

**Securing the Containerized Supply Chain:
Analysis of Government Incentives for Private Investment**

Nitin Bakshi

London Business School
Regent's Park, London
nbakshi@london.edu

Noah Gans

The Wharton School
University of Pennsylvania
gans@wharton.upenn.edu

January 2009

Working Paper # 2007-11-15

Risk Management and Decision Processes Center
The Wharton School, University of Pennsylvania
3730 Walnut Street, Jon Huntsman Hall, Suite 500
Philadelphia, PA, 19104
USA
Phone: 215-898-4589
Fax: 215-573-2130
<http://opim.wharton.upenn.edu/risk/>

THE WHARTON RISK MANAGEMENT AND DECISION PROCESSES CENTER

Established in 1984, the Wharton Risk Management and Decision Processes Center develops and promotes effective corporate and public policies for low-probability events with potentially catastrophic consequences through the integration of risk assessment, and risk perception with risk management strategies. Natural disasters, technological hazards, and national and international security issues (e.g., terrorism risk insurance markets, protection of critical infrastructure, global security) are among the extreme events that are the focus of the Center's research.

The Risk Center's neutrality allows it to undertake large-scale projects in conjunction with other researchers and organizations in the public and private sectors. Building on the disciplines of economics, decision sciences, finance, insurance, marketing and psychology, the Center supports and undertakes field and experimental studies of risk and uncertainty to better understand how individuals and organizations make choices under conditions of risk and uncertainty. Risk Center research also investigates the effectiveness of strategies such as risk communication, information sharing, incentive systems, insurance, regulation and public-private collaborations at a national and international scale. From these findings, the Wharton Risk Center's research team – over 50 faculty, fellows and doctoral students – is able to design new approaches to enable individuals and organizations to make better decisions regarding risk under various regulatory and market conditions.

The Center is also concerned with training leading decision makers. It actively engages multiple viewpoints, including top-level representatives from industry, government, international organizations, interest groups and academics through its research and policy publications, and through sponsored seminars, roundtables and forums.

More information is available at <http://opim.wharton.upenn.edu/risk>.

Securing the Containerized Supply Chain: Analysis of Government Incentives for Private Investment

Nitin Bakshi
London Business School
Regent's Park, London
nbakshi@london.edu

Noah Gans
The Wharton School
University of Pennsylvania
gans@wharton.upenn.edu

January, 2009

Abstract

To mitigate the threat that terrorists smuggle weapons of mass destruction (WMD) into the US through maritime containers, the US Bureau of Customs and Border Protection (CBP) inspects containers upon entry to domestic ports. Inspection-driven congestion is costly, and CBP provides incentives to firms to improve security upstream in the supply chain, thereby reducing the inspection burden at US ports. We perform an economic analysis of this incentive program, called Customs-Trade Partnership Against Terrorism (C-TPAT), modeling in a game-theoretic framework the strategic interaction between CBP, trading firms and terrorists. Our equilibrium results show that a properly run program can efficiently shift some of CBP's security burden to private industry. These results also suggest that there exists the possibility for CBP to use strategic delay as an incentive for firms to join. Analysis of comparative statics show that, with increasing capacity, membership in C-TPAT systematically declines.

1 Introduction

The volume and value of containerized goods entering the US through ports is enormous, and it continues to grow.¹ In 2004, \$423 billion in goods entered the US in 15.8 M containers (GAO 2007-a). Almost half of the \$2 trillion in international goods transported through the US in 2000 was shipped in containers, and the international tonnage of trade through the US is expected to double by 2020 (Greenberg et al. 2006).

Given the large numbers and value of containers entering US ports each year, concern about their use by terrorists is high. Only one of millions of containers need be compromised to cost the US billions of dollars in lost trade and endanger thousands of lives. For instance, Abt (2003) estimates that the detonation of a

¹A container is a sealed, reusable metal box (generally 20' or 40' long) in which goods are shipped by vessel, rail, or truck.

nuclear device in a port may lead to losses in the range of \$55 - 220 billion. Abt et al. (2003) estimate the economic losses from a similar bio-terrorist attack to be in the range of \$15 - 40 billion. Using a simulated war game, Gerencser et al. (2003) estimate the economic losses stemming from a coordinated “dirty bomb” attack on US ports to be \$58 billion.² The reported loss figures do not include the value of lives lost.

The Bureau of Customs and Border Protection (CBP) is responsible for ensuring the security of US ports against these types of attacks. To promote port security, CBP uses risk management techniques to screen containerized cargo for potential anomalies. Its Automated Targeting System (ATS) assigns a risk score to each container entering US waters and, based on these scores, a fraction of incoming containers is marked for rigorous inspection (GAO 2004). Containers may be subject to inspection at the port of origin, outside the US, as well as at the port of entry into the US. The focus of this paper is on the latter.

CBP is charged with securing ports with least possible hindrance to commerce, and there are inherent economic tradeoffs between the frequency and rigor with which containers can be inspected and the speed with which they can be turned around. The more containers inspected, and the more time spent inspecting each container, the smaller the probability of a hazard, such as a nuclear bomb or biological weapon, going undetected. But as the number of containers subject to detailed inspection increases, the resulting congestion can also be detrimental to trade. In the short run, unanticipated container delays can cause costly supply-chain disruptions. For example, Spencer (2004) estimates the cost of delay per day to approach 0.5% of the value of a container. Even in the long run, when inspection-induced delays can be anticipated, the extra pipeline inventory required to accommodate delays can be costly. For example, given an annual flow of \$423 billion in goods, a day of pipeline inventory is worth \$1.16 billion. At a cost of capital of 15%, that day of pipeline inventory would, in turn require \$174 M per year to finance.

Customs-Trade Partnership Against Terrorism (C-TPAT) is a federal initiative intended to induce private companies to help address this trade-off. Companies that join C-TPAT agree to take specific steps that improve the security of the containers they ship to US ports (GAO 2004). By improving the risk profile of

²A dirty bomb, also called a “radiological dispersal device” (RDD), combines a conventional explosive, such as dynamite, with radioactive material. When the conventional explosive detonates, it disperses the radioactive material, and the dispersion contaminates the surrounding area.

these containers, CBP aims to reduce the number of containers it needs to inspect and, at the same time, reduce the overall level of terrorism-related risks associated with containers entering the US. Thus, members of C-TPAT bear out-of-pocket security expenses that allow CBP to reduce costs and risks associated with container hazards and inspections.

C-TPAT membership is voluntary, and a central economic incentive for joining the program is the reduction in inspection burden to which members are entitled (C-TPAT Strategic Plan 2004). Another (more speculative) benefit is the prospect that, in the event of a disaster, C-TPAT members would be “at the head of the line” once the target port were to resume operations.

For many companies, the program’s benefits appear to outweigh its costs. More than 7,000 companies have joined C-TPAT since its inception in November, 2001 (Basham 2007). A survey of 1,240 C-TPAT members, conducted by University of Virginia on behalf of CBP, found that the respondents spent, on average, about \$54,000 per year in compliance costs as compared to about \$25,000 in security-related expenditure during the last full year before joining C-TPAT (Diop et al. 2007). The survey also found that 39% of the firms experienced a reduction in inspection frequency, while 53% reported no change. CBP is encouraged by these results because it has quadrupled inspection levels since September 11, 2001.

At the same time, both trade magazines and federal-government reviews of C-TPAT cite widespread dissatisfaction with the program (Keane 2005, GAO 2005). These reviews consistently cite two sets of concerns: 1) the benefits to participating members have not been clearly outlined; and 2) effective validation of security profiles, and regular audit of members to ensure compliance, is lacking.

Even more alarming is the apparent lack of rigor with which security inspections themselves can be conducted. Laxity in inspections have resulted in a breach in border security more than once. For example, on two occasions journalists from ABC News have managed to ship nuclear material in cargo containers into the US (Kurtz 2003). Similarly, the GAO reports that its investigators have twice used forged documents to import radioactive material through inland borders (GAO 2006-a).

The goal of this paper is to provide a modeling framework to understand the economic trade-offs embedded in container-inspection decisions and to use this framework to analyze policy initiatives such as C-TPAT.

For a private company there exists a trade-off between the cost of compliance with C-TPAT and the benefit of reduced congestion costs associated with the inspection of its containers. The US government faces a trade-off, between the security benefit derived from increased inspection of incoming containers and the adverse impact of the resulting congestion. The government must also consider the financial burden stemming from the need for additional security infrastructure. Given the actions of CBP and of trading firms, terrorists trade off the costs and benefits of infiltrating a container.

We model the interaction among CBP, trading firms, and terrorists as a multi-player sequential game, using the Principal-Agent framework. CBP (the principal) acts first, followed by the trading firms (agents) and subsequently the terrorists. CBP first sets the levels of inspection frequency and intensity (rigor), as well as parameters for the audit of members. Trading firms then decide whether or not to join C-TPAT, based on their idiosyncratic costs of complying with the security guidelines laid out in the program. Finally, terrorists choose which set of containers to target for infiltration.

Elementary considerations within our modeling approach imply that members' potential for Moral Hazard (shirking) requires CBP to audit them for compliance. Further analysis demonstrates that an equilibrium outcome exists and has the following properties:

- There is a threshold cost of compliance which separates firms that join and do not join C-TPAT.
- The optimal audit policy can be determined independently of the optimal inspection policy. CBP imposes the highest permissible penalty on a non-compliant member firm.
- The intensity of container inspections drives the surplus of non-member firms to zero.
- The expected cost to member firms, due to security measures under C-TPAT, varies with their firm-specific compliance-costs, and non-members end up with a higher expected cost than members.
- For any given (fixed) level of inspection capacity, implementation of C-TPAT results in a reduction in the costs incurred by both CBP and trading firms, relative to a Base-Case scenario, without C-TPAT.

Analysis of the game's equilibrium outcome also suggests that there may exist outcomes in which CBP

deliberately inspects some containers more frequently than is required for security purposes. This over-inspection increases congestion levels and is a means of inducing of *strategic delay*. The delay benefits CBP by providing a stronger incentive for trading firms to join C-TPAT.

Comparative statics with respect to inspection capacity show the following.

- As expected, an increase in inspection capacity results in lower expected cost for CBP.
- Surprisingly, increased capacity results in lower C-TPAT membership levels, in equilibrium.

The remainder of this paper is organized as follows. Section 2 presents a literature review. Section 3 describes a Base-Case scenario, without C-TPAT, against which the outcomes of the C-TPAT program can be compared. Section 4 models the principal-agent interactions between CBP and the trading firms and develops our equilibrium results. The role of inspection capacity is analyzed in §5. Finally, we present a brief discussion of the general scope of our work in §6.

2 Literature Review

Government documents are a comprehensive source for background information on port-security measures, such as C-TPAT, as well as inspection considerations related to border security. Details on C-TPAT can be found in the C-TPAT Strategic Plan (2004). More documents are available on CBP's web site. A comprehensive treatment of inspection issues at the various ports of entry into the US can be found in Wasem et al. (2004). Government Accountability Office (GAO) reports on maritime security (GAO 2004, GAO 2005, GAO 2006-a, GAO 2006-b) highlight implementation challenges.

Issues relating to port security and container inspections lie in the overlap between public policy and operations management, and researchers from both sides have contributed to the growing literature in the field. Some examples of policy work on this issue are Greenberg et al. (2006), Martonosi et al. (2006), and Boske (2006). Examples of the OM approach can be found in Wein et al. (2007) and Wein et al. (2006). Our work is closest in spirit to the latter.

Wein et al. (2006) develop and analyze a mathematical model of the entire multi-layered port-security

system. The paper takes a computational approach to evaluating CBP's optimal inspection strategy when faced with the risk of importation of illicit nuclear material into the US. Its aim is to prescribe the level of investment (in radiation detection equipment and personnel) required to meet a safety target, given a *predefined* flow of containers to be inspected.

In contrast, ours is an analytical treatment of the strategic interaction – between CBP, trading firms and terrorists – that *generates* the flow of containers to be inspected. Our treatment is stylized and at a higher level: it is not concerned with the specific details of the detection of nuclear threats, and our results apply to a broad range of risks, including nuclear, biological, and chemical threats.

Our model has three key components: risk assessment of containers, the impact of inspections on the economics of terrorist activity, and the effectiveness of inspections. We discuss each in turn.

CBP performs a risk assessment for terrorist threats for the entire population of incoming containers and assigns a score, that we refer to as the ATS score, to each individual container.³ This score is a probabilistic representation of the threat posed by a container. It is generated using manifest information as well as targeting rules that are based on strategic intelligence and anomalies (GAO 2004, Wasem et al. 2004, Bettge 2006). Statistics has a rich literature in screening and classification methodology, including the use of techniques such as ROC, receiver-operating curves (Fawcett 2006, Marshall and Olkin 1968). For a related treatment in OM see Shumsky and Pinker (2003). Ours is also an example of a classification problem in which the ATS score is the screening variable used to segment the container population into a “high risk” and a “low risk” category.

The decision regarding whether or not to inspect the container at the US domestic port is a function of its ATS score. The effectiveness of a container inspection can be measured through the residual probability of risk post inspection. We use a speed-accuracy-tradeoff (SAT) function to associate the expected inspection time with CBP's capacity/technology choice and the residual risk. Literature on SAT functions includes McClelland (1979), Ghylin et al. (2006) and Hopp et al. (2007).

Finally, we mention three related but distinct streams of literature. First is research on airline and pas-

³ATS stands for Automated Targeting System – the software used by CBP to help in risk assessment.

senger security, in which passengers are the analogues of shipping containers. An example from this stream is Martonosi and Barnett (2006). Second is more traditional work on the optimization of container-terminal operations. Steenken et al. (2004) provides a comprehensive survey of this literature. Third is the evolving body of work on managing supply chain disruptions. A few notable contributions on this front include Kleindorfer and Saad (2005), Sheffi (2005), and Tomlin (2006). Lee and Whang (2005) highlights the parallels between Quality Management and the creation of supply-chain security.

3 Port Security and Congestion

In this section we lay out the key features of port security that are relevant to our analysis. We also discuss the form of the container inspection policy and its impact on terrorist activity and congestion at ports. The model presented in this section is an abstraction of reality which helps us to generate insights into the trade-offs inherent in the container-inspection problem, as well as to provide a benchmark against which we can judge the effectiveness of C-TPAT.

3.1 The Shipping and Inspection Process

The flows of containers belonging to different firms follow a similar pattern. After leaving the shipper’s premises, containers are brought to the port of embarkation. From there, they are sent on an ocean-going vessel which visits a US port of debarkation. At this port of debarkation, all containers undergo some form of “passive” screening, a non-intrusive inspection which may include neutron and gamma-ray radiation monitoring. We refer to this stage as *primary inspection*.⁴ Based on prior information on the source and handling of the container, as well as the results of these tests, a fraction of these containers is tagged by CBP for more intensive, *secondary inspection*. Secondary inspection can include tests such as gamma and x-ray radiography, as well as a *devanning* of the container for a comprehensive manual inspection. For more details on inspection strategies see Wein et al. (2006). Finally, when a container is determined to be safe, it

⁴Recent initiatives suggest that, in the future, primary inspection for most US-bound containers may be completed at the port of embarkation itself (Bakshi and Gans 2008).

is allowed into the country.

3.2 Terrorist Considerations

We model terrorists as rational agents who have the means to infiltrate a container with WMD and select a target only among those containers which offer the greatest chance of success. In choosing a container, terrorists trade-off the expected benefit from an attack with the cost of planning and execution.

The cost of attack is a function of the combined abilities of the US and other foreign governments to detect and deter the planning and execution of terrorist activities in various parts of the world and is not limited to the context of maritime security alone. Two major costs that terrorists would incur are the cost of procuring a weapon of mass destruction (WMD) and the cost of recruitment and training of a team that carries out the attack. In the context of CBP's inspection problem, we model the cost of mounting an attack, c_a , as exogenously specified.

The benefit that terrorists derive from their efforts depends on the eventual disposition of the container. If the contraband *escapes* detection, then it may be used for a large-scale terrorist attack, at which point the US suffers loss L_e (e for escapes detection). If the contraband is *found* inside a container before it crosses the US borders, then the US suffers losses L_f (f for found). We note that the discovery of WMD in a maritime container can, itself, trigger economic losses.⁵ For ease of exposition, we assume that US's loss is the terrorists' gain, and to avoid trivial results we assume that $L_e > L_f$.

Given that we are considering only large-scale acts of terrorism, we posit that terrorists have the wherewithal to launch only one such attack in the period of interest, e.g., one year.⁶

3.3 Risk Scoring and Deterrence Threshold

CBP's Automated Targeting System (ATS) uses manifest information and targeting rules, based on expert judgment and historical shipment information, to determine the probability that a container poses a "high

⁵For example, there may be a port slowdown or lockdown until the source of the security breach is discovered.

⁶A different choice for the period of interest would not have a qualitative impact on the insights generated. Since the US security budget is determined annually, working with a 1-year horizon seems natural.

risk” and should be scrutinized thoroughly. ATS scores drive inspection decisions at the port of entry.

We model the ATS score as the product of two factors. First, we let b denote the (exogenously specified) base-rate probability of a terrorist attack in the period of interest. For instance, a recent congressionally mandated report (Graham et al. 2008) estimates a higher than 50% chance of a WMD attack launched by terrorists, over the next 5 years. Previous estimates of the probability of such an attack include the work by Lugar (2005). In turn, we define the *risk score*, x , to be the conditional probability that, given a container conceals terrorist weapons, it would escape detection by security precautions in place up through the primary inspection at the port of debarkation: $P\{\text{no alarm} \mid \text{threat}\}$. Thus, ATS score = bx .

In the analysis that follows, we do not vary the base rate, b , across containers, an assumption that presumes that terrorists have exogenously decided which containers are more or less likely to be successfully infiltrated. Rather, we explicitly model the terrorists decision regarding which containers might be most profitably compromised, a decision that depends on its risk score, x . In turn, the set of containers that the terrorists target for possible infiltration emerges through the equilibrium outcome determined by our analysis.

Given our assumptions, it is immediate that, if $c_a > L_f$, then there exists a so-called “deterrence threshold” for the risk scores. For containers with a risk score below this deterrence threshold, the probability that a compromised container avoids detection is low enough that the cost to terrorists of trying to infiltrate the container is greater than the expected benefit:

$$\tau = \sup\{x \mid xL_e + (1 - x)L_f \leq c_a\}. \quad (\text{DT})$$

Thus, these containers do not provide terrorists with a high enough chance of success to make the effort of introducing a hazard into them worthwhile. In turn, they are considered to be without threat. (For details on this approach see Martonosi and Barnett (2006).)

If $c_a < L_f$, then no amount of inspection will deter terrorists from attempting to infiltrate a container, and the best that outcome that CBP can attain is to find an infiltrated container with probability one. In the exposition that follows, we will assume that $L_e > c_a > L_f$, so that CBP’s aim is to eliminate the terrorist threat by reducing terrorists’ expected gains to c_a . If c_a were less than L_f , then CBP would analogously

aim to reduce terrorists' expected gains to L_f .

If primary inspection does not trigger an alarm, and a container's risk score falls below some threshold, bounded below by τ , then the container is not inspected further. If, however, one of these conditions does not hold, then CBP tags the container for more intensive secondary inspection.

Figure 4 in Appendix A pictures an example of the CDF of risk scores, $G_n(x)$, with $x \in [0, 1]$. We denote the associated density function as $g_n(x)$. For simplicity we assume that $g_n(x) > 0, \forall x \in [0, 1]$. Here, the subscript "n" is used to signify firms that are not members of C-TPAT. In this section, which analyzes a "base case" without C-TPAT, all firms are non-members, and in §4 we distinguish members from non-members by using the subscript "m."

3.4 Secondary Inspection Time and Residual Risk

Huizenga (2005) notes that, even though current technology is quite effective in detecting most nuclear material, it is less effective in detecting certain configurations of shielded highly-enriched uranium. The diversity of the nuclear threat, in conjunction with often hard-to-detect threats from chemical and biological weapons, requires CBP to determine not only which containers to inspect, but also the rigor of the inspection process for containers identified as risky.

The effectiveness of inspections depends on the time and care with which they are conducted. As we noted in the introduction, Kurtz (2003) and GAO (2006-a) report instances in which lax inspections allowed nuclear materials to be clandestinely slipped into the US. USA TODAY (2007) and Ghylis et al. (2006) note analogous problems with the screening of passengers and baggage at airports.

For containers, the time required for secondary inspections can range widely. For example, the time needed to properly interpret x-ray images may vary. More significantly, the rigor with which a container is "devanned" can extend broadly: from a cursory look inside the back doors, to a more thorough emptying out of a center "aisle" through which inspectors move, to the removal of all contents stored within the container, even to the opening and inspection of the cartons or flats that have been removed.

Thus, a key decision that CBP makes is the extent or rigor of inspection of "high risk" containers. We let

S denote the time required to perform a secondary inspection and ε denote the residual probability that there exists a hazard that remains undetected after secondary inspection. We then use a speed-accuracy-tradeoff (SAT) function to model expected inspection time as a function of capacity/technology choice and ε :

$$S = \psi(\varepsilon, \kappa) + \phi, \quad (1)$$

where κ represents the appropriately scaled inspection capacity. The random variable, ϕ , has mean 0 and variance σ^2 , which captures the randomness introduced by container-specific characteristics, such as the type of goods being shipped and the quality of documentation of manifest information. From (1) we have $E(S) = \psi$, and $E(S^2) = \psi^2 + \sigma^2$.

The inspection capacity is meant to represent a composite of equipment and human resources devoted to the secondary inspection process. In this section and §4 we assume that κ is fixed. In §5 we then analyze the impact of capacity, κ , on the equilibrium outcome.

We make three mild sets of assumptions concerning the form of $\psi(\varepsilon, \kappa)$. First, time spent on inspection is strictly decreasing in both the residual risk and capacity: $\psi_\varepsilon \equiv \partial\psi/\partial\varepsilon < 0$ and $\psi_\kappa \equiv \partial\psi/\partial\kappa < 0$. To appreciate the motivation for the latter, consider the scenario wherein 2 inspectors would be able to inspect a devanned container faster than just 1 inspector acting alone, while maintaining the same residual risk, ε , constant across the two scenarios. Second, for any finite capacity level, κ , we assume that $\psi(1, \kappa) = 0$ and $\lim_{\varepsilon \rightarrow 0} \psi(\varepsilon, \kappa) = \infty$.

Remark 1 *As an example, consider the following specific functional form for ψ :*

$$S = -\frac{\ln \varepsilon}{\kappa} + \phi. \quad (2)$$

This functional form satisfies the first two of our assumptions. It also is consistent with the classic model for SATs presented in McClelland (1979), as well as with recent higher-level models of speed-accuracy tradeoffs used in the OM literature (see Hopp et al. (2007)). Similar tradeoffs are observed by Ghylis et al. (2006) for the problem of passenger-baggage screening.

3.5 Container Inspection Policy and Congestion

We model a policy in which CBP inspects containers with risk score x , with probability $p(x)$.⁷ We represent the fraction of containers that are tagged for secondary inspection by θ_n and observe that

$$\theta_n = \int_0^1 p(x)g_n(x)dx. \quad (3)$$

Let Λ denote the “raw” (or “base”) arrival rate of containers into a port. Given that containers are marked for secondary inspection with probability θ_n , the resulting effective arrival rate for secondary inspection is $\lambda = \Lambda\theta_n$.

We model the process of secondary inspections as an M/G/1 queue, with Poisson arrival rate $\Lambda\theta_n$, service times S , as determined by (1), and expected delay in queue:

$$E(D) = \frac{\lambda E(S^2)}{2(1 - \lambda E(S))} = \frac{\lambda(\psi^2 + \sigma^2)}{2(1 - \lambda\psi)}. \quad (4)$$

The queuing discipline followed is first-come, first-served.

The M/G/1 queuing model is an approximation of the real world, where more than one station might process the containers tagged for secondary inspection. This assumption allows us to include an analytically tractable expression for expected delay within our broader economic analysis. Furthermore, in the current context – in which a small number of servers is highly utilized – the single-server assumption is reasonable, as is explained in Kollerstrom (1974) and also Chapter 11-10 (p. 518) in Wolff (1989).

Suppose that firm i incurs an idiosyncratic per-container delay cost, d_i , per unit of time and that the average dollar value per container is r_i for firm i . Then we assume that waiting cost per dollar of revenue, $w = d_i/r_i$, is a constant, for all i . To the extent that delay costs are driven by the cost of capital (and other value-driven factors) such a constant ratio is a natural assumption. For example, see Martonosi et al. (2006).

⁷Another potential degree of freedom is offered by modeling risk-score-specific inspection protocols $\varepsilon(x)$, but our limited experience with inspection systems suggests that this scheme would be very difficult to operationalize.

3.6 Analysis of the Base Case

The Base Case refers to the scenario without C-TPAT. Containers come into a port at arrival rate Λ and are picked up for secondary inspection at a rate $\lambda = \Lambda\theta_n$. We model the interaction between CBP and terrorists as a Stackelberg game (Laffont and Martimort 2001). CBP acts as the leader and decides its inspection policy first: $\{p(x) \mid x \in [0, 1]\}$ and ε_b , the Base-Case residual risk. Terrorists act next to determine which container to target for infiltration. We assume that CBP and terrorists are risk neutral

As is typical in the backward induction process that leads to the characterization of an equilibrium outcome in a Stackelberg game, we first determine the “best response” of terrorists. From the terrorists’ point-of-view, it is optimal to target the container that offers them the most favorable prospects. If there is more than one such container, then the terrorists’ equilibrium strategy will be to target any one of these containers for infiltration, with equal likelihood. Given a container with risk score x , inspection probability $p(x)$, and residual risk ε_b , the expected benefit to the terrorists from targeting it is

$$x\{p(x)[\varepsilon_b L_e + (1 - \varepsilon_b)L_f] + (1 - p(x))L_e\} + (1 - x)L_f. \quad (5)$$

We next determine CBP’s equilibrium strategy. Given capacity, κ , CBP’s choice of residual risk, ε_b , then yields an expected inspection time, $\psi(\varepsilon_b, \kappa)$. CBP’s objective is to choose an inspection policy, $\{p(x) \mid x \in [0, 1]\}$ and ε_b , to minimize the expected losses due to a container harboring a terrorist threat entering a port. Therefore, its objective is

$$\min_{\varepsilon_b, p(x): [0,1] \rightarrow [0,1]} \left[O_P = \max_{x \in [0,1]} x\{p(x)[\varepsilon_b L_e + (1 - \varepsilon_b)L_f] + (1 - p(x))L_e\} + (1 - x)L_f \right]. \quad (6)$$

While this objective naturally leads CBP to make ε_b as small as possible, for any choice of $p(x)$, concern for the economic viability of the trading firms that use the port prevent it from simply setting $\varepsilon_b = 0$.

Specifically, firm i is willing to participate in ocean trade as long as, on a per container basis, the expected cost incurred from inspection-induced congestion is bounded above by some fraction ($\Delta > 0$) of the container’s dollar value, $r_i: \theta_n d_i(E(D) + E(S)) \leq \Delta r_i$. Since $d_i/r_i = w$, we can rewrite the inequality as

$$\theta_n w(E(D) + E(S)) \leq \Delta. \quad (\text{IR}_b)$$

The above constraint acts as an upper bound on the expected cost that a firm is willing to bear. It is the natural analogue of the participation or “individual rationality” constraint used in economic theory.

The effective arrival rate at the secondary inspection facility is $\lambda = \Lambda\theta_n$. From (4) we see that (IR_b) requires that $\frac{\lambda(\psi^2 + \sigma^2)}{2(1 - \lambda\psi)} + \psi \leq \frac{\Delta}{w\theta_n}$, which implies $\lambda \leq \frac{2\Delta}{\sigma^2 w\theta_n}$ must be satisfied for the mean service time to be non-negative. A sufficient condition for this to be the case is $\Lambda \leq \frac{2\Delta}{w\sigma^2}$, and we assume that this condition is met. Similarly (4), (IR_b) , and $\Delta < \infty$ imply that, if $\theta_n > 0$, then $\rho \equiv \lambda\psi < 1$. Thus, any feasible solution, with $\theta_n > 0$, must have a stable inspection queue.

If CBP had enough inspection capacity, then it would inspect each container down to a residual probability, ε_b , that eliminates terrorist threat, or equivalently $O_P = c_a$, without creating excessive congestion for trading firms. We rule out this possibility by assuming that CBP has limited inspection capacity. This assumption is consistent with the conclusion in the various GAO and media reports that review maritime security. So, for given inspection capacity (or budget), the optimization problem faced by CBP is as follows:

$$O_P^* = \min_{\varepsilon_b, p(x): [0,1] \rightarrow [0,1]} \{O_P \mid \theta_n w(E(D) + E(S)) \leq \Delta; 0 \leq \varepsilon_b \leq 1\}. \quad (7)$$

Our first result characterizes basic properties of the optimal solution.

Lemma 1 Suppose $\Lambda \leq \frac{2\Delta}{w\sigma^2}$.

- i. If $\exists \varepsilon < 1$ such that $\psi(\varepsilon, \kappa) < \infty$, then any optimal solution to (7) has $0 < \varepsilon_b < 1$ and $\theta_n > 0$.
- ii. If, in addition, $O_P^* > c_a$, then the (IR_b) constraint is binding in the optimal solution.

Proof All proofs are in Appendix B. □

Part (i) shows that we can always assume that, at optimum, ε_b and θ_n are interior to the problem boundaries. That is, if CBP has inspection capacity then it will use it. Part (ii) shows that if CBP does not have enough capacity to drive its expected losses – and the terrorists expected gains – down to c_a , then it will inspect as intensively as possible, consuming all of the trading firms’ surplus. To see this, note that for a given level of container traffic, λ , determined by a certain choice of $p(x)$, the LHS of (IR_b) is monotonically decreasing in ε_b . Since the objective function in (6) is strictly increasing in ε_b , for a fixed λ ; it would be

optimal for CBP to set ε_b to be as low as possible, i.e., drive the expected cost incurred by a trading firm up to its upper bound.

The fact that (IR_b) is binding in equilibrium allows us to precisely characterize the optimal form of $p(x)$:

Proposition 1 Let $\mathcal{A}_b^* \triangleq \{x|x \text{ offers the maximum expected benefit to terrorists}\}$ and $x_{lb}^* \triangleq \inf \mathcal{A}_b^*$. Suppose $O_P^* > c_a$ and that $\Lambda \leq \frac{2\Delta}{w\sigma^2}$. Then the optimal form of $p(x)$ is given by the following:

- i. $p(x) = \begin{cases} 0, & x \in [0, x_{lb}^*] \\ \frac{1-x_{lb}^*}{1-\varepsilon_b^*}, & x \in [x_{lb}^*, 1] \end{cases}$
- ii. $O_P^* = x_{lb}^* L_e + (1 - x_{lb}^*) L_f$;
- iii. $\varepsilon_b^* \leq x_{lb}^*$.

If there are multiple solutions that satisfy the conditions in Lemma 1 and Proposition 1, then the solutions with the smallest value of x_{lb}^* are the relevant candidate optimal solutions. This is because we know from Proposition 1 that $O_P^* = x_{lb}^* L_e + (1 - x_{lb}^*) L_f$, which is strictly increasing in x_{lb}^* . Indeed, if the candidate optimal solutions are such that ε_b varies continuously in x_{lb}^* , then, with limited inspection capacity, the optimal solution is unique, with $\varepsilon_b^* = x_{lb}^*$ and $p(1) = 1$.⁸ As it stands, we find that the optimal $p(x)$ takes the form depicted in Figure 1, below.

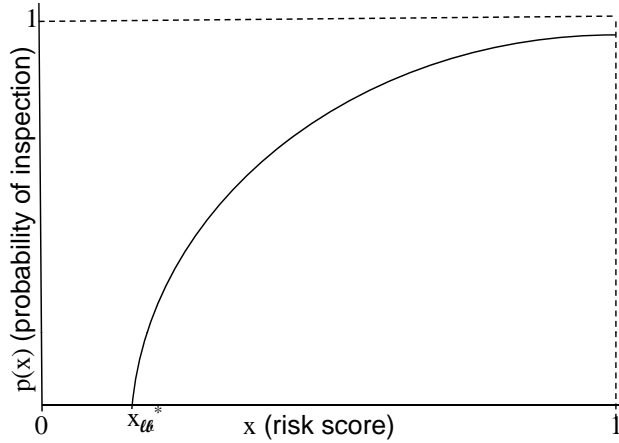


Figure 1: Optimal Form of the Inspection Probability, $p(x)$, for the Base Case.

The intuition behind the choice of the optimal form of $p(x)$ is that CBP tries to equalize its risk exposure across containers, such that the expected benefit offered to terrorists is the same for every container. If the

⁸For sufficient conditions under which ε_b varies continuously in x_{lb}^* , see Appendix C.

terrorists were not indifferent among the containers, then CBP could lower its expected cost by reducing the value of $p(x)$ for containers that offer lower than maximum expected benefit (and hence would be ignored by terrorists in equilibrium) until either the expected benefit increases to the maximal value, or $p(x) = 0$.

The results of this Base Case serve as a benchmark with which to compare and contrast the results of the security scenario with C-TPAT, as described in Section 4.

4 C-TPAT

4.1 Background on C-TPAT

CBP asks C-TPAT members to ensure the integrity of their supply chain security practices and to communicate and verify the security practices of their supply chain partners (GAO 2005). CBP specifies standards, such as infrastructure requirements and procedures to be followed while preparing a container for shipping. For example, a C-TPAT member may be required to secure its premises with patrols and video surveillance, undertake an extensive exercise in risk assessment and take remedial measures based on the results, use electronic tamper-proof seals on its containers, verify the background of all employees and contractors working for it, and adhere to other guidelines in the program.

C-TPAT and Security-Related Effort

Whether or not a firm joins C-TPAT, it may perform some due diligence of its own accord, to prevent pilferage, ensure visibility of the container during its journey to its destination, or facilitate reconciliation of contents upon delivery. To ensure compliance with C-TPAT guidelines, a firm may need to exert additional effort. We normalize the effort exerted by non-member firms to be 0 and define $\gamma_i \in [0, \infty)$ to be the extra cost per container that firm i incurs to comply with C-TPAT guidelines.

Risk Profile of Members

As in §3, the CDFs $G_m(x)$ and $G_n(x)$ describe the distribution of risk scores in the container populations of C-TPAT members and non-members, respectively. The distribution $G_n(x)$ is the same as that in the Base Case. We assume that the CDFs are differentiable, with corresponding density functions $g_m(x)$ and $g_n(x)$.

Once again, we assume that $g_n(x), g_m(x) > 0, \forall x \in [0, 1]$.

Given C-TPAT's aim of motivating companies to reduce container risk, we expect the distribution of G_m and G_n to differ, and we assume that $G_m(x) > G_n(x)$, for all $x \in [0, 1]$. This relationship is referred to as a strict First Order Stochastic Dominance (FOSD) ordering (Shaked and Shanthikumar 1994).

Fraction of Containers Inspected

Whether or not a firm joins C-TPAT, the flow of its containers follows a similar pattern. θ_m represents the fraction of a C-TPAT member's (m for members) containers that undergo more intensive secondary inspection. Likewise θ_n represents the fraction of a non-member's (n for non-members) containers that are tagged for secondary inspection. The values of θ_m and θ_n are functions of $p(x)$ - the probability of tagging a container with risk score x , for secondary inspection - and the density functions $g_m(x)$ and $g_n(x)$ respectively. While the value of θ_n is as described in (3), the value of θ_m is similarly defined:

$$\theta_m = \int_0^1 p(x)g_m(x)dx. \quad (8)$$

Observe that, for non-decreasing $p(x)$, the strict FOSD ordering implies that $\theta_m < \theta_n$. We will verify in §4.2 that this is indeed the case for the optimal choice of $p(x)$. Thus, by joining C-TPAT, a firm improves its risk profile, and the improvement leads to a reduction in the fraction of its containers that undergo secondary inspection. The savings associated with this reduction are an important incentive to join.

Audit of Members

To prevent C-TPAT members from shirking (i.e., not exerting the extra security effort required of members), CBP may conduct an audit of member firms. The audit determines whether or not the guidelines laid out in C-TPAT are being diligently followed. Use of damaged electronic container seals, use of contract labor without background checks, and absence of video surveillance at facilities are examples of the types of lapses that might be encountered during an audit. We assume that, once an audit has been undertaken, it can be determined with certainty whether or not a firm has shirked.

CBP audits member firms with an annual relative frequency, q , and it then imposes a penalty if a deviation is discovered. The audit frequency can also be thought of as the fraction of C-TPAT members that are audited in any given time period. We denote the per-container cost of auditing a member firm i as $c_i(q)$,

with $c'_i(q) \geq 0$. For example, a firm with an annual volume of container traffic, V_i , incurs an expected cost of audit of $qc_i(q)V_i$, which translates to a per-container expected cost of $qc_i(q)$. Similarly, we let P_i represent the per-container allocation of the penalty assessed should firm i be found to be shirking. (The total penalty is assessed on the firms's container traffic, from the start of the period until the time shirking is discovered, since this is the set of containers that benefited from a lower inspection frequency.) This allows us to account for all costs on a per-container basis. A shirking firm is also relegated to non-member status for the remainder of the period.⁹

We model audit costs as being borne by trading firms. Specifically, the SAFE Port Act (2006) mandates a pilot for a third-party audit program. Under this scheme, CBP-authorized third-party auditors conduct audits, and C-TPAT participants pay for the audits. The third parties need to be audited by CBP in turn. Since only a small fraction of the staff resources are required to audit the auditors, relative to auditing the member firms directly, we make the simplifying assumption that the cost associated with the auditing of third-party auditors is fixed, i.e., it is independent of the membership level in C-TPAT. Hence, we do not explicitly include it in our model.

Such a third-party scheme is attractive to CBP for two reasons. First, with an increasing number of firms signing up for C-TPAT, CBP is falling short of staff required to effectively validate membership and later audit firms (GAO 2005).¹⁰ Second, CBP auditors do not have access to certain trade lanes in the international supply chain, for political and sovereignty reasons. CBP launched its pilot program for third-party audits in June 2007 (Basham 2007).¹¹

4.2 A Principal-Agent model of C-TPAT

We model the interaction between CBP, trading firms, and terrorists as a multi-player sequential game. The terrorists act last and their equilibrium strategy is to target one of the containers that offers maximum ex-

⁹An equivalent penalty scheme would be to penalize all of the deviating firm's container traffic in the period of interest, but then allow the firm to sign-up as a member again, immediately after failing the audit.

¹⁰In CBP's parlance "revalidation" of C-TPAT membership is equivalent to an audit, as described in this paper.

¹¹Similar third-party audit mechanisms have been used successfully in other contexts, such as the promotion of industrial safety and the enforcement of environmental regulations (Kunreuther et al. 2002).

pected benefit, as explained in the Base-Case analysis in §3.6. Incorporating the best response of terrorists in this manner, the interaction between CBP and the trading firms can thereafter be thought of as a Stackelberg game in which CBP (the principal) acts as the leader and the trading firms (agents) are followers. Both CBP and the trading firms are assumed to be risk neutral.

CBP first decides on the secondary inspection parameters θ_m, θ_n and ε , and the audit parameters q and P_i . It then offers the contract $\{q, P_i, \varepsilon, \theta_m\}$ to members and $\{\varepsilon, \theta_n\}$ to non-members who use the port facilities. Firms decide whether or not to join C-TPAT, based on their respective costs of compliance and the expected congestion costs due to secondary inspection. Once firms have decided whether or not to join C-TPAT, members are expected to comply with the security-related guidelines prescribed in the agreement. A pictorial representation of the sequence of events is presented in the Figure 2 below.

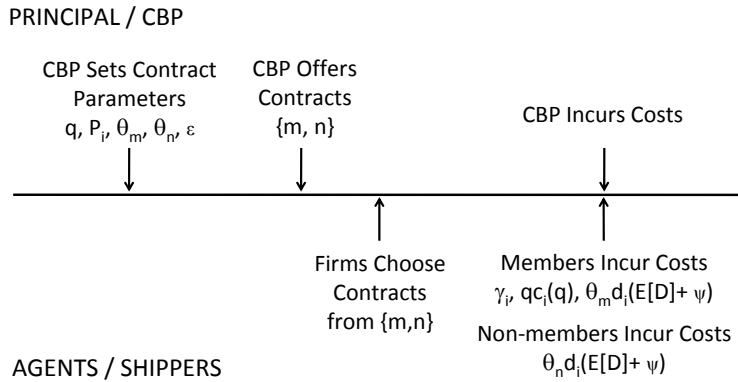


Figure 2: Dynamics of the Principal-Agent Stackelberg Game.

Agent's Problem

The decision of whether or not to join C-TPAT is largely governed by the agents' cost of compliance with the program. Firms with cost of compliance γ_i are faced with two choices: either sign up for C-TPAT at an expected per-container expense of $\gamma_i + qc_i(q)$ and experience an expected system waiting time of $E(D) + E(S)$ with probability θ_m , or remain a non-member and experience an expected wait of $E(D) + E(S)$ with probability, θ_n . The condition that must be satisfied for a firm to sign up for C-TPAT is therefore

$$\gamma_i + qc_i(q) + \theta_m d_i(E(D) + E(S)) \leq \theta_n d_i(E(D) + E(S)). \quad (9)$$

The above condition necessarily requires that $\theta_n \geq \theta_m$. Observe that the expected delay, $E(D)$, is the same

on both sides of the inequality. Implicitly, we are assuming that each firm is an infinitesimal player, whose individual decisions do not impact the overall congestion levels in the system. This assumption is similar in spirit to the treatment in a Wardrop Equilibrium (Altman et al. 2006).

Recalling that the dollar value of revenue associated with a container is r_i for firm i , we now define $\alpha(q) \equiv (\gamma_i + qc_i(q))/r_i$, as member i 's cost of compliance per dollar of revenue, or simply the compliance cost. For $\gamma_i \in [0, \infty)$ we see that $\alpha(q) \in [\frac{qc_i(q)}{r_i}, \infty)$. For fixed q , we can also define the cumulative distribution function (CDF) $F(\alpha)$ to be the fraction of the total volume of containers shipped to the US which come from firms with a compliance cost no more than α . We assume that for any fixed q , $F(\alpha)$ is differentiable everywhere, and $dF(\alpha) = f(\alpha)d\alpha$ represents the relative likelihood that a container comes from a firm with compliance cost α . Implicit here, again, is the assumption that each firm contributes an infinitesimal amount to the cumulative volume of container trade.

For a given $E(D)$, let α_t denote a threshold compliance cost (t for threshold), below which (9) is satisfied and above which it is not. In turn, for a given α_t , the fraction of C-TPAT certified containers is $F(\alpha_t)$, which yields the effective arrival rate at the secondary-inspection queue:

$$\lambda = \Lambda[F(\alpha_t)\theta_m + (1 - F(\alpha_t))\theta_n]. \quad (10)$$

Substituting this value of λ into (4) yields the corresponding expression for expected delay, $E(D)$.

As described above, the definitions of α_t and $E(D)$ are circular, since each depends on the other. Nevertheless, we can show that, for given q , θ_m , θ_n and ε , these two equilibrium quantities are well defined.

Proposition 2 *For given q , θ_m , θ_n and ε , the threshold compliance cost exists, is unique, and is given by*

$$\alpha_t = (\theta_n - \theta_m)w(E(D) + E(S)). \quad (11)$$

The Principal's Problem

As before, the principal tries to minimize the expected cost of a disaster:

$$O_P^* = \min_{\varepsilon, P_i, q \in [0,1], p(x): [0,1] \rightarrow [0,1]} \left[\max_{x \in [0,1]} x \{p(x)[\varepsilon L_e + (1 - \varepsilon)L_f] + (1 - p(x))L_e\} + (1 - x)L_f \right]. \quad (12)$$

The solution to the principal's problem should be such that it provides the appropriate incentives for the

agents to participate without shirking.

Participation Constraint for Agents

The participation constraint for non-members remains the same as described in condition (IR_b) in the Base Case. Satisfying (IR_b) is also sufficient to ensure participation of member firms, as is apparent from (9), provided $\theta_n \geq \theta_m$.

Incentive-Compatibility Constraint for Agents

A firm that has signed up for membership in C-TPAT may find it beneficial to shirk by not putting in the effort required for compliance with C-TPAT guidelines while, at the same time, continuing to enjoy reduced congestion costs afforded to members only. An incentive compatibility constraint ensures that such a situation does not arise. The principal uses audit as a means to achieve incentive compatibility: a member firm i which fails an audit is penalized an amount P_i , which is bounded above by some $B_i < \infty$.

The upper bound, B_i , is set to the benefit accruing to the participating firm from joining C-TPAT. This captures the idea that the penalty cannot be larger than the non-compliant agent's benefit from its false announcement. (See page 123 in Laffont and Martimort (2001).) We consider a more general upper bound, a constant multiple β (≥ 1) of the benefit from non-compliance, minus the cost of the audit itself.¹² A member firm that fails an audit, no matter when it is conducted, forgoes the benefit accrued due to member status and is relegated to non-member status for the rest of the period. Thus, for $\alpha(q) \in [0, \alpha_t]$, where $\alpha(q) = (\gamma_i + qc_i(q))/r_i$, condition (9), along with incentive compatibility considerations, implies that for each member firm i :

$$\gamma_i + qc_i(q) + \theta_m d_i(E(D) + E(S)) \leq (1 - q)[\theta_m d_i(E(D) + E(S))] + q[\theta_n d_i(E(D) + E(S)) + c_i(q) + P_i], \quad (13)$$

where:

$$0 \leq P_i \leq B_i = \beta(\theta_n - \theta_m)d_i(E(D) + E(S)) - c_i(q). \quad (14)$$

We assume that β is large enough so that $B_i \geq 0$. Dividing (13) by r_i , we observe that, without audit, $q \equiv 0$ and (13) can be satisfied only for $\alpha = 0$. Thus, without some form of audit (or analogous mechanism), CBP

¹²Once a firm is audited, it has to incur the audit cost irrespective of whether or not it failed the audit. This expected cost is already accounted for in the compliance cost.

cannot prevent shirking among member firms.

In fact, CBP has an incentive to make the audit penalty, P_i , as large as possible.

Proposition 3 *In equilibrium, $P_i^* = B_i$.*

Thus, at optimum P_i will achieve its upper bound B_i . In the economics literature, this is known as the *principle of maximal punishment*. (See pages 121-126 in Laffont and Martimort (2001).) Indeed, a finite upper bound B_i is required to make the audit mechanism reasonable, lest CBP impose an infinite penalty with probability zero.

Using (11), (13) and (14) to simplify the incentive-compatibility (IC) constraint, we obtain:

$$\frac{\gamma_i + qc_i(q)}{r_i} \equiv \alpha(q) \leq q(1 + \beta)\alpha_t(q) \quad \forall \alpha(q) \leq \alpha_t(q). \quad (\text{IC})$$

In turn, we have:

Proposition 4 *In the Stackelberg game between CBP and trading firms, the optimal fraction of members to be audited is:*

$$q^* = \frac{1}{1 + \beta} \quad \text{if } \alpha_t(q^*) > 0; \quad q^* = 0 \quad \text{if } \alpha_t(q^*) = 0.$$

Proposition 4 implies that the value of q^* is independent of the choice of $p(x)$ and ε . Thus, CBP can fix q^* and then optimize over $p(x)$ and ε alone. Also, given q^* , we have $\alpha \equiv (\gamma_i + q^*c_i(q^*))/r_i$, and the compliance-cost distribution function $F(\alpha)$ is well defined.

Proposition 4 also provides insight into the effectiveness of audit practises. For example, suppose $\beta = 1$, so that the penalty for shirking equals the expected benefit from joining the program. This implies that $q^* = 0.5$, in which case a 50% chance of audit is optimal.

Thus, the optimization problem faced by the principal is

$$O_P^* = \min_{\varepsilon, p(x): [0,1] \rightarrow [0,1]} \left[\max_{x \in [0,1]} x \{p(x)[\varepsilon L_e + (1 - \varepsilon)L_f] + (1 - p(x))L_e\} + (1 - x)L_f \right],$$

s.t.

$$\theta_n w(E(D) + E(S)) \leq \Delta, \quad (\text{IR})$$

$$\theta_n \geq \theta_m, \quad (\text{IC}')$$

$$0 \leq \varepsilon \leq 1, \quad (\text{FEAS})$$

and we obtain an initial characterization of the equilibrium behavior induced by C-TPAT that parallels that of the base case.

Lemma 2 Suppose $\Lambda \leq \frac{2\Delta}{w\sigma^2}$.

- i. If $\exists \varepsilon < 1$ such that $\psi(\varepsilon, \kappa) < 1$, then any equilibrium solution has $0 < \varepsilon < 1$ and $\theta_n > 0$.
- ii. If, in addition, $O_P^* > c_a$ then the (IR) constraint is binding in equilibrium.

The intuition for this result is similar to that for the result in Lemma 1, and with it we can more completely characterize the optimal form of $p(x)$. To do so, we first define the following quantity:

$$\zeta = \frac{f(\alpha_t)\alpha_t + F(\alpha_t)}{f(\alpha_t)\alpha_t(1 - \frac{\alpha_t}{\Delta}) + F(\alpha_t) - 1 - \left(\frac{\Delta\lambda(1-\lambda\psi)}{\theta_n^2 w \Lambda E(D)}\right)}. \quad (15)$$

Proposition 5 Suppose $\Lambda \leq \frac{2\Delta}{w\sigma^2}$. Let $\mathcal{A}^* \triangleq \{x | x \text{ offers the maximum expected benefit to terrorists}\}$, and let $x_l^* \triangleq \inf \mathcal{A}^*$. If $O_P^* > c_a$ then we have the following.

- i. $\forall x \in \mathcal{A}^*$, $p(x)$ is strictly increasing in x , and:

$$x\{p(x)[\varepsilon L_e + (1 - \varepsilon)L_f] + (1 - p(x))L_e\} + (1 - x)L_f = k, a \text{ constant}. \quad (16)$$

- ii. If $\zeta \leq 0$, then:

- a) $\forall x \in [0, x_l^*]$, we have $p(x) = 0$;
- b) $\forall x \in [x_l^*, 1)$, we have $x \in \mathcal{A}^*$;
- c) $\forall x \in (x_l^*, 1)$, we have $p(x) \in (0, 1)$, and the relationship in (16) is satisfied.

- iii. If $\zeta > 0$, then $\forall x \in [0, 1)$:

- a) If $g_n(x)/g_m(x) < \zeta$, then $p(x)$ behaves as in (ii);
- b) If $g_n(x)/g_m(x) \geq \zeta$, then $p(x) = 1$.

iv. $\theta_n > \theta_m$.

The left panel of Figure 3 provides an illustration of the general form of the optimal $p(x)$.

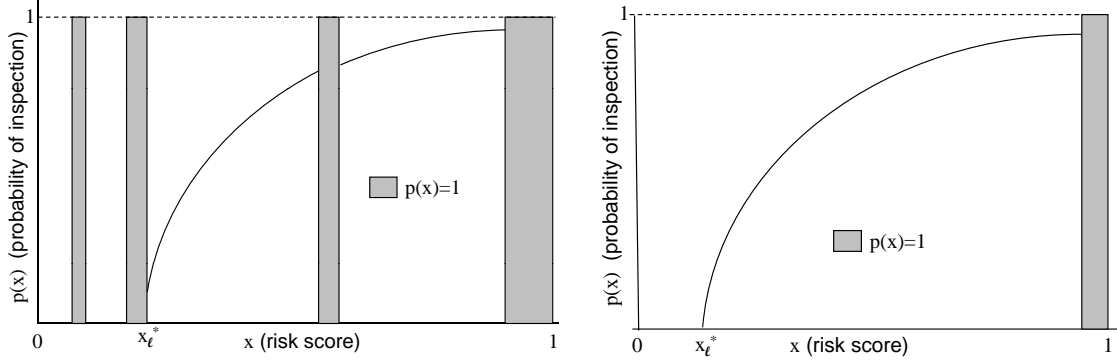


Figure 3: Optimal Form of Inspection Probability, $p(x)$

Once again, the intuition behind the nature of the optimal form of $p(x)$ is similar to that for the result in Proposition 1: CBP tries to equalize its risk exposure across all containers in order to minimize its expected cost. However, in this case it is possible that CBP makes strategic use of its inspection capacity to influence the membership level in C-TPAT. When $g_n(x)/g_m(x) > \zeta$ an increase in $p(x)$ drives up congestion costs enough for non-members, relative to members, leading to a higher α_t and additional participation. When the condition is satisfied, $p(x) = 1$, even though these containers offer less than maximum expected benefit to terrorists. The additional membership benefits CBP enough to offset the additional burden of inspecting containers that are not the terrorists' preferred targets. This is an instance of the use of *strategic delay* by CBP (Afeche 2006).

Corollary 1 *Suppose $F(x) > 0$, $\forall x > 0$. If $O_P^* > c_a$, then implementation of C-TPAT results in strictly lower costs for CBP and weakly lower costs for the trading firms, relative to the base case.*

This result confirms the economic intuition behind C-TPAT. The main purpose of the program is to transfer the burden of securing the containerized supply chain, in a cost-effective manner, from the congestion-causing step of secondary inspections at US ports to security investments by importers further upstream in the supply chain. If implemented judiciously, it ought to be a win-win solution for CBP and the trading community.

Even though we have made progress towards characterizing the optimal solution for the case with strict FOSD ordering between the distribution $G_n(x)$ and $G_m(x)$, we can obtain sharper results if we assume the stronger condition of strict monotone likelihood ratio (MLR) ordering, i.e., $g_n(x)/g_m(x)$ is strictly increasing in x (Shaked and Shanthikumar 1994).¹³ The MLR property implies that compliance with C-TPAT *systematically* reduces the distribution of risk across a given company's containers. Assuming that there isn't enough inspection capacity to eliminate terrorist threat completely, the nature of the optimal solution is formalized as follows:

Corollary 2 *Suppose $\Lambda \leq \frac{2\Delta}{w\sigma^2}$ and that $g_n(x)$ and $g_m(x)$ obey a strict MLR ordering. If $O_P^* > c_a$, at the optimal solution to the principal's problem, the results in Proposition 5 hold, and:*

- i. *If $\zeta > 0$, then there is at most a single value, x_u^* , such that $0 \leq x_l^* < x_u^* \leq 1$, and $\forall x \in (x_u^*, 1]$, $g_n(x)/g_m(x) > \zeta$.*
- ii. *$O_P^* = x_l^* L_e + (1 - x_l^*) L_f$.*
- iii. *$\varepsilon^* \leq x_l^*$.*

The optimal form of $p(x)$ is depicted in the right panel of Figure 3.

Since $O_P^* = k = x_l^* L_e + (1 - x_l^*) L_f$, the principal's objective function is strictly increasing in x_l^* . If there are multiple solutions that satisfy the conditions in Corollary 2, then we can restrict attention to the candidate solutions with the smallest value of x_l^* . Indeed, if ε^* varies continuously in x_l^* , then, assuming $O_P^* > c_a$, the optimal solution is unique, with $\varepsilon^* = x_l^*$, $p(1) = 1$ and $(x = 1) \in \mathcal{A}^*$.¹⁴

5 Comparative Statics with Capacity

Installed inspection capacity is a crucial determinant of overall security in the containerized supply chain. It can be thought of in terms of the number of customs inspectors available for container inspections at ports, along with the technology infrastructure in place, such as x-ray and gamma-ray scanners. Both more

¹³Strict MLR ordering implies strict FOSD, but not the other way around.

¹⁴For details, refer to Appendix C.

inspectors and better technology can allow for quicker and more precise inspections and thereby enable lower inspection times for a given ε . While greater capacity can provide for greater security, it is expensive, and a key decision CBP must make is how much to invest.

In this section we characterize the impact of changes in capacity, κ , on the equilibrium outcome. Using a mix of analytical and numerical approaches, we analyze the sensitivity of our optimal solution to the installed inspection capacity. We use the results of Proposition 5 and Corollary 2 as our starting point. Our first analytic result states:

Proposition 6 *Suppose $\Lambda \leq \frac{2\Delta}{w\sigma^2}$ and that $g_n(x)$ and $g_m(x)$ obey a strict MLR ordering. If $O_P^* > c_a$, then greater inspection capacity results in lower cost for CBP, and a lower x_l^* . Furthermore, if for any x_l and x_u that are candidates for the optimal solution in Corollary 2, the corresponding ε varies continuously with x_l , then greater capacity results in:*

- i. Higher θ_n and θ_m ;
- ii. Lower C-TPAT membership, $F(\alpha_t)$;
- iii. Higher effective arrival rate of containers to inspection facility, λ .

Appendix C characterizes sufficient conditions under which the optimal ε is continuous in x_l . In this case, $\varepsilon^* = x_l^*$ and $x_u^* = 1$.

Thus, for this special case, we find a somewhat surprising outcome: an increase in inspection capacity results in lower C-TPAT membership. The intuition for this result is that greater capacity provides CBP with the ability to inspect a higher volume of containers at the secondary inspection facility, thereby reducing the need for upstream security measures, as encompassed in C-TPAT. We have yet to study the impact of increasing capacity in the more general case, when the assumption in Proposition 6 may not hold. We do so through numerical experiments.

Numerical Study

We assign the following values to the model primitives. $G_n(x)$ is Beta(2,4) and $G_m(x)$ is Beta(2,7), giving

rise to strict MLR ordering, and mean risk scores of 0.33 and 0.22 respectively. $F(\alpha)$ is Gamma(9,1/63,000) resulting in a mean compliance cost of 1/7,000. All 3 distributions are bell shaped.

A heuristic rationale for the choice of the mean of the Gamma distribution is as follows. The mean annual cost of C-TPAT compliance for an importer is around \$70,000 (Diop et al. 2007), and we estimate the mean annual container traffic for an importer to be about 7,000. (Lee and Whang (2005) use the example of a high-tech manufacturer with a tradelane of 4,300 containers a year.) Then assuming negligible contribution from audit costs ($q = 0.5$ and typical $c_i(q) \approx \$1$), we estimate $\gamma_i \approx 70,000/7,000 = \10 . If the average container value, r_i , is about \$70,000, then a typical $\alpha \approx 10/70,000 = 1/7,000$.

The waiting cost per dollar per hour, w , can be estimated using a 20% cost of capital as: $1.2^{1/(365*24)} - 1 = 2.08 \times 10^{-5}$. To estimate the upper bound on per container cost, Δ , we assume that the maximum acceptable mean system wait time is 6 hours, the maximum $\theta_n = 1$, which leaves us with $\Delta = 6 \times w = 1.25 \times 10^{-4}$. We let the raw arrival rate of containers, Λ , be 100 per hour, which translates to 876,000 containers per annum.

Among the other parameters that we estimate: standard deviation of inspection time, $\sigma = 0.03$ hours; losses from a WMD that escapes detection, $L_e = \$520$ billion; losses from finding a WMD during inspection, $L_f = \$36.7$ billion. We use Abt (2003) as the basis for the loss estimates. L_e is the worst case projection that includes value of lives lost, while L_f is the indirect loss component (2/3 of total) of the best-case estimate presented therein. It is worth noting that the values of L_e and L_f do not influence the computation of the optimal policy for fixed capacity.

We use the functional form for SAT described in 2. In the numerical study that follows, we conduct a sensitivity analysis with respect to capacity, κ , and find that the results of Proposition 6 hold. In Figure 5, shown in Appendix A, we plot O_P^* versus capacity and find that it is strictly convex.

This gives us some insight into how the optimal capacity can be determined. Given a linear cost of capacity, h , CBP would choose capacity to optimize the following objective: $\min_{\kappa} [bO_P^* + h\kappa]$. If O_P^* is strictly convex in κ , then the overall objective is strictly convex as well, and the first-order-condition will specify the optimal capacity κ .

Example Number	κ	θ_n	θ_m	α_t	$F(\alpha_t)$	E(D) (hours)	λ (per hour)	ψ (hours)	ε^*	x_l^*	x_u^*	O_P^* (\$ billions)
1	1	0.027	0.002	1.17×10^{-4}	0.320	220.11	1.92	0.521	0.594	0.594	1	324.0
2	2	0.044	0.004	1.13×10^{-4}	0.285	136.13	3.26	0.306	0.542	0.542	1	299.0
3	5	0.081	0.012	1.06×10^{-4}	0.229	73.73	6.54	0.153	0.466	0.466	1	262.0
4	10	0.127	0.027	9.83×10^{-5}	0.173	47.08	10.99	0.091	0.403	0.403	1	231.0
5	20	0.197	0.058	8.79×10^{-5}	0.109	30.38	18.20	0.055	0.334	0.334	1	198.0
6	40	0.303	0.124	7.38×10^{-5}	0.048	19.75	29.47	0.034	0.258	0.258	1	161.0
7	60	0.388	0.191	6.34×10^{-5}	0.021	15.42	38.43	0.026	0.211	0.211	1	138.0
8	80	0.461	0.258	5.50×10^{-5}	0.009	12.99	45.93	0.022	0.176	0.176	1	122.0
9	100	0.525	0.323	4.79×10^{-5}	0.004	11.42	52.37	0.019	0.149	0.149	1	109.0
10	500	0.975	0.957	2.30×10^{-6}	0.000	6.14	97.55	0.010	0.006	0.006	1	39.7

Table 1: Sensitivity Analysis with Respect to Inspection Capacity

6 Discussion and Future Research

We have used a stylized model of port-security operations to obtain insights into the strategic considerations of CBP, trading firms that participate in C-TPAT, and terrorists. Our analysis points out that, for any given level of capacity, the program results in an improvement in the costs incurred by CBP and trading firms. Therefore, we can conclude that, even though security mandates might seem to be the easiest way to bolster homeland security, a creative use of economic mechanisms – ones that provide the right incentives for private sector (and individual) participation in security initiatives – can yield important benefits.

At the same time, it is important to remember that C-TPAT’s effectiveness is critically dependent on the improvement in the risk profile induced by the supply-chain practices included in the program, as well as the efficacy of ATS. These aspects are treated as exogenous to our model. Prospective changes on both of these fronts may lead to new operational challenges and to new opportunities for analysis.

From the trading firms’ point of view, the benefits of joining C-TPAT must offset the additional investment required to comply with the security guidelines. In this paper we focus our attention on the benefit related to reduced inspection frequency. An additional level of benefits pertains to a proposed tiered membership of C-TPAT. The highest performing members of C-TPAT would be eligible to have access to an inspection-free shipping process. This concept of expedited processing has been referred to as the “green lane” concept (C-TPAT Strategic Plan 2004). However, implementation of this scheme is contingent on R&D advances and successful roll-out of “smart” containers. Challenges remain, and it is yet to be ascertained whether green lanes will ever become a reality (Downey 2006). Also on the horizon is the benefit

associated with “restart priority” in the event of port closure due to a disaster. An economic analysis of both of these benefits present further opportunities for future work.

Our analysis generates useful high-level insights by characterizing the nature of the equilibrium outcome. However, the contrasting of our findings with reality presents a challenge in terms of accurately estimating the model parameters: the distributions G_n , G_m and F ; the cost of capacity h ; etc. Nevertheless, the numerical study in §5 illustrates how our model might be used to determine an optimal inspection policy. The current inspection frequency for non-members is about 5-6% (Marine Link 2004, McClure 2007), and our numerical results highlight the possibility that the optimal θ_n could be much larger, although not necessarily close to 100%. While the current membership level of C-TPAT is about 30% in terms of container traffic (GAO 2008), it is hard to draw a meaningful comparison with our model results, owing to the estimation problems described above.

Since the audit-policy parameters are determined independently of the optimal inspection policy, these are less affected by difficulty in estimating the true value of the model primitives. Here, we find that a 50% annual audit rate is optimal given $\beta = 1$, which assumes that the only benefit obtained from the program is via reduced inspection. In contrast, CBP plans to revalidate (or audit) the security profile of member firms only once every 3 years (GAO 2008). While it is possible that our ignoring of other benefits of C-TPAT, besides reduced inspection frequency, leads us to find a higher optimal audit rate in our analysis, the GAO has also raised concerns pertaining to the inadequacy of CBP’s revalidation strategy.

It is also worth noting that the idea of reduced inspections of trusted entities crossing US borders is applicable to other domains besides port and cargo security. CBP has trusted traveler programs (e.g., SENTRI, NEXUS) for frequent, low-risk border crossers. The program entitles trusted travelers to expedited inspection at the ports of entry (SENTRI 2006). In an analogue to the compliance cost trading firms incur when joining C-TPAT, these trusted travelers incur a dis-utility from subjecting themselves to an extensive background check, a pre-requisite for enrollment in the program. Similar ideas may be applicable to international mail as well. Although the scope of CBP’s mandate for inspections covers international mail (Wasem et al. 2004), it has not yet become a priority issue.

References

- Abt, C. C. 2003. The Economic Impact of Nuclear Terrorist Attacks on Freight Transport Systems in an Age of Seaport Vulnerability. Report by Abt Associates Inc.
- Abt, C. C., W. Rhodes, R. Casagrande and G. Gaumer. 2003. The Economic Impacts of Bioterrorist Attacks on Freight Transport Systems in an Age of Seaport Vulnerability. Report by Abt Associates Inc.
- Afeche, P. 2006. Incentive-compatible revenue management in queueing systems: optimal strategic delay and other delay tactics. Working paper, Rotman School, University of Toronto.
- Altman, E., T. Boulogne, R. El-Alzouzi, T. Jimenez and L. Wynter. 2006. A Survey of Networking Games in Telecommunications. *Computers and Operations Research*, 33:2, 286-311.
- Bakshi, N. and N. Gans. 2008. Measuring the Operational Impact of Container Inspections at International Ports. Working paper, The Wharton School, University of Pennsylvania.
- Basham, W. R. 2007. Remarks by CBP Commissioner W. Ralph Basham on Container Security at the Center for Strategic and International Studies. www.cbp.gov/xp/cgov/newsroom/commissioner/speeches_statements/commish_remarks_csc.xml.
- Bettge, J. 2006. Private Communication.
- Bonner, R. C. 2003. Testimony of Commissioner Robert C. Bonner in U.S. Congress. http://commerce.senate.gov/hearings/testimony.cfm?id=912&wit_id=2545.
- Boske, L. B. 2006. Port and Supply-Chain Security Initiatives in the United States and Abroad. Report by Lyndon B. Johnson School of Public Affairs, UT Austin.
- C-TPAT Strategic Plan. Customs-Trade Partnership Against Terrorism: Strategic Plan. www.cbp.gov/xp/cgov/import/commercial_enforcement/ctpat/.
- Diop, A. D., D. Hartman and D. Rexrode. 2007. C-TPAT: Cost/Benefit Survey. Report by Center for Survey Research, University of Virginia.
- Downey, L. 2006. International Cargo Conundrum: How much investment in security is enough? *RFID Journal*, Feb. 6, 2006.
- FAST. 2007. Free and Secure Trade Program. www.cbp.gov/xp/cgov/import/commercial_enforcement/ctpat/fast/
- Fawcett, T. 2006. An Introduction to ROC Analysis. *Pattern Recognition Letters*, 27, 861-874.
- GAO. 2004. Summary of Challenges Faced in Targeting Oceangoing Cargo Containers for Inspection. Report GAO-04-557T.
- GAO. 2005. Homeland Security: Key Cargo Security Programs Can Be Improved. Report GAO-05-466T.
- GAO. 2006-a. Border Security: Investigators Transported Radioactive Sources Across Our Nation's Borders at Two Locations. Report GAO-06-583T.
- GAO. 2006-b. Cargo Container Inspections: Preliminary Observations on the Status to Efforts to Improve the Automated Targeting System. Report GAO-06-591T.

- GAO. 2007-a. Maritime Security: Observations on Selected Aspects of the SAFE Port Act. Report GAO-07-754T.
- GAO. 2007-b. Homeland Security: Applying Risk Management Principles to Guide Federal Investments. Report GAO-07-754T.
- GAO. 2008. Supply Chain Security: US Customs and Border Protection Has Enhanced Its Partnership with Import Trade Sectors, but Challenges Remain in Verifying Security Practices. Report GAO-08-240.
- Gerencser, M., J. Weinberg and D. Vincent. 2003. Port Security War Game: Implications for U.S. Supply Chains. Report by Booz Allen Hamilton.
- Ghylin, K. M., C. G. Drury and A. Schwaninger. 2006. Two-component model of security inspection: application and findings. 16th World Congress of Ergonomics, IEA, Maastricht, The Netherlands.
- Graham, B., J. Talent, G. Allison, R. Cleveland, S. Rademaker, T. Roemer, W. Sherman, H. Sokolski and R. Verma. 2008. World at Risk. Report of the Commission on the Prevention of WMD Proliferation and Terrorism.
- Greenberg, M. D., P. Chalk, H. H. Willis, I. Khilko and D. S. Ortiz. 2006. Maritime Terrorism, Risk and Liability. Report by RAND Center for Terrorism Risk Management Policy.
- Hopp, W. J., G. Yuen, and S. M. R. Iravani. 2007. Operation Systems with Discretionary Task Completion. *Management Science*, 53:1, 61-77.
- Huizenga, D. 2005. Detecting nuclear weapons and radiological materials: How effective is available technology? Testimony before the House Committee on Homeland Security.
- Keane, A. G. Where's the Incentive? 2005. *Traffic World*, p12, April 11.
- Kleindorfer, P. R. and G. H. Saad. 2005. Managing Disruption Risks in Supply Chains. *Production and Operations Management*, 14:1, p53-98.
- Kollerstrom, J. 1974. Heavy Traffic Theory for Queues with Several Servers. I. *Journal of Applied Probability*, 11:3, p544-552.
- Kunreuther, H. C., P. J. McNulty, and Y. Kang. 2002. Third-Party Inspection as an Alternative Command and Control Regulation. *Risk Analysis*, 22:2, p309-318.
- Kurtz, H. 2003. ABC ships Uranium overseas for story. *Washington Post*, p.A21, September 11.
- Laffont, J. J., and D. Martimort. 2001. The Theory of Incentives, The Principal-Agent Model. Princeton University Press.
- Lee, H. L., and S. Whang. 2005. Higher Supply Chain Security with Lower Cost: Lessons from Total Quality Management. *International Journal of Production Economics*, 96:3, 289-300.
- Lugar, R. G. 2005. The Lugar Survey on Proliferation Threats and Responses. Report by the Chairman, Senate Foreign Relations Committee.
- Marine Link. 2004. The 5% Myth. <http://www.marinelink.com/Story/The+5%25+Myth-15423.html>. Accessed online on 08/29/08.
- Marshall, A. W., and I. Olkin. 1968. A General Approach to Some Screening and Classification Problems. *Journal of the Royal Statistical Society. Series B (Methodological)*, 30:3, 407-443.

- Martonosi, S. E., and A. Barnett. 2006. How Effective Is Security Screening of Airline Passengers? *Interfaces*, 36:6, 545-552.
- Martonosi, S. E., D. S. Ortiz and H. H. Willis. 2006. Evaluating the Viability of 100 Percent Container Inspection at America's Ports. Report by RAND Corporation.
- McClelland, J. L. 1979. On the Time Relations of Mental Processes: An Examination of Systems of Processes in Cascade. *Psychological Review*, 86:4, 287-330.
- McClure, G. 2007. How Safe Are Our Ports? <http://www.todaysengineer.org/2007/Sep/port-security.asp>. Accessed online on 08/14/08.
- SENTRI. 2006. Secure Electronic Network for Travelers Rapid Inspection. U.S. CBP www.cbp.gov/xp/cgov/travel/frequent_traveler/sentri/sentri.xml.
- Shaked, M. and G. J. Shanthikumar. 1994. Stochastic Orders and their Applications. San Diego: Academic Press.
- Sheffi, Yossi. 2005. The Resilient Enterprise, MIT press.
- Shumsky, R. A. and E. J. Pinker. 2003. Gatekeepers and Referrals in Services. *Management Science*, 49:7, 839-856.
- Spencer, C. 2004. International supply chain security regulatory programs presented at McCombs School of Business, University of Texas at Austin, 16 October 2003, IMS Worldwide. <http://mba.mcombs.utexas.edu/plus/Academies/BAB/CurtisSpencer.ppt>, 17 August.
- Stanford Study Group. 2002. Detecting Nuclear Material in International Container Shipping: Criteria for Secure Systems. Report of Center for International Security and Cooperation. Stanford University.
- Steenken, D., S. Voß, and R. Stahlbock. 2004. Container terminal operation and operations research - a classification and literature review. *OR Spectrum*, 26:1, 3-49.
- Tomlin, Brian. 2006. On the Value of Mitigation and Contingency Strategies for Managing Supply Chain Disruption Risks. *Management Science*, 52:5, p639-657.
- USA TODAY. 2007. Most fake bombs missed by screeners. www.usatoday.com/news/nation/2007-10-17-airport-security_N.htm, October 17.
- Wasem, R. E., J. Lake, L. Seghetti, J. Monke, and S. Vina. 2004. Border Security: Inspection Practices, Policies, and Issues. Congressional Research Service Report.
- Wein, L. M., A. H. Wilkins, M. Baveja, and S. E. Flynn. 2006. Preventing the Importation of Illicit Nuclear Materials in Shipping Containers. *Risk Analysis*, 26:5.
- Wein, L. M., Y. Liu, Z. Cao, and S. E. Flynn. 2007. The Optimal Spatiotemporal Deployment of Radiation Portal Monitors Can Improve Nuclear Detection at Overseas Ports. To appear in *Science & Global Security*.
- Whitt, Ward. Uniform Conditional Stochastic Order. *Journal of Applied Probability*, 17:1, 112-123, 1980.
- Wickelgren, W. 1977. Speed Accuracy Tradeoff and Information Processing Dynamics. *Acta Psychologica*, 41, 67-85.
- Wolff, R. W. 1989. Stochastic Modeling and the Theory of Queues. Prentice Hall.

A Appendix: Figures

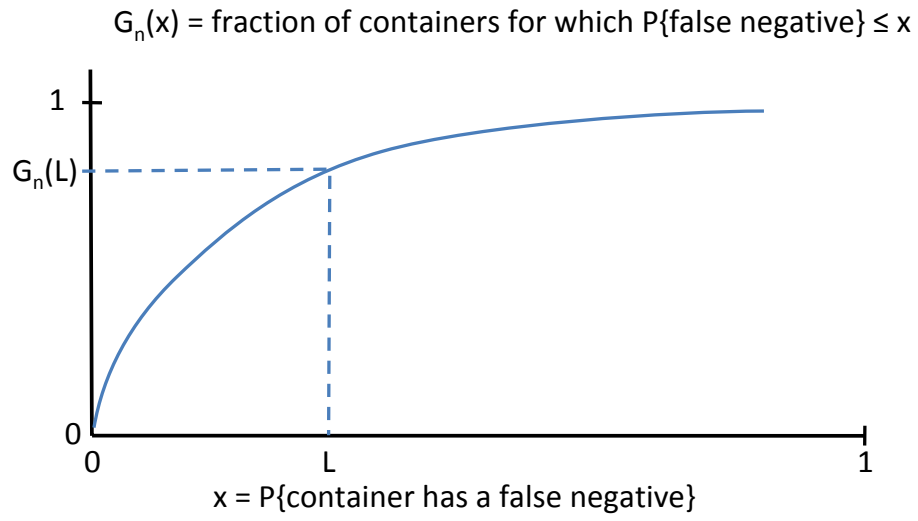


Figure 4: Sample CDF for risk scores.

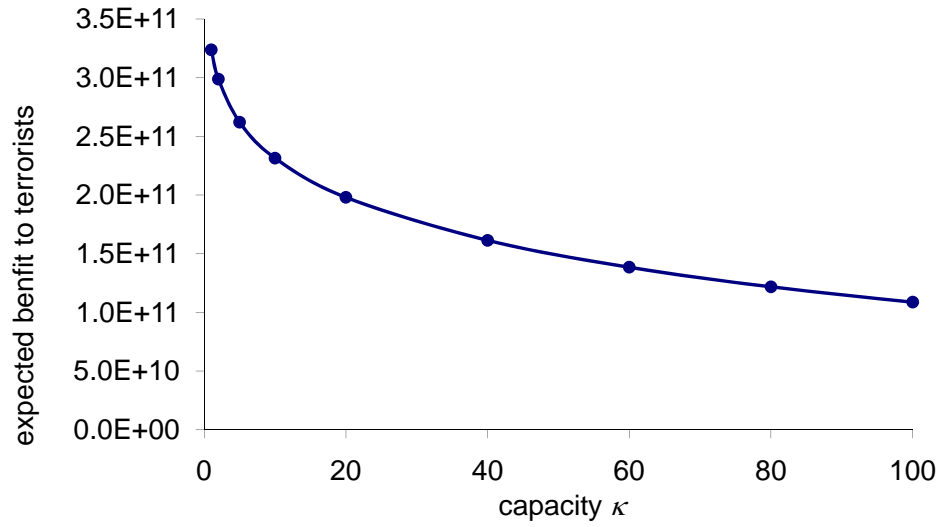


Figure 5: Plot of Expected Benefit to Terrorists (O_P^*) Vs. Inspection Capacity (κ)

B Appendix: Proofs

Lemma 1

Proof Part (i). First, we show that CBP never need consider inspection plans with $\varepsilon = 0$. Recall that $\lim_{\varepsilon \rightarrow 0} \psi(\varepsilon, \kappa) = \infty$. Therefore, for $\varepsilon = 0$ the only feasible solution is $p(x) = 0$ for $x \in [0, 1]$, so that $\theta_n = 0$. From (5) we see that, here, the expected benefit to terrorists is

$$x\{0[0L_e + (1 - 0)L_f] + (1 - 0)L_e\} + (1 - x)L_f = xL_e + (1 - x)L_f,$$

which is maximized at $x = 1$. In this case CBP can do as well by letting $\varepsilon = 1$ and $p(x) = 1$ for $x \in [0, 1]$ so that the expected benefit to terrorists is

$$x\{1[1L_e + (1 - 1)L_f] + (1 - 1)L_e\} + (1 - x)L_f = xL_e + (1 - x)L_f,$$

which again is maximized at $x = 1$. Thus, the latter plan performs as well as the former, and CBP need only consider it.

Second, recall that there exists $\varepsilon < 1$ for which $\psi(\varepsilon, \kappa) < \infty$. We show that for this ε , we can set $p(x) = p \in (0, 1)$ for all $x \in [0, 1]$, where p is feasible for the (IR_b) constraint and provides a lower expected benefit to terrorists than the schemes above. Specifically, for a given p , we have $\theta_n = \int_0^1 p g_n(x) dx = p$ and $\lambda = \Lambda \theta_n = \Lambda p$, so that (IR_b) becomes

$$p w \left[\frac{\Lambda p (\psi^2 + \sigma^2)}{2(1 - \Lambda p \psi)} \right] \leq \Delta.$$

Observe that the left-hand side of the inequality is strictly decreasing in p and that, for any $\psi < \infty$, there exists a $p > 0$ small enough to satisfy (IR_b) . Then using this p and ε in (5) we have

$$x\{p[\varepsilon L_e + (1 - \varepsilon)L_f] + (1 - p)L_e\} + (1 - x)L_f,$$

which is strictly increasing in x . At $x = 1$ we have

$$1\{p[\varepsilon L_e + (1 - \varepsilon)L_f] + (1 - p)L_e\} + (1 - 1)L_f = p[\varepsilon L_e + (1 - \varepsilon)L_f] + (1 - p)L_e\},$$

which is strictly less than L_e when $p > 0$. Thus any optimal solution will have $\varepsilon < 1$.

Finally, note that any solution with $\theta_n = 0$ will have $p(x) = 0$ for all $x \in [0, 1]$, and an

$$x\{0[\varepsilon L_e + (1 - \varepsilon)L_f] + (1 - 0)L_e\} + (1 - x)L_f = xL_e + (1 - x)L_f,$$

with maximum to terrorists of L_e at $x = 1$. Thus, this solution is outperformed by that with $\varepsilon < 1$ and $p > 0$, above. Therefore, no optimal solution can have $\theta_n = 0$.

Part (ii). By contradiction suppose $O_P^* > c_a$ and the (IR_b) were not binding:

$$\theta_n w \left[\frac{\Lambda \theta_n (\psi^2 + \sigma^2)}{2(1 - \Lambda \theta_n \psi)} \right] < \Delta.$$

We will construct another inspection scheme that outperforms O_P^* .

From part (i) recall that any optimal solution must have $\varepsilon < 1$ and $\theta_n > 0$. Recalling (5), we let

$$t(x) = x\{p(x)[\varepsilon L_e + (1 - \varepsilon)L_f] + (1 - p(x))L_e\} + (1 - x)L_f,$$

and let $\mathcal{A}^* = \arg \max\{t(x) \mid x \in [0, 1]\}$.

First, we note that, without loss of generality, we can assume that, if $x \notin \mathcal{A}^*$ then $p(x) = 0$. Note that $t(x)$ is decreasing in $p(x)$. Therefore, given some $x \notin \mathcal{A}^*$, we can lower $p(x)$ until either $x \in \mathcal{A}^*$ or $p(x) = 0$, whichever comes first.

Second, let $\mathcal{A}_0^* = \{x \in \mathcal{A}^* \mid p(x) = 0\}$, and let $\mathcal{A}_{>0}^* = \mathcal{A}^* \setminus \mathcal{A}_0^*$. Then we claim that $\mathcal{A}_{>0}^* \neq \emptyset$. If the set were empty, then we would have $p(x) = 0$ for all $x \in [0, 1]$ and $\theta_n = 0$, which contradicts the fact that $\theta_n > 0$, established in part (i).

Third, we observe that $(x = 0) \notin \mathcal{A}^*$. Note that $t(0) = L_f$, so if it were then $O_P^* = L_f < c_a$, which contradicts the assumption that $O_P^* > c_a$.

Fourth we observe that \mathcal{A}_0^* is either empty or a singleton. If not, then we would have $x_1, x_2 \in \mathcal{A}_0^*$ with $x_1 < x_2$ and $t(x_1) < t(x_2)$, which contradicts the definition of \mathcal{A}^* and \mathcal{A}_0^* .

If \mathcal{A}_0^* is empty, then $\mathcal{A}^* = \mathcal{A}_{>0}^*$ and all $x \in \mathcal{A}^*$ have $p(x) > 0$. If we decrease ε by a small amount, say δ , which is small enough that the (IR_b) constraint is still not binding, then we can reduce

$$t(x) = x\{p(x)[(\varepsilon - \delta)L_e + (1 - (\varepsilon - \delta))L_f] + (1 - p(x))L_e\} + (1 - x)L_f,$$

for all $x \in \mathcal{A}^*$. In this case, we have further reduced the maximum $t(x)$, which contradicts the assumption that O_P^* was optimal.

If \mathcal{A}_0^* is a singleton, then the above decrease in ε leaves the point $x \in \mathcal{A}_0^*$ as the single $x \in \mathcal{A}^*$. Because there remains slack in the (IR_b) we can increase $p(x) > 0$ so that, $x \notin \mathcal{A}^*$ and we obtain the desired result. □

Proposition 1

Proof In order to prove the results in this Proposition, we first establish the results listed in the following steps:

1. $p(x)$ is strictly increasing on the set \mathcal{A}_b^* .
2. If $0 < p(x) < 1$, then $x \in \mathcal{A}_b^*$.
3. $\forall x \in [0, 1)$, $p(x) < 1$.
4. $(x = 1) \in \mathcal{A}_b^*$.
5. $p(x_{lb}^*) = 0$ and $x_{lb}^* \in \mathcal{A}_b^*$.

It can be verified that the results in the above steps imply the claims in (i), (ii) and (iii). The definition of x_{lb}^* and step 5 imply the result in (ii). Thereafter, steps 2, 3 and 5 together imply that $p(x) =$

$0, \forall x \in [0, x_{lb}^*]; 0 < p(x) < 1, \forall x \in [x_{lb}^*, 1];$ and $[x_{lb}^*, 1) \subset \mathcal{A}_b^*$. These observations coupled with step 4 imply that $p(x) = \frac{1 - \frac{x_{lb}^*}{x}}{1 - \varepsilon_b^*}, \forall x \in [x_{lb}^*, 1]$. Thus, we know that the result in (i) is true. The feasibility condition $p(1) \leq 1$ then gives us the result in (iii).

We now proceed with the proof:

Step 1. By definition, the expected benefit offered by all elements of \mathcal{A}_b^* is equal to some constant, which we denote by k_b . Using (5) we note that for $x_1, x_2 \in \mathcal{A}_b^*$ such that $x_1 < x_2$, we have:

$$\begin{aligned} x_1 \{p(x_1)[\varepsilon_b L_e + (1 - \varepsilon_b)L_f] + (1 - p(x_1))L_e\} + (1 - x_1)L_f &= k_b \\ &= x_2 \{p(x_2)[\varepsilon_b L_e + (1 - \varepsilon_b)L_f] + (1 - p(x_2))L_e\} + (1 - x_2)L_f. \end{aligned}$$

Then $\varepsilon < 1, x_1 < x_2$ and $L_f < L_e$ imply

$$p(x_1)[\varepsilon_b L_e + (1 - \varepsilon_b)L_f] + (1 - p(x_1))L_e > p(x_2)[\varepsilon_b L_e + (1 - \varepsilon_b)L_f] + (1 - p(x_2))L_e,$$

which finally gives us the result $p(x_1) < p(x_2)$. Thus, we conclude that $p(x)$ is strictly increasing in x on the set \mathcal{A}_b^* .

Step 2. We can prove this result by contradiction. Assume that the result is not true. Then for the optimal solution, consider $x^* \in \mathcal{A}_b^*$ and $\hat{x} \notin \mathcal{A}_b^*$ such that $0 < p(\hat{x}) < 1$. Then we have:

$$\begin{aligned} x^* \{p(x^*)[\varepsilon_b L_e + (1 - \varepsilon_b)L_f] + (1 - p(x^*))L_e\} + (1 - x^*)L_f \\ > \hat{x} \{p(\hat{x})[\varepsilon_b L_e + (1 - \varepsilon_b)L_f] + (1 - p(\hat{x}))L_e\} + (1 - \hat{x})L_f. \end{aligned}$$

Using the result in Lemma 1 and (11) we know that any candidate optimal solution must satisfy the following two conditions:

$$\theta_n w(E(D) + E(S)) = \Delta, \tag{17}$$

$$0 \leq p(x) \leq 1, \forall x. \quad (18)$$

Denoting a small perturbation in the value of $p(\hat{x})$ by $dp(\hat{x})$, we obtain:

$$\frac{d\theta_n}{dp(\hat{x})} = g_n(\hat{x})d\hat{x}. \quad (19)$$

We rewrite (17) as follows:

$$\phi(\theta_n, \varepsilon_b) = \Delta - \theta_n w(E(D) + E(S)) = 0. \quad (20)$$

Using the Implicit Function Theorem on the above equation we obtain:

$$\frac{d\phi}{dp(\hat{x})} = 0 = \frac{\partial \phi}{\partial \theta_n} \frac{d\theta_n}{dp(\hat{x})} + \frac{\partial \phi}{\partial \varepsilon_b} \frac{d\varepsilon_b}{dp(\hat{x})},$$

or equivalently:

$$\frac{d\varepsilon_b}{dp(\hat{x})} = -\frac{\frac{\partial \phi}{\partial \theta_n} \frac{d\theta_n}{dp(\hat{x})}}{\frac{\partial \phi}{\partial \varepsilon_b}}. \quad (21)$$

To evaluate the expression in (21), the following relationships are required:

$$\frac{\partial \phi}{\partial \theta_n} = -w(E(D) + \psi) - \theta_n w \frac{\partial E(D)}{\partial \lambda} \frac{\partial \lambda}{\partial \theta_n} = -w(E(D) + \psi) - \theta_n w \frac{\psi^2 + \sigma^2}{2(1 - \lambda\psi)^2} \Lambda < 0, \quad (22)$$

$$\frac{\partial \phi}{\partial \varepsilon_b} = -\theta_n w \left[\frac{\partial E(D)}{\partial \varepsilon_b} + \psi_{\varepsilon_b} \right] = -\theta_n w \left[\frac{(1 - \lambda\psi)2\lambda\psi + \lambda^2(\psi^2 + \sigma^2)}{2(1 - \lambda\psi)^2} + 1 \right] \psi_{\varepsilon_b} > 0. \quad (23)$$

Using (19), along with the above two relationships in (21) along with the result of Lemma 1, we obtain

$\forall x \in [0, 1]$:

$$\frac{d\varepsilon_b}{dp(x)} = \frac{\left(-\frac{\partial \phi}{\partial \theta_n} \right) g_n(x) dx}{\frac{\partial \phi}{\partial \varepsilon_b}} > 0 \quad (24)$$

Now choose the magnitude of $dp(\hat{x})$ such that x^* would continue to offer the highest benefit to terrorists if ε_b were left unaltered. Also, in (21), choose the sign of $dp(\hat{x})$ such that $d\varepsilon_b < 0$, i.e., decrease

$p(\hat{x})$. It is always possible to do so because $0 < p(\hat{x}) < 1$.

After choosing the appropriate $dp(\hat{x})$ we find that $d\varepsilon_b < 0$, i.e., the new value of residual risk is $\varepsilon'_b = \varepsilon_b + d\varepsilon_b$ such that $\varepsilon'_b < \varepsilon_b$.¹⁵ Two cases can arise:

Case 2.1: x^* continues to offer the greatest benefit to terrorists. Since $\varepsilon'_b < \varepsilon_b$, we have managed to improve the principal's objective function value, while ensuring feasibility. Hence, the previous solution could not have been optimal.

Case 2.2: x' now offers the greatest benefit to terrorists, where $x' \neq x^*$. We have:

$$\begin{aligned} & x^* \{p(x^*)[\varepsilon_b L_e + (1 - \varepsilon_b)L_f] + (1 - p(x^*))L_e\} + (1 - x^*)L_f \\ & \geq x' \{p(x')[\varepsilon_b L_e + (1 - \varepsilon_b)L_f] + (1 - p(x'))L_e\} + (1 - x')L_f \\ & > x' \{p(x')[\varepsilon'_b L_e + (1 - \varepsilon'_b)L_f] + (1 - p(x'))L_e\} + (1 - x')L_f. \end{aligned}$$

Once again, since we have improved the objective function value, the previous solution could not have been optimal. Thus, we have shown that at the optimal solution, any container with risk score x , such that $0 < p(x) < 1$, offers the same (maximal) expected benefit to terrorists, denoted by k_b .

Step 3. In order to prove the claim in this step we observe that for $x \notin \mathcal{A}^*$ it is possible for $p(x)$ to take on value 1, only if $d\varepsilon_b/dp(x) \leq 0$. Otherwise a reduction in $p(x)$ would lead to a lowering of ε_b , thereby improving the objective function value. (Note that the reduction in $p(x)$ should be small enough such that x continues to be outside of \mathcal{A}^* .) Using (24) we can now conclude that $\forall x \in [0, 1)$, if $x \notin \mathcal{A}_b^*$, then $p(x) < 1$.

On the other hand, if $x \in \mathcal{A}_b^*$, then the only way to have $p(x) = 1$ is if $x = \sup \mathcal{A}_b^*$, otherwise the monotonicity of $p(x)$ on \mathcal{A}_b^* (result in Step 1) would be violated. Moreover, $\sup \mathcal{A}_b^* = 1$, since for points with $p(x) = 1$, the expected benefit to terrorists has to be strictly increasing in x . Hence, from

¹⁵To get around the technical difficulty posed by $\frac{d\varepsilon_b}{dp(\hat{x})} \rightarrow 0$, we can consider the set $I \triangleq \{\hat{x} | 0 < p(\hat{x}) < 1, \hat{x} \notin \mathcal{A}_b^*\}$. Then choosing the appropriate $dp(\hat{x})$ for every element of I , we define $\Delta\varepsilon_b = \int_{\hat{x} \in I} \frac{d\varepsilon_b}{dp(\hat{x})} d\hat{x} < 0$. Then $\varepsilon'_b = \varepsilon_b + \Delta\varepsilon_b$. For simplicity in exposition we will not repeat this technical argument at later points in the paper when such issues arise around infinitesimal changes.

amongst these risk scores, only the highest one could belong to \mathcal{A}_b^* . Thus, $\forall x \in [0, 1), p(x) < 1$.

Step 4. We rule out the possibility that $p(1) = 0$, otherwise $(x = 1)$ will offer the theoretically maximum possible benefit - L_e , and therefore $\mathcal{A}_b^* = \{1\}$. Using the result from Step 2 this would imply that $\theta_n = 0$. We know from Lemma 1 that this would be suboptimal. At optimality, we can also rule out the possibility that $p(1) = 1$ and $(x = 1) \notin \mathcal{A}_b^*$. This is because we showed in (24) that $\frac{d\varepsilon_b}{dp(x)} > 0$ at $x = 1$, hence a reduction in the value of $p(1)$ would improve the value of the objective function, thus violating optimality. Finally, we are left with the possibility that $0 < p(1) < 1$. Then by the result in Step 2, we can conclude that $(x = 1) \in \mathcal{A}_b^*$.

Step 5. Consider $x \in [0, x_{lb}^*)$. Using the results from Steps 2 and 3, we know that $p(x) = 0$. Then:

$$\lim_{x \rightarrow x_{lb}^*} [xL_e + (1-x)L_f] \leq x_{lb}^* \{p(x_{lb}^*)[\varepsilon_b L_e + (1-\varepsilon_b)L_f] + (1-p(x_{lb}^*))L_e\} + (1-x_{lb}^*)L_f. \quad (25)$$

Since $p(x_{lb}^*) \geq 0$, therefore L_e is necessarily greater than or equal to $p(x_{lb}^*)[\varepsilon_b L_e + (1-\varepsilon_b)L_f] + (1-p(x_{lb}^*))L_e$. The relationship in (25) can hold only if $p(x_{lb}^*) \rightarrow 0$. Hence, $\forall x \in [0, x_{lb}^*]$, we have $p(x) = 0$. Now if $x_{lb}^* \notin \mathcal{A}_b^*$, then:

$$\lim_{x \rightarrow x_{lb}^*} [x\{p(x)[\varepsilon_b L_e + (1-\varepsilon_b)L_f] + (1-p(x))L_e\} + (1-x)L_f] > [x_{lb}^* L_e + (1-x_{lb}^*)L_f], \quad (26)$$

where $p(x) > 0$. This leads to a contradiction. Hence, $x_{lb}^* \in \mathcal{A}_b^*$.

□

Proposition 2

Proof At the threshold compliance cost α_t , the condition in (9) is binding. Using $\alpha = (\gamma_i + qc_i(q))/r_i$ and $w = d_i/r_i$ in the binding version of (9), we obtain

$$\alpha_t = (\theta_n - \theta_m)w(E(D) + E(S)). \quad (27)$$

The LHS of the threshold condition is monotonically increasing (strictly) in α_t . We claim that the right-hand side (RHS) is monotonically decreasing. Recalling (28), we see that this is because $\frac{\partial E(D)}{\partial \alpha_t} \leq 0$, which ensures that the equation has a unique solution, if it exists.

To show existence, we observe that when $\alpha_t = 0$, the RHS is some positive number, $E(D) + \psi$, where $E(D)$ is evaluated with $\lambda = \Lambda\theta_n$. Based on the described monotonicity in α_t , and since $\alpha_t \in [0, \infty)$, a solution to the equation must exist. Note that, in the case that compliance-costs have a finite upper bound, B , i.e., $\alpha \in [0, B]$, if it turns out that the solution to (11) results in $\alpha_t > B$, then using the obtained value of α_t for all subsequent calculations would still provide $F(\alpha_t) = F(B) = 1$. \square

Proposition 3

Proof To see this, first note that an increase in P_i allows the inequality (13) to be satisfied for a smaller q , and, in turn, a lower cost $c_i(q)$. Since, in equilibrium, the agent incurs the expected audit cost $qc_i(q)$, but not the penalty P_i , any given agent would want q to be as small as possible. Second, observe that from Proposition 2 we also know that the threshold compliance cost, α_t , is independent of the choice of q . Recalling that each trading firm has $\alpha(q) \equiv (\gamma_i + qc_i(q))/r_i$, we see that a reduction in q therefore allows for greater γ_i and increases participation in C-TPAT. Further, because

$$\frac{\partial E(D)}{\partial \alpha_t} = \frac{\partial \lambda}{\partial \alpha_t} \left[\frac{\psi^2 + \sigma^2}{2(1 - \lambda\psi)^2} \right] = -\Lambda f(\alpha_t)(\theta_n - \theta_m) \left[\frac{\psi^2 + \sigma^2}{2(1 - \lambda\psi)^2} \right] \leq 0 \quad (28)$$

and

$$\frac{\partial E(D)}{\partial \varepsilon} = \frac{(1 - \lambda\psi)2\lambda\psi + \lambda^2(\psi^2 + \sigma^2)}{2(1 - \lambda\psi)^2} \psi_\varepsilon < 0, \quad (29)$$

we can see that, while maintaining $p(x)$ and system congestion at a constant level (which ensures feasibility), an increase in participation in C-TPAT will have to result in a smaller arrival rate and therefore a lower value of ε . From (12) we see that lower ε will lead, in turn, to a lower value of O_P . Therefore, CBP should minimize q to lower the compliance cost of trading firms, maximize participation, and thereby minimize its expected costs. Thus, $P_i^* = B_i$. \square

Proposition 4

Proof The result follows from ensuring (IC) is satisfied for $\alpha = \alpha_t$, and the assumption that $c'_i(q) \geq 0$. We claim that, for any given level of system congestion, it is optimal for CBP to maximize participation in C-TPAT. Recall from (28) that $\frac{\partial E(D)}{\partial \alpha_t} \leq 0$. This implies that for a given ε and $p(x)$, an increase in C-TPAT membership will result in a decrease in delay. Since we plan to keep system congestion, or $E(D) + E(S)$, fixed and, from (29), we have $\frac{\partial E(D)}{\partial \varepsilon} < 0$, we can therefore conclude that, if we leave $p(x)$ fixed, then there will be a reduction in ε in response to an increase in α_t . This leads to an improvement in CBP's objective function, while maintaining feasibility as required by the (IR) constraint.

Next we claim that for any fixed value of $E(D) + E(S)$, CBP can maximize membership by making q as small as possible. In particular, recall from (11) that for fixed ε , we have $\alpha_t = (\theta_n - \theta_m)w(E(D) + E(S))$, where $\alpha_t = (\gamma_t + qc_t(q))/r_t$ for some $i = t$, although this index i may not be unique. Further, for all $\alpha \leq \alpha_t$ we obtain $\alpha = (\gamma_i + qc_i(q))/r_i \leq (\gamma_t + qc_t(q))/r_t$.

Now recall that $c'_i(q) \geq 0$. Then if we reduce q we will also reduce α , for all $\alpha \leq \alpha_t$. Hence, these companies continue to be members. Similarly, for $\alpha > \alpha_t$, a decrease in q will reduce $\alpha = (\gamma_i + qc_i(q))/r_i$, possibly permitting more firms' α 's to fall below the original threshold α_t .

Thus, membership is weakly decreasing in q , so the value of q will be driven down until the (IC) constraint is tight. At this point $\alpha_t \leq q(1 + \beta)\alpha_t$, or $q^* = 1/(1 + \beta)$, which is independent of ε . \square

Lemma 2

Proof Part (i). The argument used for part (i) of Lemma 1 holds here as well and is omitted.

Part (ii). First, note that any given choice of $p(x)$'s determines θ_n and θ_m . Second, observe that, for a fixed choice of θ_m and θ_n , a large enough decrease in ε will cause the LHS of (IR) to become binding on the right hand side, Δ . This occurs no matter how α or $E[D]$ changes with ε , because $\psi_\varepsilon < 0$ and $\lim_{\varepsilon \rightarrow 0} \psi = \infty$. Third, note that, for fixed $p(x)$'s with $\theta_n, \theta_m > 0$, the objective function O_P is strictly increasing in ε . Therefore, it is optimal for CBP to set ε to be as low as possible, driving the expected cost incurred by a trading firm up to its upper bound, Δ . Since capacity is limited, we have ruled out the

possibility that terrorist threat is completely eliminated by choosing a high enough $p(x)$ and a low enough ε , without violating the (IR) constraint. \square

Proposition 5

Proof

- i. By definition, the expected benefit offered by all elements of \mathcal{A}^* is equal to some constant, which we denote by k . On inspecting (16) we note that for $x_1, x_2 \in \mathcal{A}^*$ such that $x_1 < x_2$, we have:

$$\begin{aligned} x_1 \{p(x_1)[\varepsilon L_e + (1 - \varepsilon)L_f] + (1 - p(x_1))L_e\} + (1 - x_1)L_f &= k \\ &= x_2 \{p(x_2)[\varepsilon L_e + (1 - \varepsilon)L_f] + (1 - p(x_2))L_e\} + (1 - x_2)L_f, \end{aligned}$$

which implies:

$$p(x_1)[\varepsilon L_e + (1 - \varepsilon)L_f] + (1 - p(x_1))L_e > p(x_2)[\varepsilon L_e + (1 - \varepsilon)L_f] + (1 - p(x_2))L_e,$$

which finally gives us the result $p(x_1) < p(x_2)$. Thus, we conclude that $p(x)$ is strictly increasing in x on the set \mathcal{A}^* .

- ii. We first introduce some new notation for expositional convenience:

$$\mathcal{A}^1 \triangleq \{x | p(x) = 1\}; \mathcal{A}^0 \triangleq \{x | p(x) = 0\}.$$

$$x_u^* \triangleq \sup \mathcal{A}^*; x_u^1 \triangleq \sup \mathcal{A}^1; x_u^0 \triangleq \sup \mathcal{A}^0.$$

The three sets - \mathcal{A}^* , \mathcal{A}^1 , \mathcal{A}^0 - are not necessarily disjoint but they are exhaustive, i.e., $\mathcal{A}^* \cup \mathcal{A}^1 \cup \mathcal{A}^0 = [0, 1]$. It is straightforward to see that, *for the sets \mathcal{A}^1 and \mathcal{A}^0 , the expected benefit offered to terrorists is monotonic in the risk score*. Hence, the only risk scores that could potentially belong to the intersection of these sets with \mathcal{A}^* are x_u^1 and x_u^0 .

In order to prove the results in (ii), we first establish to results listed in the following steps:

1. If $0 < p(x) < 1$, then $x \in \mathcal{A}^*$ and (16) is satisfied.
2. $\forall x \in [0, 1)$, $p(x) = 1 \iff \frac{d\varepsilon}{dp(x)} \leq 0$.
3. $\zeta \leq 0 \iff \frac{d\varepsilon}{dp(x)} > 0 \implies p(x) < 1, \forall x \in [0, 1)$, and $p(x_l^*) = 0$.

It can be verified that the results in the above steps imply the claims in (ii)-a), b) and c). Using the result in step 3 we can conclude that $p(x) < 1, \forall x \in [0, 1)$ and $p(x_l^*) = 0$. Then using the result from part (i) along with the definition of x_l^* and the results in steps 1 and 3, we find that $p(x) = 0, \forall x \in [0, x_l^*], [x_l^*, 1) \subset \mathcal{A}^*$, and $0 < p(x) < 1, \forall x \in (x_l^*, 1)$. This is enough to prove (ii).

We now proceed with the proof:

Step 1. We can prove this result by contradiction. Assume that the result is not true. Then for the optimal solution, consider $x^* \in \mathcal{A}^*$ and $\hat{x} \notin \mathcal{A}^*$ such that $0 < p(\hat{x}) < 1$. Then we have:

$$\begin{aligned} & x^* \{p(x^*)[\varepsilon L_e + (1 - \varepsilon)L_f] + (1 - p(x^*))L_e\} + (1 - x^*)L_f \\ & > \hat{x} \{p(\hat{x})[\varepsilon L_e + (1 - \varepsilon)L_f] + (1 - p(\hat{x}))L_e\} + (1 - \hat{x})L_f. \end{aligned}$$

Using the result in Lemma 2 and (11) we know that any candidate optimal solution must satisfy the following three conditions:

$$\alpha_t = \left(1 - \frac{\theta_m}{\theta_n}\right) \Delta, \quad (30)$$

$$\theta_n w(E(D) + E(S)) = \Delta, \quad (31)$$

$$0 \leq p(x) \leq 1, \forall x. \quad (32)$$

Denoting a small perturbation in the value of $p(\hat{x})$ by $dp(\hat{x})$, and using (30) and (31), we obtain:

$$\frac{d\theta_n}{dp(\hat{x})} = g_n(\hat{x})d\hat{x}; \quad \frac{d\theta_m}{dp(\hat{x})} = g_m(\hat{x})d\hat{x}, \quad (33)$$

$$\frac{d\alpha_t(\theta_n, \theta_m)}{dp(\hat{x})} = \frac{\partial\alpha_t}{\partial\theta_n} \frac{d\theta_n}{dp(\hat{x})} + \frac{\partial\alpha_t}{\partial\theta_m} \frac{d\theta_m}{dp(\hat{x})} = [\theta_m g_n(\hat{x}) - \theta_n g_m(\hat{x})] \frac{\Delta d\hat{x}}{\theta_n^2}, \quad (34)$$

We can use (30) in (31) to rewrite (31) as follows:

$$\phi(\theta_m, \theta_n, \alpha_t, \varepsilon) = \Delta - \theta_n w(E(D) + E(S)) = 0.$$

Using the Implicit Function Theorem on the above equation we obtain:

$$\frac{d\phi}{dp(\hat{x})} = 0 = \frac{\partial\phi}{\partial\theta_m} \frac{d\theta_m}{dp(\hat{x})} + \frac{\partial\phi}{\partial\theta_n} \frac{d\theta_n}{dp(\hat{x})} + \frac{\partial\phi}{\partial\alpha_t} \frac{d\alpha_t}{dp(\hat{x})} + \frac{\partial\phi}{\partial\varepsilon} \frac{d\varepsilon}{dp(\hat{x})},$$

or equivalently:

$$\frac{d\varepsilon}{dp(\hat{x})} = - \left[\frac{\frac{\partial\phi}{\partial\theta_m} \frac{d\theta_m}{dp(\hat{x})} + \frac{\partial\phi}{\partial\theta_n} \frac{d\theta_n}{dp(\hat{x})} + \frac{\partial\phi}{\partial\alpha_t} \frac{d\alpha_t}{dp(\hat{x})}}{\frac{\partial\phi}{\partial\varepsilon}} \right]. \quad (35)$$

Now choose the magnitude of $dp(\hat{x})$ such that x^* would continue to offer the highest benefit to terrorists if ε were left unaltered. Also, in (35), choose the sign of $dp(\hat{x})$ such that $d\varepsilon < 0$. It is always possible to do so because $0 < p(\hat{x}) < 1$, unless $d\varepsilon/dp(\hat{x}) = 0$. In the latter case, we adopt the convention that the principal would increase $p(\hat{x})$ to a value such that $d\varepsilon/dp(\hat{x}) \neq 0$, or set $p(\hat{x}) = 1$. This can be done because, while $d\varepsilon/dp(\hat{x}) = 0$, a change in $p(\hat{x})$ does not affect the principal's objective function value.

After choosing the appropriate $dp(\hat{x})$ we find that $d\varepsilon < 0$, i.e., the new value of residual risk is $\varepsilon' = \varepsilon + d\varepsilon$ such that $\varepsilon' < \varepsilon$. Two cases can arise:

Case 2.1: x^* continues to offer the greatest benefit to terrorists. Since $\varepsilon' < \varepsilon$, we have managed to improve the principal's objective function value, while ensuring feasibility. Hence, the previous solution could not have been optimal.

Case 2.2: x' now offers the greatest benefit to terrorists, where $x' \neq x^*$. We have:

$$\begin{aligned}
& x^* \{p(x^*)[\varepsilon L_e + (1 - \varepsilon)L_f] + (1 - p(x^*))L_e\} + (1 - x^*)L_f \\
& \geq x' \{p(x')[\varepsilon L_e + (1 - \varepsilon)L_f] + (1 - p(x'))L_e\} + (1 - x')L_f \\
& > x' \{p(x')[\varepsilon' L_e + (1 - \varepsilon')L_f] + (1 - p(x'))L_e\} + (1 - x')L_f.
\end{aligned}$$

Once again, since we have improved the objective function value, the previous solution could not have been optimal. Thus, we have shown that at the optimal solution, any container with risk score x , such that $0 < p(x) < 1$, offers the same (maximal) expected benefit to terrorists, denoted by k .

Step 2. To prove the “necessary” part of the claim made in this step, we need to establish that at optimality $p(x) = 1$ for some $x \in [0, 1) \implies d\varepsilon/dp(x) \leq 0$.

For such an x , if $x \notin \mathcal{A}^*$ and $d\varepsilon/dp(x) > 0$, then a reduction in the value of $p(x)$ would result in a lower ε , and therefore in an improvement in the principal’s objective function value, which would imply that the solution we started out with, could not have been optimal. (The reduction in $p(x)$ is small enough that x continues to be outside \mathcal{A}^* .)

We now rule out the possibility that $x \in \mathcal{A}^*$. If $x \in \mathcal{A}^*$, then we can eliminate the case that $p(1) = 0$. Otherwise ($x = 1$) offers benefit L_e and therefore has to belong to \mathcal{A}^* . This violates the strict monotonicity of $p(x)$ on \mathcal{A}^* , as shown in part(i) of the proposition. Similarly, we can also rule out the possibility that $0 < p(1) < 1$, because, if $0 < p(1) < 1$, then using the result in (i) we can conclude that $(x = 1) \in \mathcal{A}^*$. Once again, that would violate the same monotonicity property of $p(x)$. Finally we rule out the possibility that $p(1) = 1$ since we already showed that amongst the elements of \mathcal{A}^1 , only x_u^1 can belong to \mathcal{A}^* . Therefore, x cannot belong to \mathcal{A}^* . This concludes the proof of the “necessary” part of the claim.

To prove the “sufficiency” part of the claim we need to establish that at optimality, $\forall x \in [0, 1)$, the relationship $d\varepsilon/dp(x) \leq 0 \implies p(x) = 1$ holds.

Since an increase in $p(x)$ results in a reduction of ε provided $d\varepsilon/dp(x) < 0$, this necessarily implies an improvement in the objective function value. Hence, the principal will increase $p(x)$ till it hits the upper bound 1. Also, recalling our convention of setting $p(x)$ to 1 in case $d\varepsilon/dp(x) = 0$, we have now proved the sufficiency condition.

Step 3. We now relate the condition $d\varepsilon/dp(x) \leq 0$ to the existence of ζ . The critical value, ζ , can be obtained by evaluating the condition $\frac{d\varepsilon}{dp(x)} \leq 0$ in (35). The following should be helpful in evaluating (35):

$$\frac{\partial \phi}{\partial \theta_m} = -\theta_n w \frac{\partial E(D)}{\partial \lambda} \frac{\partial \lambda}{\partial \theta_m} = -\theta_n w \frac{\psi^2 + \sigma^2}{2(1 - \lambda\psi)^2} \Lambda F(\alpha_t) \leq 0, \quad (36)$$

$$\frac{\partial \phi}{\partial \theta_n} = -w(E(D) + \psi) - \theta_n w \frac{\partial E(D)}{\partial \lambda} \frac{\partial \lambda}{\partial \theta_n} = -w(E(D) + \psi) - \theta_n w \frac{\psi^2 + \sigma^2}{2(1 - \lambda\psi)^2} \Lambda(1 - F(\alpha_t)) \leq 0, \quad (37)$$

$$\frac{\partial \phi}{\partial \alpha_t} = -\theta_n w \frac{\partial E(D)}{\partial \lambda} \frac{\partial \lambda}{\partial \alpha_t} = -\theta_n w \frac{\psi^2 + \sigma^2}{2(1 - \lambda\psi)^2} \Lambda f(\alpha_t)(\theta_m - \theta_n) \geq 0, \quad (38)$$

$$\frac{\partial \phi}{\partial \varepsilon} = -\theta_n w \left[\frac{\partial E(D)}{\partial \varepsilon} + \psi_\varepsilon \right] = -\theta_n w \left[\frac{(1 - \lambda\psi)2\lambda\psi + \lambda^2(\psi^2 + \sigma^2)}{2(1 - \lambda\psi)^2} + 1 \right] \psi_\varepsilon > 0. \quad (39)$$

Using (33) - (39), we find that:

$$\begin{aligned} \frac{d\varepsilon}{dp(x)} = \frac{g_m(x)dx}{\partial \phi / \partial \varepsilon} & \left[\frac{\theta_n w \Lambda F(\alpha_t) E(D)}{\lambda(1 - \lambda\psi)} + \left(\frac{\Delta}{\theta_n} + \frac{\theta_n w \Lambda(1 - F(\alpha_t)) E(D)}{\lambda(1 - \lambda\psi)} \right) \frac{g_n(x)}{g_m(x)} \right] \\ & - \frac{g_m(x)dx}{\partial \phi / \partial \varepsilon} \left[\left(\frac{\theta_n w \Lambda f(\alpha_t) \alpha_t E(D)}{\lambda(1 - \lambda\psi)} \right) \left(\left(1 - \frac{\alpha_t}{\Delta} \right) \frac{g_n(x)}{g_m(x)} - 1 \right) \right] \end{aligned} \quad (40)$$

Defining $\eta = \frac{\Delta \lambda(1 - \lambda\psi)}{\theta_n^2 w \Lambda E(D)} > 0$, and using (39) and (40), we find that $\frac{d\varepsilon}{dp(x)} \leq 0$ if and only if:

$$\frac{g_n(x)}{g_m(x)} \geq \zeta = \frac{f(\alpha_t) \alpha_t + F(\alpha_t)}{f(\alpha_t) \alpha_t (1 - \frac{\alpha_t}{\Delta}) + F(\alpha_t) - 1 - \eta}. \quad (41)$$

If $\zeta \leq 0$, then we require that the denominator $f(\alpha_t) \alpha_t (1 - \frac{\alpha_t}{\Delta}) + F(\alpha_t) - 1 - \eta < 0$, which would imply that $\frac{d\varepsilon}{dp(x)} > 0$, $\forall x \in [0, 1]$. Then, together with the result in Step 2, we conclude

that $p(x) < 1$, $\forall x \in [0, 1)$.

Now consider $x \in \mathcal{A}^0$. Then:

$$\lim_{x \rightarrow x_l^*} [xL_e + (1-x)L_f] \leq x_l^* \{p(x_l^*)[\varepsilon L_e + (1-\varepsilon)L_f] + (1-p(x_l^*))L_e\} + (1-x_l^*)L_f. \quad (42)$$

Since $p(x_l^*) \geq 0$, therefore L_e is necessarily greater than or equal to $p(x_l^*)[\varepsilon L_e + (1-\varepsilon)L_f] + (1-p(x_l^*))L_e$. The relationship in (42) can hold only if $p(x_l^*) \rightarrow 0$. Hence, $\forall x \in [0, x_l^*]$, we have $p(x) = 0$.

iii In order to characterize the properties of the optimal $p(x)$ when $\zeta > 0$, we use the proofs of the steps in part (ii) of the proposition.

a) Using (40) and (41), we conclude that $\zeta > 0$ implies that for $x \in [0, 1)$ if $g_n(x)/g_m(x) < \zeta$, then $p(x) < 1$. Therefore, if $x \in [0, x_l^*]$, then $p(x) = 0$. On the other hand, if $x \in (x_l^*, 1)$ then we can rule out the possibility that $p(x) = 0$. Otherwise the result in part (i) of the proposition would be violated. Hence, $0 < p(x) < 1$, and using the result of Step 1 in the proof of part (ii), we can conclude that $x \in \mathcal{A}^*$ and satisfies (16). In summary, if $0 \leq g_n(x)/g_m(x) < \zeta$ then x and $p(x)$ satisfy the same properties as any other point when $\zeta \leq 0$.

b) For $x \in [0, 1)$, when $g_n(x)/g_m(x) \geq \zeta > 0$, then using (40) and (41), we conclude that $p(x) = 1$.

iv. We now verify that the optimal solution thus obtained will necessarily satisfy the condition $\theta_n > \theta_m$.

Assume that this is not the case. Then (9) implies that $\alpha_t = 0$, or that there is no C-TPAT membership.

We are now in the situation equivalent to the Base Case. We already know that the optimal $p(x)$ for the Base Case is a non-decreasing function of x . The strict FOSD relationship between $G_n(x)$ and $G_m(x)$ implies $E_n[p(x)] > E_m[p(x)]$ (where $E_\nu[p(x)]$ represents the expectation of $p(x)$ with respect to the function $G_\nu(x)$). The result $\theta_n > \theta_m$ follows, leading to a contradiction. Hence, the optimal solution in the above proposition will satisfy the condition $\theta_n > \theta_m$.

□

Corollary 1:

Proof We start with the optimal $p(x)$ obtained from the Base Case and use it in the context of C-TPAT. This means that θ_n is common to both scenarios. Since the optimal $p(x)$ from the Base Case is a non-decreasing function of x , therefore using the strict FOSD ordering of $G_n(x)$ and $G_m(x)$ we find that $\theta_m < \theta_n$. Using (11) and the result that $\varepsilon < 1$, from Lemma 2, we find that $\alpha_t > 0$. Then using our assumption that $F(x) > 0, \forall x > 0$, we infer that $F(\alpha_t) > 0$. Then using (10), we conclude that the effective arrival rate, λ , is lower in the C-TPAT case. We also know that the (IR_b) and (IR) constraints are identical, and that:

$$\frac{\partial E(D)}{\partial \lambda} = \frac{\psi^2 + \sigma^2}{2(1 - \lambda\psi)^2} > 0; \quad \frac{\partial E(D)}{\partial \varepsilon} = \left[\frac{(1 - \lambda\psi)2\lambda\psi + \lambda^2(\psi^2 + \sigma^2)}{2(1 - \lambda\psi)^2} \right] \psi_\varepsilon < 0.$$

Hence, we can conclude that a strictly lower ε is sustainable with C-TPAT, while using a common $p(x)$ (the optimal form from the Base Case) for both scenarios. From Lemma 1 we know that $\theta_n > 0$ and hence, $x_{lb}^* < 1$. Since, in this example, ε is lower for the C-TPAT scenario, we find that the expected benefit offered by each element of \mathcal{A}_b^* , except x_{lb}^* , is strictly lower than before, the latter being held constant. Hence, we can increase $p(x_{lb}^*)$ slightly in the C-TPAT scenario and achieve a strictly lower expected cost relative to the Base Case. This change does not violate feasibility, in particular the (IR) constraint is not violated, as the increase is made on a set of measure zero.

Alternately, such an increase can be made feasibly, if the reduction in ε for the C-TPAT scenario is chosen such that the (IR) constraint is just short of being tight.

With regard to the trading firms, we know, from Lemmas 1 and 2, that the participation constraint is binding in both scenarios. Hence, the non-member firms achieve zero surplus in both cases. However, the member firms incur a lower cost than member firms, as is evident from (9). □

Corollary 2:

Proof

- i. Follows directly from the result in Proposition 5 and the strict MLR property.
- ii. & iii. We know that for $x \in [x_l^*, x_u^*)$, the condition in (16) is satisfied. We also know that $p(x_l^*) = 0$. Hence,

$$O_P = k = x_l^* + (1 - x_l^*)L_f. \text{ Therefore, for } x \in [x_l^*, x_u^*), \text{ we find that } p(x) = \frac{1 - \frac{x_l^*}{x}}{1 - \varepsilon^*}.$$

Two cases arise: either the point $(x = 1) \in \mathcal{A}^*$, or $(x = 1) \notin \mathcal{A}^*$. In the former case, $p(1) = \frac{1 - x_l^*}{1 - \varepsilon^*} \leq 1 \implies \varepsilon^* \leq x_l^*$. For the latter case, using the result in part (i) of the proposition, we know that $p(1) = 1$, and the expected benefit offered at $(x = 1)$ is lower than $x_l^*L_e + (1 - x_l^*)L_f$, i.e.,

$$\varepsilon^*L_e + (1 - \varepsilon^*)L_f < x_l^*L_e + (1 - x_l^*)L_f \implies \varepsilon^* < x_l^*.$$

□

Proposition 6:

Proof

- i. Starting with the optimal solution for a given value of inspection capacity κ , if we increase κ slightly while keeping all other model parameters fixed, then since $\psi_\kappa < 0$, and

$$\frac{\partial E(D)}{\partial \psi} = \frac{(1 - \lambda\psi)2\lambda\psi + \lambda^2(\psi^2 + \sigma^2)}{2(1 - \lambda\psi)^2} > 0,$$

it is clear that the (IR) constraint is now slack. Using (29) and the assumption that $\psi_\varepsilon < 0$, we conclude that a strictly lower value of ε is now sustainable with the original $p(x)$. Since, in this example, ε is lower for the new scenario with greater capacity, we find that the expected benefit offered by each element of the original \mathcal{A}^* , except x_l^* , is strictly lower than before, the latter being held constant. Hence, we can increase $p(x_l^*)$ slightly in the scenario with greater capacity, and thereby achieve a strictly lower expected cost relative to the original scenario. This change does not violate feasibility, in particular the (IR) constraint is not violated, as the increase is made on a set of measure zero.

Alternately, such an increase can be made feasibly, if the reduction in ε , for the scenario with greater

capacity, is chosen such that the (IR) constraint is just short of being tight. Hence, there is an improvement in the objective function value. Since we have shown earlier that $O_P^* = x_l^* L_e + (1 - x_l^*) L_f$, therefore an increase in κ must result in a strict decrease in x_l^* .

ii. For the unique equilibrium, we know that:

$$\theta_n = \int_{x_l^*}^1 \frac{1 - \frac{x_l^*}{x}}{1 - \varepsilon^*} g_n(x) dx = \int_{x_l^*}^1 \frac{1 - \frac{x_l^*}{x}}{1 - x_l^*} g_n(x) dx; \quad \theta_m = \int_{x_l^*}^1 \frac{1 - \frac{x_l^*}{x}}{1 - x_l^*} g_m(x) dx.$$

Therefore,

$$\frac{d\theta_n}{dx_l^*} = \int_{x_l^*}^1 \frac{1 - \frac{1}{x}}{(1 - x_l^*)^2} g_n(x) dx < 0 \implies \frac{d\theta_n}{d\kappa} = \frac{d\theta_n}{dx_l^*} \frac{dx_l^*}{d\kappa} > 0.$$

By a similar analysis for θ_m we can conclude that $\frac{d\theta_m}{d\kappa} > 0$.

iii. We define:

$$\beta = \frac{\theta_m}{\theta_n} = \frac{\int_{x_l^*}^1 \left(1 - \frac{x_l^*}{x}\right) g_m(x) dx}{\int_{x_l^*}^1 \left(1 - \frac{x_l^*}{x}\right) g_n(x) dx}.$$

Then:

$$\frac{d\beta}{dx_l^*} = \frac{\left[\int_{x_l^*}^1 \left(1 - \frac{x_l^*}{x}\right) g_n(x) dx \right] \left[\int_{x_l^*}^1 \left(-\frac{1}{x}\right) g_m(x) dx \right] - \left[\int_{x_l^*}^1 \left(1 - \frac{x_l^*}{x}\right) g_m(x) dx \right] \left[\int_{x_l^*}^1 \left(-\frac{1}{x}\right) g_n(x) dx \right]}{\left[\int_{x_l^*}^1 \left(1 - \frac{x_l^*}{x}\right) g_n(x) dx \right]^2},$$

or equivalently:

$$\frac{d\beta}{dx_l^*} = \frac{[1 - G_m(x_l^*)] \int_{x_l^*}^1 \frac{1}{x} g_n(x) dx - [1 - G_n(x_l^*)] \int_{x_l^*}^1 \frac{1}{x} g_m(x) dx}{\left[\int_{x_l^*}^1 \left(1 - \frac{x_l^*}{x}\right) g_n(x) dx \right]^2}.$$

To evaluate the sign of $\frac{d\beta}{dx_l^*}$, we rely on the equivalence of MLR ordering and Uniform Conditional Stochastic Ordering (UCSO) (Whitt 1980), for the univariate case. Then for the non-decreasing func-

tion $-1/x$ where $x \in [x_l^*, 1]$, this implies:

$$\begin{aligned} \frac{\int_{x_l^*}^1 \left(\frac{-1}{x}\right) g_n(x) dx}{1 - G_n(x_l^*)} &\geq \frac{\int_{x_l^*}^1 \left(\frac{-1}{x}\right) g_m(x) dx}{1 - G_m(x_l^*)} \\ \implies [1 - G_m(x_l^*)] \int_{x_l^*}^1 \left(\frac{1}{x}\right) g_n(x) dx &\leq [1 - G_n(x_l^*)] \int_{x_l^*}^1 \left(\frac{1}{x}\right) g_m(x) dx. \end{aligned}$$

Therefore, we can conclude that $\frac{d\beta}{dx_l^*} \leq 0$, and $\frac{d\beta}{d\kappa} = \frac{d\beta}{dx_l^*} \frac{dx_l^*}{d\kappa} \geq 0$. Using (30) we know that:

$$\frac{d[F(\alpha_t)]}{d\kappa} = f(\alpha_t) \frac{d\alpha_t}{d\kappa} = -\Delta f(\alpha_t) \frac{d\beta}{d\kappa} \leq 0.$$

iv. Note that using (10) and the results in (ii) and (iii), we find that:

$$\begin{aligned} \frac{d\lambda}{d\kappa} &= \frac{\partial \lambda}{\partial \alpha_t} \frac{d\alpha_t}{d\kappa} + \frac{\partial \lambda}{\partial \theta_m} \frac{d\theta_m}{d\kappa} + \frac{\partial \lambda}{\partial \theta_n} \frac{d\theta_n}{d\kappa} \\ &= \Lambda f(\alpha_t)(\theta_m - \theta_n) \frac{d\alpha_t}{d\kappa} + \Lambda F(\alpha_t) \frac{d\theta_m}{d\kappa} + \Lambda(1 - F(\alpha_t)) \frac{d\theta_n}{d\kappa} > 0. \end{aligned}$$

□

C Appendix: The Last Mile

In this Appendix we discuss what more is required to fully characterize the optimal solution to the container inspection problem.

C.1 Base Case

Using Proposition 1-(ii) it is clear that the objective function value, O_P is strictly increasing in x_{lb}^* . Hence, the principal will try to make it as small as possible at the optimal solution. Therefore at optimality, either $x_{lb}^* = 0$ or the feasibility constraint on $p(x)$ is tight, i.e., $\varepsilon_b = x_{lb}^*$, provided ε_b is a continuous function of x_{lb}^* . We rule out the possibility that $x_{lb}^* = 0$, since we have assumed that there isn't enough inspection capacity to achieve that outcome. We know that if $\frac{d\varepsilon_b}{dx_{lb}^*}$ exists and is well defined everywhere then ε_b is continuous in x_{lb}^* .

Using (20), and differentiating with respect to x_{lb}^* we find:

$$\frac{d\phi}{d\theta_n} = 0 = \frac{\partial\phi}{\partial\theta_n} \frac{d\theta_n}{dx_{lb}^*} + \frac{\partial\phi}{\partial\varepsilon_b} \frac{d\varepsilon_b}{dx_{lb}^*}, \quad (43)$$

Using (3) and the expression for $p(x)$, we find:

$$\frac{d\theta_n}{dx_{lb}^*} = \frac{d\left(\int_{x_{lb}^*}^1 \frac{1 - \frac{x_{lb}^*}{x}}{1 - \varepsilon_b} g_n(x) dx\right)}{dx_{lb}^*} = \int_{x_{lb}^*}^1 \frac{-1}{1 - \varepsilon_b} g_n(x) dx + \frac{d\varepsilon_b}{dx_{lb}^*} \left[\int_{x_{lb}^*}^1 \frac{1 - \frac{x_{lb}^*}{x}}{(1 - \varepsilon_b)^2} g_n(x) dx \right]. \quad (44)$$

Now using the above result in (43), we find:

$$\frac{d\varepsilon_b}{dx_{lb}^*} = \frac{\frac{\partial\phi}{\partial\theta_n} \int_{x_{lb}^*}^1 \frac{g_n(x) dx}{x(1 - \varepsilon_b)}}{\frac{\partial\phi}{\partial\theta_n} \int_{x_{lb}^*}^1 \frac{1 - \frac{x_{lb}^*}{x}}{(1 - \varepsilon_b)^2} g_n(x) dx + \frac{\partial\phi}{\partial\varepsilon_b}} \quad (45)$$

Using Lemma 1 we rule out the possibility that $\varepsilon_b = 1$, which is equivalent to no inspection. Now plugging in the values of $\frac{\partial\phi}{\partial\theta_n}$ and $\frac{\partial\phi}{\partial\varepsilon_b}$ from (22) and (23) respectively, we find that $\frac{d\varepsilon_b}{dx_{lb}^*}$ is well defined, and therefore ε_b is continuous in x_{lb}^* , provided the denominator in the RHS of (45) is never equal to zero. If the latter

condition is true, then we can claim that at optimality $\varepsilon_b^* = x_{lb}^*$. Also, since the objective function is strictly increasing in x_{lb}^* , therefore the optimal value of x_{lb}^* would be unique. This implies that the optimal value $\varepsilon_b^* = x_{lb}^*$ would be unique, and using the expression $p(x) = \frac{1 - \frac{x_{lb}^*}{x}}{1 - \varepsilon_b^*}$, $\forall x \in [x_{lb}^*, 1)$ along with (3), we can conclude that the optimal value for θ_n would be unique.

C.2 C-TPAT

We identify sufficient conditions, subsequent to the results in Corollary 2, under which ε^* varies continuously with x_l^* , and results in a unique equilibrium outcome for the container inspection problem with C-TPAT.

CBP is still left with the following optimization problem:

$$\min O_P = \min_{x_l^* \in [0,1], \varepsilon^*} [x_l^* L_e + (1 - x_l^*) L_f],$$

s.t.

$$\varepsilon^* \leq x_l^*.$$

The other constraints that have to be adhered to are (30) and (31). Note that the objective function in the optimization problem is strictly increasing in x_l^* . Also, if ε^* is a continuous function of x_l^* , then at the optimal solution either $x_l^* = \tau$ or $\varepsilon^* = x_l^*$. We have assumed that there isn't enough inspection capacity to achieve the former outcome, hence, it must be the case that $\varepsilon^* = x_l^*$ at optimality. This necessarily implies that $x_u^* = 1$ and $p(x_u^*) = 1$, where $x_u^* \triangleq \sup \mathcal{A}^*$.

In order to establish that ε^* is a continuous function of x_l^* , it is sufficient to show that $d\varepsilon^*/dx_l^*$ exists and is well defined. Using (31) and (41), we define the following two functions:

$$\phi(\theta_m, \theta_n, \alpha_t, \varepsilon^*) = \Delta - w\theta_n[E(D) + \psi] = 0, \quad (46)$$

$$\Psi(\theta_m, \theta_n, \alpha_t, \varepsilon^*, \zeta) = \zeta - \frac{f(\alpha_t)\alpha_t + F(\alpha_t)}{f(\alpha_t)\alpha_t(1 - \frac{\alpha_t}{\Delta}) + F(\alpha_t) - 1 - \eta} = 0, \quad (47)$$

We now define the sets: $\chi = \{x | g_n(x)/g_m(x) \geq \zeta\}$, and $\Phi = \text{Null Set}$. Then:

$$\begin{aligned} x_u^* &= 1, \text{ if } \zeta \leq 0; \\ &= 1, \text{ if } \zeta > 0 \text{ and } \chi = \Phi. \\ &= \inf \chi, \text{ if } \zeta > 0 \text{ and } \chi \neq \Phi. \end{aligned}$$

This relationship can be represented graphically in the Figure below. Assuming $g_n(x)/g_m(x)$ is a smooth

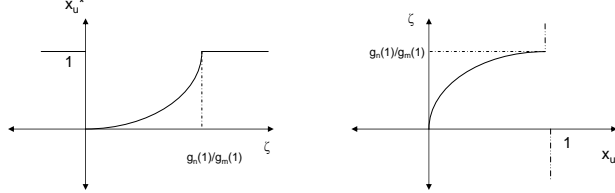


Figure 6: The relationship between ζ and x_u^* .

function of x , we can conclude that $d\zeta/dx_u^*$ exists and is well defined $\forall x_u^* \in [0, 1)$. We now use the Implicit Function Theorem on (46) and (47) by evaluating $d\phi/dx_l^* = 0$ and $d\Psi/dx_l^* = 0$. We find:

$$\frac{\partial \phi}{\partial \theta_m} \frac{d\theta_m}{dx_l^*} + \frac{\partial \phi}{\partial \theta_n} \frac{d\theta_n}{dx_l^*} + \frac{\partial \phi}{\partial \alpha_t} \frac{d\alpha_t}{dx_l^*} + \frac{\partial \phi}{\partial \varepsilon^*} \frac{d\varepsilon^*}{dx_l^*} = 0, \quad (48)$$

$$\frac{\partial \Psi}{\partial \theta_m} \frac{d\theta_m}{dx_l^*} + \frac{\partial \Psi}{\partial \theta_n} \frac{d\theta_n}{dx_l^*} + \frac{\partial \Psi}{\partial \alpha_t} \frac{d\alpha_t}{dx_l^*} + \frac{\partial \Psi}{\partial \varepsilon^*} \frac{d\varepsilon^*}{dx_l^*} + \frac{\partial \Psi}{\partial \zeta} \frac{d\zeta}{dx_l^*} = 0. \quad (49)$$

To help simplify the above expressions, we note that:

$$\frac{d\alpha_t}{dx_l^*} = \frac{d \left[\left(1 - \frac{\theta_m}{\theta_n} \right) \Delta \right]}{dx_l^*} = \frac{-\Delta}{\theta_n} \frac{d\theta_m}{dx_l^*} + \frac{\theta_m \Delta}{\theta_n^2} \frac{d\theta_n}{dx_l^*},$$

$$\frac{d\zeta}{dx_l^*} = \frac{d\zeta}{dx_u^*} \frac{dx_u^*}{dx_l^*},$$

$$\begin{aligned} \frac{d\theta_n}{dx_l^*} &= \frac{d \left[\int_{x_l^*}^{x_u^*} \left(\frac{1 - \frac{x_l^*}{x}}{1 - \varepsilon} \right) g_n(x) dx + [1 - G_n(x_u^*)] \right]}{dx_l^*} \\ &= - \int_{x_l^*}^{x_u^*} \frac{g_n(x) dx}{x(1 - \varepsilon)} + \frac{d\varepsilon}{dx_l^*} \left[\int_{x_l^*}^{x_u^*} \left(\frac{1 - \frac{x_l^*}{x}}{(1 - \varepsilon^*)^2} \right) g_n(x) dx \right] + \left(\frac{1 - \frac{x_l^*}{x_u^*}}{1 - \varepsilon^*} \right) g_n(x_u^*) \frac{dx_u^*}{dx_l^*} - g_n(x_u^*) \frac{dx_u^*}{dx_l^*} \end{aligned}$$

$$\begin{aligned}
\frac{d\theta_m}{dx_l^*} &= \frac{d \left[\int_{x_l^*}^{x_u^*} \left(\frac{1 - \frac{x_l^*}{x}}{1 - \varepsilon} \right) g_m(x) dx + [1 - G_m(x_u^*)] \right]}{dx_l^*} \\
&= - \int_{x_l^*}^{x_u^*} \frac{g_m(x) dx}{x(1 - \varepsilon)} + \frac{d\varepsilon}{dx_l^*} \left[\int_{x_l^*}^{x_u^*} \left(\frac{1 - \frac{x_l^*}{x}}{(1 - \varepsilon)^2} \right) g_m(x) dx \right] + \left(\frac{1 - \frac{x_l^*}{x_u^*}}{1 - \varepsilon^*} \right) g_m(x_u^*) \frac{dx_u^*}{dx_l^*} - g_m(x_u^*) \frac{dx_u^*}{dx_l^*}
\end{aligned}$$

Using the above relationships in (48) and (49), we observe that we have a pair of simultaneous linear equations in two variables - $\frac{d\varepsilon^*}{dx_l^*}$ and $\frac{dx_u^*}{dx_l^*}$ - which can be represented as follows:

$$\phi_1 \frac{dx_u^*}{dx_l^*} + \phi_2 \frac{d\varepsilon^*}{dx_l^*} + \phi_3 = 0,$$

$$\Psi_1 \frac{dx_u^*}{dx_l^*} + \Psi_2 \frac{d\varepsilon^*}{dx_l^*} + \Psi_3 = 0.$$

These can be solved for $\frac{d\varepsilon^*}{dx_l^*}$. Provided $\phi_1 \Psi_2 - \phi_2 \Psi_1 \neq 0$, this is sufficient to establish that ε^* is continuous in x_l^* . Finally, Lemma 2 (i) rules out the possibility that $\varepsilon = 1$. Thus, under the stated sufficient conditions, we can conclude that $\varepsilon^* = x_l^*$, at optimality.

From the optimization problem that CBP is left with, it is clear that, since we are choosing the smallest possible value of x_l^* , the optimal value of x_l^* is unique. This uniquely identifies the values of ε^* , θ_n , θ_m , and α_t ; thus bearing out the claim that the optimal solution is unique.

D Appendix: Glossary of Notation

- x : risk score, or the conditional probability that a container concealing terrorist contraband escapes detection up through the primary inspection stage.
- b : base probability of terrorist attack in the period of analysis.
- c_a : cost to terrorists of planning and carrying out an attack with WMD.
- L_e : losses inflicted on the US if terrorist contraband escapes detection and enters into the country.
- L_f : losses engendered by the economic dislocation resulting from a terrorist weapon being found inside a maritime container.
- τ : deterrence threshold, or the threshold value of risk score such that lower risk scores offer terrorists an expected benefit that is lower than their cost of carrying out the attack.
- $G_n(x)$: CDF associated with the distribution of risk scores in the container population of a non-member trading firm.
- $g_n(x)$: pdf associated with the distribution of risk scores in the container population of a non-member trading firm.
- $G_m(x)$: CDF associated with the distribution of risk scores in the container population of a member trading firm.
- $g_m(x)$: pdf associated with the distribution of risk scores in the container population of a member trading firm.
- ε_b : the residual risk post secondary inspection in the Base Case, or probability of not detecting a terrorist weapon during secondary inspection.
- ε : the residual risk post secondary inspection in the C-TPAT scenario, or probability of not detecting a terrorist weapon during secondary inspection.

- κ : scaled inspection capacity, which is a composite of equipment and labor resources available with CBP for inspecting containers.
- S : service time associated with secondary inspection of a container; $E(S)$ denotes its first moment; $E(S^2)$ denotes its second moment.
- ψ : speed-accuracy-tradeoff function that gives the value of expected service time as a function of ε and κ .
- σ^2 : variance of service time for secondary inspection.
- ϕ : random component of the service time, with mean zero and variance σ^2 .
- $p(x)$: probability of tagging a container with risk score x , for secondary inspection.
- $E(D)$: expected waiting time in the secondary inspection queue.
- Λ : raw arrival rate of containers to a US port.
- λ : effective arrival rate of containers to the secondary inspection facility.
- θ_n : probability of inspecting a container belonging to a non-member firm.
- θ_m : probability of inspecting a container belonging to a member firm.
- Δ : upper bound on the per container congestion cost that can be borne by a trading firm.
- d_i : waiting cost per unit time, for firm i .
- r_i : mean revenue per container, for firm i .
- w : waiting cost per unit time per dollar of revenue, for firm i .
- P_i : penalty for failing an audit, for member firm i .
- B_i : upper bound on penalty for failing an audit, for member firm i .

- q : frequency of audit in the period of analysis.
- $c_i(q)$: allocated cost per container of undergoing an audit, for member firm i .
- γ_i : per container cost of compliance with C-TPAT, for firm i .
- α : normalized cost of compliance, or compliance cost; $\alpha = \frac{\gamma_i + qc_i(q)}{r_i}$
- α_t : threshold value of compliance cost, below which trading firms participate in C-TPAT.
- $F(\alpha)$: CDF associated with the distribution of compliance cost amongst trading firms.
- $f(\alpha)$: pdf associated with the distribution of compliance cost amongst trading firms.
- β : multiple of per-container benefit from shirking in the specification of B_i .
- O_P : CBP's objective function for the container inspection problem.
- \mathcal{A}_b^* : set of risk scores which offer the maximum benefit to terrorists, for the Base Case.
- \mathcal{A}^* : set of risk scores which offer the maximum benefit to terrorists, for the C-TPAT scenario.
- x_{lb}^* : infimum of the set \mathcal{A}_b^* .
- x_l^* : infimum of the set \mathcal{A}^* .
- x_u^* : supremum of the set \mathcal{A}^* .