)=0$, $D'(0)=0$, and $U_A=[1-((np)/N)^m]v-m^\alpha a<0$. Therefore, the optimal $n=0$, $U_A<0$, and $U_D=e$.

IV.3.3. In Subgame III, $D'(1)=0$, $D'(0)=0$, and $U_A=[1-((np)/N)^m]v-((np)/N)^m k_A(0)-m^\alpha a<0$. Therefore, the optimal $n=0$, $U_A<0$, and $U_D=e$.

IV.3.4. In Subgame IV, $D'(1)=0$, $D'(0)=0$, and $U_A=[1-((np)/N)^m]v-m^\alpha a<0$. Therefore, the optimal $n=0$, $U_A<0$, and $U_D=0$.

We observe that the equilibrium strategy for the defender is to choose $n=0$, $D(1)=D'(1)=D(0)=D'(0)=0$, $U_A<0$, and $U_D=0$.

In summary, we verify that the results presented in Figure 2 solve the extended game model presented in Figure 1(b).

A.2. Proof of Proposition 1

As illustrated in Regions I–III of Figure 2, at some equilibriums we have credible threats. When the defender makes a decision in stage 3 in Figure 1(b), there are three possibilities for retaliation:

- If $r(1)<c(1,1,0)-c(1,1,1)$ and $r(0)<c(1,0,0)-c(1,0,1)$ (Region I of Figure 2), in Subgames I and II of Figure 1(b) the defender would choose $D'(1)=1$ over $D'(1)=0$ since $r(1)<c(1,1,0)-c(1,1,1)$, and in Subgames I and III the defender would choose $D'(0)=1$ over $D'(0)=0$ since $r(0)<c(1,0,0)-c(1,0,1)$. Therefore, at equilibrium, the defender would choose $D(1)=D'(1)=1$ and $D(0)=D'(0)=1$ (type (a) credible retaliation in Definition 1), $D(1)=D'(1)=1$ and $D(0)=D'(0)=0$ (type (b) credible retaliation in Definition 1), or $D(1)=D'(1)=0$ and $D(0)=D'(0)=1$ (type (c) credible retaliation in Definition 1).
- If $r(1)<c(1,1,0)-c(1,1,1)$ and $r(0)\ge c(1,0,0)-c(1,0,1)$ (Region II of Figure 2), in Subgames I and II of Figure 1(b) the defender would choose $D'(1)=1$ over $D'(1)=0$ since $r(1)<c(1,1,0)-c(1,1,1)$, and in Subgames I and III the defender would choose $D'(0)=0$ over $D'(0)=1$ since $F_d\ge c(1,0,0)-c(1,0,1)$. Therefore, at equilibrium, the defender would choose $D(1)=D'(1)=1$ and $D(0)=D'(0)=0$ (type (b) credible retaliation in Definition 1).
- If $r(1)\ge c(1,1,0)-c(1,1,1)$ and $r(0)<c(1,0,0)-c(1,0,1)$ (Region III of Figure 2), in Subgames I and II of Figure 1(b) the defender would choose $D'(1)=0$ over $D'(1)=1$ since $r(1)\ge c(1,1,0)-c(1,1,1)$[11], and in Subgames I and III the defender would choose $D'(0)=1$ over $D'(0)=0$ since $r(0)<c(1,0,0)-c(1,0,1)$. Therefore, at equilibrium, the defender would choose $D(1)=D'(1)=0$ and $D(0)=D'(0)=1$ (type (c) credible retaliation in Definition 1).

A.3. Proof of Proposition 2

Depending upon the value of smuggling cost (a), there are four ranges separated by three thresholds (T_1, T_2, T_3) shown in Table 1. From Figure 2, differences among Regions I–IV arise only when $T_1<a\le T_3$ holds. There are four cases:

Case a. When $a<T_1$, the smuggler will always try to smuggle nuclear weapons due to the low cost. Therefore, the defender declares no retaliation threats and either inspects all the containers or none depending on inspection cost (d) as shown in Regions I.1–IV.1 of Figure 2. We conclude that when $a<T_1$, the threshold inspection level to deter smuggling (n^*) if we have noncredible threats is equal to that we have credible threats.

Case b. When $T_1\le a<T_2$, Region I.2 of Figure 2 requires inspecting $\min\{N,n_1\}$ containers, whereas Region II.2 requires inspecting $\min\{N,n_2\}$ or $\min\{N,n_3\}$ containers, and Region III.2 requires inspecting $\min\{N,n_4\}$ containers. Note that since the declared defender retaliation threats are always noncredible in Region IV ($r(1)\ge c(1,1,0)-c(1,1,1)$ and $r(0)\ge c(1,0,0)-c(1,1,1)$) and the declaration cost is not free ($e>0$), the defender will not declare retaliation threats at equilibrium. We show that (a.1) $n_1<n_2$, (a.2) $n_1<n_3$, and (a.3) $n_1<n_4$ as follows.

(a.1) $n_1\le n_2 \Longleftrightarrow -(m^\alpha a+k_A(0))/(v-k_A(1)+k_A(0))\le -(m^\alpha a)/v \Longleftrightarrow (m^\alpha a+k_A(0))/(v-k_A(1)+k_A(0))\ge (m^\alpha a)/v \Longleftrightarrow vm^\alpha a+vk_A(0)\ge vm^\alpha a-k_A(1)m^\alpha a+k_A(0)m^\alpha a \Longleftrightarrow (k_A(1)-k_A(0))m^\alpha a\ge -vk_A(0) \Longleftrightarrow a\ge(-vk_A(0))/((k_A(1)-k_A(0))m^\alpha)$. Since we assume that $a>0$ and observe that $(-vk_A(0))/((k_A(1)-k_A(0))m^\alpha)<0$, the last inequality $a>(-vk_A(0))/((k_A(1)-k_A(0))m^\alpha)$ always holds, and thus $n_1\le n_2$.

(a.2) $n_1<n_3 \Longleftrightarrow -(m^\alpha a+k_A(0))/(v-k_A(1)+k_A(0))\le -(m^\alpha a)/(v-k_A(1)) \Longleftrightarrow (m^\alpha a+k_A(0))/(v-k_A(1)+k_A(0))\ge (m^\alpha a)/(v-k_A(1)) \Longleftrightarrow vm^\alpha a+vk_A(0)-k_A(1)m^\alpha a-k_A(1)k_A(0)\ge vm^\alpha a-k_A(1)m^\alpha a+k_A(0)m^\alpha a \Longleftrightarrow m^\alpha a\le v-k_A(1) \Longleftrightarrow a\le (v-k_A(1))/m^\alpha$. Since $T_1\ge a<T_2=(v-k_A(1))/m^\alpha$, the last inequality $a<(v-k_A(1))/m^\alpha$ always holds, and thus $n_1<n_3$.

(a.3) $n_1\le n_4 \Longleftrightarrow -(m^\alpha a+k_A(0))/(v-k_A(1)+k_A(0))\le -(m^\alpha a+k_A(0))/(v+k_A(0)) \Longleftrightarrow (m^\alpha a+k_A(0))/(v-k_A(1)+k_A(0))\ge (m^\alpha a+k_A(0))/(v+k_A(0)) \Longleftrightarrow vm^\alpha a+vk_A(0)+k_A(0)m^\alpha a+k_A(0)^2\ge vm^\alpha a+vk_A(0)-k_A(1)m^\alpha a-k_A(1)k_A(0)+k_A(0)m^\alpha a+k_A(0)^2 \Longleftrightarrow 0\ge -k_A(1)m^\alpha a-k_A(1)k_A(0)$. Since we assume that $a>0$, $m>0$, $k_A(1)>0$, and $k_A(0)\ge 0$, the last inequality $0\ge -k_A(1)m^\alpha a-k_A(1)k_A(0)$ always holds, and thus $n_1\le n_4$.

Therefore, we conclude that when $T_1\le a<T_2$, the threshold inspection level to deter smuggling (n^*) if we have noncredible threats is higher than or equal to that if we have credible threats.

Case c. If $T_2 \le a < T_3$, Regions I and II in Figure 2 require inspecting no containers when the defender declares retaliation at equilibrium, whereas Region III requires inspecting $\min\{N, n_4\}$ containers when the defender declares retaliation at equilibrium, and Region IV requires inspecting no containers when the defender does not declare retaliation (the defender always chooses not to declare retaliation in Region IV). In other words, there are noncredible threats at equilibrium in Regions II and III, and the required inspection levels are 0 and n_4, respectively. By definition of threshold inspection level n_4 in Table 1, $n_4 > 0$. Therefore, we conclude that when $T_2 \le a < T_3$, the threshold inspection level to deter smuggling (n^*) is higher if we have noncredible threats than if we have credible threats.

Case d. When $a \ge T_3$, the smuggler will not smuggle nuclear weapons due to the high cost. Therefore, the defender declares no retaliation threats, and the defender does not inspect any containers as shown in Regions I.4–IV.4 of Figure 2.

In summary, we conclude that at equilibrium, the threshold inspection level to deter smuggling (n^*) if we have noncredible threats is higher than or equal to that if we have credible threats.

A.4. Proof of Proposition 3

The defender may need to inspect a certain number of containers to deter the smuggler from smuggling nuclear weapons. Depending upon the credibility of declared retaliation threats and other conditions, the required inspection level could be n_1, n_2, n_3, or n_4 as introduced in Table 1. However, none of them involve the newly introduced parameters ($c(1,1,1)$, $c(1,1,1)$, $c(1,0,1)$, $c(1,0,1)$, and e). Therefore, the newly introduced parameters will not influence the defender's decision of how many containers to inspect to deter the smuggling attempts.

A.5. Additional Explanations to Solutions for Regions II–IV in Figure 2

The solutions in Regions II.1 and II.4 of Figure 2 are identical to the ones in Regions I.1 and I.4, respectively. In Region II.2, depending upon inspection cost d, declaration cost e, and expected damage from a smuggled nuclear weapon v, the defender has the following equilibrium strategy: (a) If d is low (condition C_4 holds), $n = \min\{N, n_2\}$ is sufficient in deterring smugglers so that there is no need to declare retaliation ($D(1) = D(0) = D'(1) = D'(0) = 0$). (b) If d is medium (condition C_5 holds), $n = \min\{N, n_3\}$ paired with credible threat against a successful smuggling attempt ($D(1) = D'(1) = 1$) is sufficient for deterrence. Note that in this case we have two equilibriums, where declared retaliation against a foiled smuggling attempt is either credible or not. And (c) if d is high (condition C_7 holds), the defender would rather suffer an expected damage (v) instead of attempting to deter the smuggling. Region II.3 is similar to Region I.3 except that there is a noncredible threat at equilibrium ($D(1) = D'(1) = 1, D(0) = 1, D'(0) = 0$).

The solutions in Regions III.1 and III.3 of Figure 2 are identical to the ones in Regions I.1 and I.4. In Region III.2, given inspection is effective ($p^m > k_A(1)/v$), if condition C_5 holds, $n = \min\{N, n_4\}$ paired with credible retaliation against a foiled smuggling attempt can deter the smuggler. Note that the retaliation against a successful smuggling attempt could either be credible or noncredible. Otherwise, a higher level of inspection ($n = \min\{N, n_2\}$) can deter smugglers so that there is no need to declare any retaliations ($D(1) = D'(1) = D(0) = D'(0) = 0$).

The solutions in Regions IV.1 and IV.3 of Figure 2 are identical to the ones in Regions I.1 and I.4. In Region IV.2, if condition C_7 holds, $n = 0$, and there is no need to inspect because of the high inspection cost. Otherwise, $n = \min\{N, n_2\}$ can deter the smuggler. In both situations, there is no need to declare any retaliations ($D(1) = D'(1) = D(0) = D'(0) = 0$).

A.6. Lemma 1 and Proof

Lemma 1 demonstrates that a smuggling attempt will happen if the defender declares a noncredible retaliation as shown in §2.5.

Lemma 1. (a) $v((m^\alpha a + k_A(0))/(v - k_A(1) + k_A(0))) - k_A(0)((v - k_A(1) - m^\alpha a)/(v - k_A(1) + k_A(0))) - m^\alpha a > 0$; (b) *when* $T_1 \le a < T_2$, $(v - k_A(1))((m^\alpha a + k_A(0))/(v - k_A(1) + k_A(0))) - m^\alpha a > 0$.

Proof. For part (a), we have $v((m^\alpha a + k_A(0))/(v - k_A(1) + k_A(0))) - k_A(0)((v - k_A(1) - m^\alpha a)/(v - k_A(1) + k_A(0))) - m^\alpha a = (vm^\alpha a + vk_A(0) - vk_A(0) + k_A(1)k_A(0) - k_A(0)^2 - vm^\alpha a + k_A(1)m^\alpha a - k_A(0)m^\alpha a)/(v - k_A(1) + k_A(0)) = (k_A(1)k_A(0) + k_A(1)m^\alpha a)/(v - k_A(1) + k_A(0))$. Using the properties of v, $k_A(1)$, $k_A(0)$, m, and α in Table 1 (i.e., $v > k_A(1) > 0$, $k_A(0) > 0$, $m > 0$, and $\alpha > 0$), we have $(k_A(1)k_A(0) + k_A(1)m^\alpha a)/(v - k_A(1) + k_A(0)) > 0$, and thus $v((m^\alpha a + k_A(0))/(v - k_A(1) + k_A(0))) - k_A(0)((v - k_A(1) - m^\alpha a)/(v - k_A(1) + k_A(0))) - m^\alpha a > 0$.

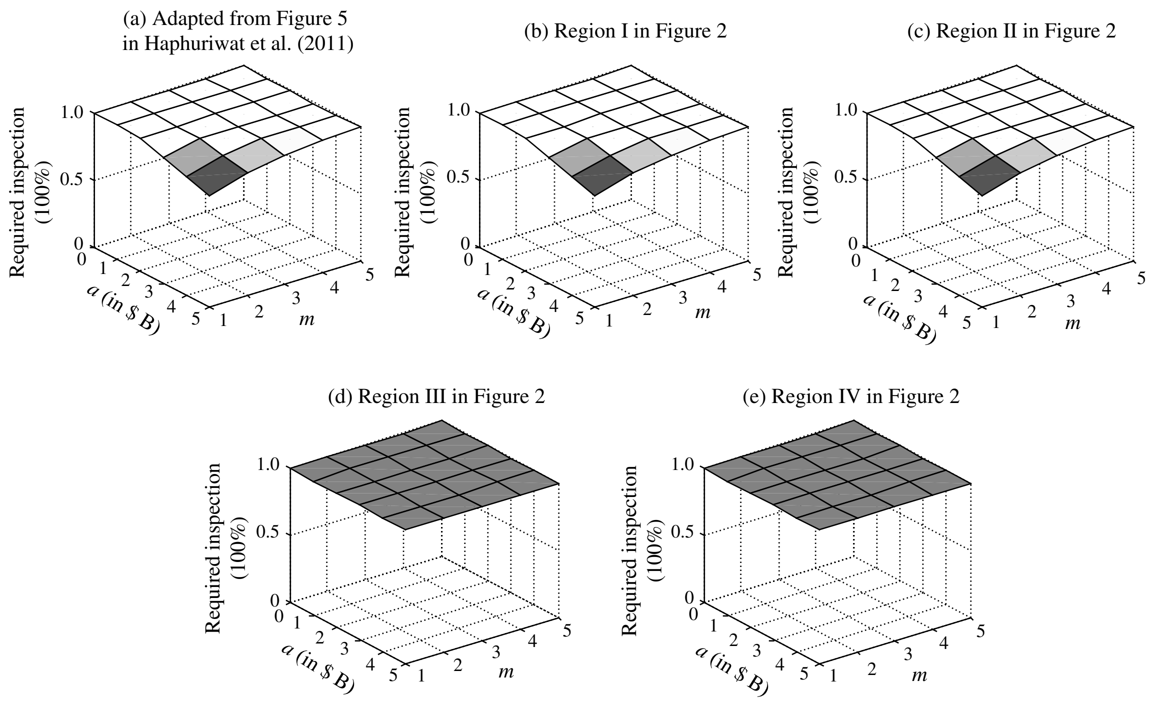
Similarly, for part (b), we have $(v - k_A(1))((m^\alpha a + k_A(0))/(v - k_A(1) + k_A(0))) - m^\alpha a = (vm^\alpha a + vk_A(0) - k_A(1)m^\alpha a - k_A(1)k_A(0) - vm^\alpha a + k_A(1)m^\alpha a - k_A(0)m^\alpha a)/(v - k_A(1) + k_A(0)) = (k_A(0)(v - k_A(1) - m^\alpha a))/(v - k_A(1) + k_A(0))$. Using the properties of v, $k_A(1)$, $k_A(0)$, m, and α in Table 1 (i.e., $v > k_A(1) > 0$, $k_A(0) \ge 0$, $m > 0$, and $\alpha > 0$) and the condition $T_1 \le a < T_2$, we have $(k_A(0)(v - k_A(1) - m^\alpha a))/(v - k_A(1) + k_A(0)) > 0$, and thus $v((m^\alpha a + k_A(0))/(v - k_A(1) + k_A(0))) - k_A(0)((v - k_A(1) - m^\alpha a)/(v - k_A(1) + k_A(0))) - m^\alpha a > 0$.

A.7. When m Is Greater Than or Equal to 1

Figure A.1 compares required inspection levels (percentage) in Figure 5 of Haphuriwat et al. (2011) and the four regions in Figure 2 using the same baseline values with $k_A(0) = 0$ and $p = 0.9$.

If retaliation after a successful attack is credible, the required inspection levels are the same among Figure A.1(a)–(c). In these cases, partial inspection could successfully deter the smuggler under circumstances. In contrast, if retaliation against a successful attack is not credible, a 100% level of inspection is always required in Figure A.1, (d) and (e).

Figure A.1 Comparing Required Inspection Levels (%) in Figure 5 of Haphuriwat et al. (2011) and Four Regions in Figure 2, Using the Same Baseline Values with $k_A(0) = 0$ and $p = 0.9$



References

Allison GT (2005) *Nuclear Terrorism: The Ultimate Preventable Catastrophe* (Time Books, Henry Holt and Company, New York).

Axelrod RM (1997) *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration* (Princeton University Press, Princeton, NJ).

Azad S (2011) Defence against terrorist attacks: Game theory can help. Accessed January 12, 2014, http://www.idsa.in/node/6468/2099.

Bakır NO (2008) A decision tree model for evaluating countermeasures to secure cargo at United States southwestern ports of entry. *Decision Anal.* 5(4):230–248.

Barrett AM (2010) Cost effectiveness of on-site chlorine generation for chlorine truck attack prevention. *Decision Anal.* 7(4): 366–377.

Bier VM, Haphuriwat N (2011) Analytical method to identify the number of containers to inspect at US ports to deter terrorist attacks. *Ann. Oper. Res.* 187(1):137–158.

Bier VM, Haphuriwat N, Menoyo J, Zimmerman R, Culpen AM (2008) Optimal resource allocation for defense of targets based on differing measures of attractiveness. *Risk Anal.* 28(3): 763–770.

Brown G, Carlyle M, Salmerón J, Wood K (2006) Defending critical infrastructure. *Interfaces* 36(6):530–544.

Camerer CF (2003) *Behavioral Game Theory: Experiments in Strategic Interaction* (Princeton University Press, Princeton, NJ).

Cavusoglu H, Raghunathan S (2004) Configuration of detection software: A comparison of decision and game theory approaches. *Decision Anal.* 1(3):131–148.

CNN (2001) Bush announces opening of attacks. Accessed January 14, 2014, http://edition.cnn.com/2001/US/10/07/ret.attack .bush/.

Cox LA (2009) Game theory and risk analysis. *Risk Anal.* 29(8): 1062–1068.

Dixit AK, Nalebuff BJ (1991) *Thinking Strategically: The Competitive Edge in Business, Politics, and Everyday Life* (W. W. Norton & Company, New York).

Goldstein A (2000) *Deterrence and Security in the 21st Century: China, Britain, France, and the Enduring Legacy of the Nuclear Revolution* (Stanford University Press, Redwood City, CA).

Hall JR (2009) The elephant in the room is called game theory. *Risk Anal.* 29(8):1061.

Haller NM, ed. (2013) *U.S. Air Force Strategic Deterrence Capabilities in the 21st Century Security Environment: A Workshop Summary* (National Academies Press, Washington, DC).

Hao M, Jin S, Zhuang J (2009) Robustness of optimal defensive resource allocations in the face of less fully rational attacker. *Proc. 2009 Indust. Engrg. Res. Conf., Miami, FL*, 886–891.

Haphuriwat N, Bier VM, Willis HH (2011) Deterring the smuggling of nuclear weapons in container freight through detection and retaliation. *Decision Anal.* 8(2):88–102.

Hausken K, Zhuang J (2011a) Governments' and terrorists' defense and attack in a *T*-period game. *Decision Anal.* 8(1):46–70.

Hausken K, Zhuang J (2011b) The timing and deterrence of terrorist attacks due to exogenous dynamics. *J. Oper. Res. Soc.* 63(6): 726–735.

Helene C (2011) Obama announces killing of Osama Bin Laden. *New York Times* (May 1), http://thelede.blogs.nytimes.com/2011/05/ 01/bin-laden-dead-u-s-official-says/?_r=1.

Jose VRR, Zhuang J (2013) Technology adoption, accumulation, and competition in multi-period attacker-defender games. *Military Oper. Res.* 18(2):33–47.

Kahneman D, Tversky A (1979) Prospect theory: An analysis of decision under risk. *Econometrica* 47(2):263–291.

Keeney GL, von Winterfeldt D (2011) A value model for evaluating homeland security decisions. *Risk Anal.* 31(9):1470–1487.

Keeney RL (2007) Modeling values for anti-terrorism analysis. *Risk Anal.* 27(3):585–596.

Knopf JW (2010) The fourth wave in deterrence research. *Contemporary Security Policy* 31(1):1–33.

Kreps DM (1990) *Game Theory and Economic Modelling* (Oxford University Press, New York).

Levi MA (2004) Deterring nuclear terrorism. *Issues Sci. Tech.* 20(3): 70–74.

Marshall E (2012) Parsing terrorism. *Science* 336(6083):820–823.

Mas-Colell A, Whinston MD, Green JR (1995) *Microeconomic Theory* (Oxford University Press, New York).

McLay L, Rothschild C, Guikema S (2012) Robust adversarial risk analysis: A level-*k* approach. *Decision Anal.* 9(1):41–54.

Merrick JRW, McLay LA (2010) Is screening cargo containers for smuggled nuclear threats worthwhile? *Decision Anal.* 7(2): 155–171.

Miller G (2011) CIA spied on Bin Laden from safe house. *Washington Post* (May 5), http://www.washingtonpost.com/world/cia-spied-on-bin-laden-from-safe-house/2011/05/05/AFXbG31F_story.html.

Morton DP, Pan F, Saeger KJ (2007) Models for nuclear smuggling interdiction. *IIE Trans.* 39(1):3–14.

National Commission on Terrorist Attacks upon the United States (2004) *The 9/11 Commission report: Final report of the National Commission on Terrorist Attacks upon the United States* (W. W. Norton & Company, New York).

Nikoofal ME, Zhuang J (2012) Robust allocation of a defensive budget considering an attacker's private information. *Risk Anal.* 32(5):930–943.

Poundstone W (1992) *Prisoner's Dilemma* (Bantam Doubleday Dell, New York).

Powell R (2007) Defending against terrorist attacks with limited resources. *Amer. Political Sci. Rev.* 101(3):527–541.

Sandler T, Arce DG (2003) Terrorism and game theory. *Simulation Gaming* 34(3):319–337.

Schelling TC (1984) *Choice and Consequence* (Harvard University Press, Cambridge, MA).

Schelling TC (2008) *Arms and Influence* (Yale University Press, New Haven, CT).

Sevillano JC, Rios Insua D, Rios J (2012) Adversarial risk analysis: The Somali pirates case. *Decision Anal.* 9(2):86–95.

Shan X, Zhuang J (2013a) Cost of equity in homeland security resource allocations in the face of a strategic attacker. *Risk Anal.* 33(6):1083–1099.

Shan X, Zhuang J (2013b) Hybrid defensive resource allocations in the face of a partially strategic attacker. *Eur. J. Oper. Res.* 228(1):262–272.

Shan X, Zhuang J (2013c) Subsidizing to disrupt a terrorism supply chain—A four-player game. *J. Oper. Res. Soc.*, ePub ahead of print May 22, http://www.palgrave-journals.com/jors/journal/vaop/ncurrent/abs/jors201353a.html.

Trager RF, Zagorcheva DP (2006) Deterring terrorism: It can be done. *Internat. Security* 30(3):87–123.

Train K (2003) *Discrete Choice Methods with Simulation* (Cambridge University Press, Cambridge, UK).

Wang C, Bier VM (2011) Target-hardening decisions based on uncertain multiattribute terrorist utility. *Decision Anal.* 8(4): 286–302.

Wang X, Zhuang J (2011) Balancing congestion and security in the presence of strategic applicants with private information. *Eur. J. Oper. Res.* 212(1):100–111.

Yang R, Kiekintveld C, Ordóñez F, Tambe M, John R (2013) Improving resource allocation strategies against human adversaries in security games: An extended study. *Artificial Intelligence* 195: 440–469.

Zhuang J, Bier VM (2007) Balancing terrorism and natural disasters—Defensive strategy with endogenous attacker effort. *Oper. Res.* 55(5):976–991.

Zhuang J, Bier VM, Alagoz O (2010) Modeling secrecy and deception in a multiple-period attacker-defender signaling game. *Eur. J. Oper. Res.* 203(2):409–418.

Xiaojun (Gene) Shan conducts cutting-edge research in risk management against natural and man-made disasters using game theory, decision analysis, and operations research techniques. He was a postdoctoral research associate at the University of Delaware, modeling natural disaster risk management. He obtained his Ph.D. in industrial and systems engineering from the University at Buffalo, The State University of New York, and his dissertation work was honored as a finalist of INFORMS Decision Analysis Society Student Paper competition in 2010 and 2012.

Jun Zhuang is assistant professor of industrial and systems engineering at the University at Buffalo, the State University of New York (UB, or SUNY-Buffalo), since he obtained his Ph.D. in industrial engineering in 2008 from the University of Wisconsin–Madison. His long-term research goal is to integrate operations research, game theory, and decision analysis to improve mitigation, preparedness, response, and recovery for natural and man-made disasters. Other areas of interest include healthcare, sports, transportation, supply chain management, and sustainability. His research has been supported by the U.S. National Science Foundation (NSF), the U.S. Department of Homeland Security through the National Center for Risk and Economic Analysis of Terrorism Events and the National Consortium for the Study of Terrorism and Responses to Terrorism, the U.S. Department of Energy through the Oak Ridge National Laboratory, and by the U.S. Air Force Office of Scientific Research (AFOSR) through the Air Force Research Laboratory. He is on editorial board of *Decision Analysis* and is the coeditor of *Decision Analysis Today*. Jun Zhuang is the corresponding author.