How to Quantify Deterrence and Reduce Critical Infrastructure Risk

Eric F. Taquechel and Ted G. Lewis

ABSTRACT

We propose a definition of critical infrastructure deterrence and develop a methodology to explicitly quantify the deterrent effects of critical infrastructure security strategies. We leverage historical work on analyzing deterrence, game theory and utility theory. Our methodology quantifies deterrence as the extent to which an attacker's expected utility from an infrastructure attack changes after a defender has invested to deter attacks, as compared to their expected utility absent deterrence. We derive expected utilities from a modified game theory approach, which uses probabilistic utility functions, wherein utility function probabilities are functions of investment. We vary these functions based on different information availability assumptions (e.g., perfect vs imperfect attacker information). We produce evidence that it is quantifiably more advantageous to overtly deter, rather than conceal security information, under specific conditions. We also leverage these utility functions to determine the unconditional risk to a defender if deterrence strategies fail, and we determine cost efficiency of those strategies.

INTRODUCTION

DHS policy advocates investing in critical infrastructure (CI) protection to deter terrorist attacks.[1] The *National Infrastructure Protection Plan* is replete with references to deterrence, most notably on page 38: "account for the adversary's ability to recognize the target and the deterrence value of existing security measures." Also, Homeland Security Presidential Directive (HSPD) 7 emphasizes deterrence of CI attacks.[2] Since DHS also advocates quantitative terrorism risk analysis, to support the goal to deter attacks, DHS would benefit from a methodology to quantify the deterrence value of CI terrorism risk reduction investments in a way that complements and enhances existing traditional approaches to risk analysis. Also, DHS would benefit from a working definition of CI deterrence, the lack of which is identified by Morral and Jackson.[3]

Given this policy context, various elements within DHS have begun efforts to analyze deterrence, or influence adversary decision-making before a CI attack is executed. This necessarily involves considering human factors; specifically, thinking about the adversary's approach to terrorism planning. For example, the US Coast Guard, in conjunction with the University of Southern California's Center for Risk and Economic Analysis of Terrorism Events (CREATE), has begun the development of PROTECT, a model intended to help Coast Guard units deter adversary planning by patrolling CI in a random fashion.[4] Recent work at the Naval Postgraduate School in Monterey, CA has produced a model that measures the resiliency of supply chains, accounting for perceived attacker preferences for disrupting a supply chain and increasing commodity shipment costs.[5] Additionally, DHS Science and Technology (S&T) recently undertook a review of multiple methodological approaches to enhancing traditional risk analysis by incorporating insights from intelligent adversary modeling, which accounts for adversary planning and goals.[6]

While well intentioned and promising, these efforts do not explicitly quantify deterrence in the context of critical infrastructure protection (CIP) or explain how to incorporate measurable deterrence into CIP risk equations to show the CI risk reduction effects of deterrence.[7] If DHS components are to implement policies to deter attacks, eventually they may be asked to account to Congress on how they are measuring the effectiveness of that implementation and how that implementation plays with traditional risk reduction. At present, some DHS agencies such as the US Coast Guard report risk-based performance metrics without explicitly accounting for measurable deterrence. Looking ahead, it is possible that the Coast Guard and other DHS agencies will have to start reporting deterrence effectiveness and resulting risk reduction accounting for this effectiveness.[8] This is our present motivation to develop a quantitative approach to deterrence measurement and to develop a method to incorporate that approach into traditional CIP risk analysis.

In the spirit of this DHS emphasis on deterrence, and assuming DHS components may eventually have to report measures of effectiveness, we propose that CI deterrence can be quantified as the extent to which an attacker's intent to attack a CI target changes. This change is a result of changes in expected utility from a potential attack on a CI after a CI defender attempts to deter an attack, as compared to what the attacker's expected utility from a potential attack would have been before the defender deterred. Utility is the value of an outcome.

Our approach is intended to complement existing, traditional CIP risk analysis tools, not replace them. In fact, our approach relies on input from reputable risk analysis tools. Our example of how we apply this approach focuses on deterring terrorism. However, as long as existing risk analysis tools or best practices can be used to analyze the threat, vulnerability, and consequences of other intentional hazards such as sabotage and theft, our general principles can be applied to measuring deterrence value of measures against any sort of attack on a CI. Indeed, some of our literature review will examine how deterrence principles, which we leverage in our own approach, have been applied to issues other than terrorism. Importantly, our approach does not consider unintentional threats to these CIKR such as natural disasters or industrial accidents[9]. Finally, our approach incorporates elements of the DHS prevent, protect, respond, and recover approach, consistent with Homeland Security Presidential Directive (HSPD) 8.[10]

DETERRENCE AND RISK: MUTUALLY INCLUSIVE IN THEORY

Deterrent tactics include investments to reduce attacker expected utility by influencing attacker capability, target vulnerability, and target consequence. These terms, along with intent to attack, make up the basic DHS risk equation *Risk=T*V*C*. Risk is expected loss to a CI defender from a potential CI attack, but this loss constitutes part of an attacker's gain, or utility. This equation incorporates all aspects of prevention protection, response and recovery as articulated in HSPD-8: reducing threat is preventing an attack, reducing vulnerability is protecting against an attack, and reducing consequences is responding to and recovering from an attack. Thus, if we measure deterrence as a function of changing components of threat, vulnerability and consequence, our deterrence efforts encompass the entire HSPD-8 spectrum.

By influencing these components of attacker utility, a defender may attempt to influence an attacker's belief that the costs of an attack outweigh the benefits or utility, or that the probability of successful attack is too low, so they don't execute that course of action (COA). This is operational deterrence as proposed by Morral and Jackson.[11] Thus, deterrence strategies must convince the attacker than if they execute a certain COA, it will yield benefits less than those yielded by their best course of action without deterrence. Hence, these strategies must account for our attackers' goals as well as our own, as long as we can influence their perspective on attaining their goals. This lends to a game theory approach to quantifying deterrence. If we determine changes in attacker expected utility in a game theory format, we can measure change in intent and thus can measure deterrence.

Deterrence thus inherently includes our adversary's assessment of their risk, or perceived reduction in utility, as well as our assessment of our own risk. Ultimately, when deliberating deterrence strategies, we want to know answers to three questions: (1) what is the extent to which the attacker is deterred, (2) what is the risk reduction, or change in expected defender loss resulting from that deterrence strategy; and (3) what are the cost implications, or defender deterrence investment efficiency, of that deterrence strategy. All answers must be quantifiable to be useful in quantitative deterrence analysis and terrorism risk management. Taken together, deterrence quantification, defender expected loss as a result of that quantifiable deterrence, and defender investment efficiency constitute a deterrence "portfolio" of criteria a defender must consider. This portfolio would support existing CIKR security plans and risk analysis with additional rigor and measurable deterrence effect data.

GENERAL FORM OF OUR DETERRENCE QUANTIFICATION EQUATION

Formally, quantifiable CI deterrence effectiveness of an lth defender deterrence course of action (COA) given an attacker's kth post-deterrence action, is the difference between attacker pre-deterrence intent, which we denote $Intent_i^{pre}\Big|_{\$AB}$, or probability an attacker desires to attack a certain target, and attacker post-deterrence intent $Intent_k^{post}\Big|_l$, which is then normalized by the pre-deterrence intent. The i and k subscripts denote specific attacker COA; i is an attacker's pre-deterrence COA and k is an attacker's post-deterrence COA. This is shown in Equation 1.

$$E_l\Big|_k = \frac{Intent_i^{pre}\Big|_{\$AB} - Intent_k^{post}\Big|_l}{Intent_i^{pre}\Big|_{\$AB}}$$

Equation 1. Quantification of deterrence

These intent values are functions of attacker expected utilities from an attack. These expected utility values are outputs of a deterrence game in a game theory format.

For the attacker, expected utilities are probabilities of an attack succeeding given it is already desired (attacker capability multiplied by target vulnerability), multiplied by the maximum utility (in dollars) if the attack were to succeed. The attacker's utility or defender's consequence is a combination of deaths[12] and economic consequence (in dollars of annualized profit for that target), and the economic consequence is multiplied by a mitigation factor based on the extent to which the defender could mitigate that economic consequence if an attack was successful.

Threat is generally attacker intent * attacker capability; our methodology only incorporates the capability part of the traditional threat probability equation in the attacker expected utility function. This is because we treat intent as the output of our expected utility calculation methodology, rather than an input, but we apply intent into the defender' s unconditional risk equation to reflect how CI risk changes given quantifiable deterrence so as not to wholly depart from a traditional risk analysis approach. We now elaborate on our approach to leveraging threat with utility functions.

OUR PHILOSOPHY ON THREAT AND UTILITY FUNCTIONS

The risk equation R(risk) = TVC (threat*vulnerability*consequence) is the general standard for CIKR risk analysis because it has a threat component. There are numerous vulnerable and highly consequential CIKR in the US, and leveraging threat judgments helps differentiate and reduce the defender's workload based on the likelihood that an attacker may prefer some target types more than others. However, that threat is often high level and strategic,[13] without accounting for target nuances that may influence attacker perception of specific target attractiveness. Incorporating our quantifiable deterrence, as a function of intent changes based on individual target expected attacker utility, adds this nuance and we believe this more accurately represents unconditional risk.

We claim that intent is not critical component of an attacker expected utility function. We invest to deter and influence desirability or intent. Traditionally, intent is an input to the standard risk equation, as a component of threat. However, we claim that desirability is a function of attacker beliefs about probability of successful attack initiation, execution, consummation, and consequence of that consummation.[14] Thus, deterrence investments must influence these factors, but desirability or intent is an outcome of that process. Intent then need not be included in probabilistic risk functions or utility functions in a deterrence game.

There are those, such as Louis Cox who believe that threat, or intent and capability to attack, should be an outcome of a probabilistic risk analysis product rather than an input.[15] The outcome of the games in our research will result in intent probabilities. We will show that only intent need be an outcome; we believe that capability probabilities should remain inputs within a risk or utility equation. Intent is still useful in a CIKR risk equation when considering how to allocate resources based on risk, from the defender perspective only. This means if we don't explicitly consider deterrence effects, or changes in attacker intent as a result of changing expected utility, we can justifiably leave intent in the risk equation.

Excluding intent, expected loss should include capability, or the probability an attacker can initiate an attack. The defender invests in offensive, counterterrorism actions to reduce attacker capability to initiate and execute an attack. This is one aspect of investing to deter. If capability reduction investments fail, the attacker is not deterred because they believe a high probability that the attack can be initiated. But, expected loss should also include the probability that the attack can be successfully executed, and will bring about a consequence to the target once consummated, which are components of vulnerability. If vulnerability investments fail, the attacker is not deterred because they may believe a high probability that the attack, once initiated, can be executed and consummated.

Finally, expected loss must include consequence, but only to the extent it is not reduced by defender recoverability and redundancy measures.[16] These measures constitute the final aspect of investing to deter, because if the attacker believes any consequence will be quickly mitigated, they may be deterred as they may not achieve their goals.

Thus, expected utility functions in a deterrence game must include capability probabilities, vulnerability probabilities, maximum consequence values, and probabilities of consequence mitigation. Intent or desirability probabilities are not needed. If intent probabilities were included in expected loss equations, since intent reflects the desirability to carry out an attack, and desirability is a function of probability of successful initiation, execution and consummation, then we would be double counting the deterrent effects of defender investment in those expected utility functions. Thus, our expected utility functions will treat expected utility as a conditional probability: they reflect the expected utility of attack, if the attack already desires to attack, or intent=100%.

DETERRENCE AND RISK: MUTUALLY INCLUSIVE IN PRACTICE

Our utility functions upon which the intent values in Equation 1 are based, account for the traditional risk=threat*vulnerability*consequence approach. In fact, our methodology depends on traditional risk analysis tools to provide this input; it cannot function without existing risk=TVC tools such as the Coast Guard's Maritime Security Risk Analysis Model (MSRAM),[17] TRAM, CARVER, or any legacy probabilistic CI risk analysis tool that leverages threat, vulnerability, consequence, and mitigation data. Thus, our methodology is not intended to supplant existing probabilistic CI risk methodologies and tools, but complement them and extend their functionality. Additionally, myriad CI protection efforts such as Design Basis Threat analysis, Crime Prevention through Environmental Design, and others can easily be leveraged with our proposed methodology, as they contribute to our understanding of threat, vulnerability and consequence-all critical input for our methodological approach. Thus, even if a CIP best practice is not formally considered a risk assessment tool or methodology, if that practice produces data that can meet the basic definition of threat, vulnerability, or consequence, then its input could be leveraged by our deterrence quantification methodology.

Equation 1 will help us answer the first deterrence strategy question. The deterrence game that produces those utilities that support intent values in Equation 1 can also produce defender risk equations, which can be compared to answer the second deterrence strategy question. Further, attacker and defender utilities must be modeled as functions of attacker and defender investment, since return on investment is critical for risk-informed decision making.[18] This will help answer the third deterrence strategy question.

Also, Equation 1 can be applied to quantify deterrence subject to specific theoretical analysis options for how a CI defender and an attacker perceive the utility of different COA. These analysis options include (1) assumptions about information availability circumstances, or the extent to which an attacker knows what a defender is doing to defend their CI; (2) different utility theory assumptions, or assumptions about people assess information and make decisions about the value of an outcome; and (3) assumptions about how risk probabilities are reduced with defender investment.

PRACTICAL IMPLICATIONS AND USEFULNESS OF THIS APPROACH

WHY SHOULD WE HAVE THE ABILITY TO QUANTIFY DETERRENCE?

This quantitative approach can be applied to real-world CI protection. Since we are quantifying deterrence in a way that will leverage traditional probabilistic CI risk data, those involved in CI protection can report a different metric to our policymakers, in a way that reinterprets the effect of CI attack capability, vulnerability and consequence mitigation measures.

To explain, the current performance metric that DHS components report on is how much risk, or expected loss, we have reduced to CI by implementing vulnerability and consequence reduction measures. The focus is on how these measures would defeat an attack already in progress, or how they would mitigate effects after the attack. However, if we leverage this same data to quantify the deterrent value of those measures, we could instead report how we believe we have changed the attacker's expected utility from an attack, and thus have changed their intent to attack, by implementing these measures. Thus, we gain insights into how risk changes based on our adversary's interpretation of our defenses, and we have now accounted for an intelligent adversary. This is the reinterpretation of the effects of our current measures, contextualized for policy that requires both risk analysis and deterrence. Also, current risk methodologies generally only account for an engineering approach to risk management, as they focus primarily on the details of vulnerability and consequence. They include threat as well, but often this is a very high-level, non specific threat probability estimation. So, our methodology will account for more tactical aspects of threat probabilities, such as how individual target engineering characteristics such as security measures actually influence the attacker's interest in attacking, and thus threat judgments can be more granular and target-specific.

That said, it seems necessary to be able to incorporate our quantitative deterrence effectiveness results back into the standard risk=TVC equation. There are theoretical challenges with this, but we believe they are easily overcome. In brief, adopting our deterrence quantification methodology and incorporating its results in back into traditional probabilistic risk=TVC, where T represents intent and capability, might seem to double count the effects of our capability, vulnerability, and consequence factor reduction deterrent measures. It would count those deterrence effects once against those factors themselves, and again in intent. However, we claim that unconditional risk, or risk including the probability the attacker desired to attack after deterrence measures are implemented, appropriately double-counts the effects of those deterrence investments: once in the reduction of capability, vulnerability, and/or consequence, and again in intent. We must account for the likelihood the attacker wants to attack, and the likelihood of success if they do initiate an attack. The current form of risk used for CIP is unconditional risk where we believe the intent probability modifies conditional risk, and conditional risk is capability*vulnerability*consequence: the expected loss given an attack is already desired with 100% intent.

This approach is useful for a wide audience: academics who will consider the technical aspects of our methodology, CIP practitioners who may implement the deterrence measures that inform our methodology once the methodology is accepted within academia; and DHS agencies who must report performance effectiveness measures to Congress and other oversight bodies after practitioners have implemented the measures.

We expound on how we incorporate measurable deterrence into the traditional risk equation in a case study example. We also give examples of how to apply this methodology to notional real world CI in a subsequent section of this research. Finally, we give examples of how quantifiable deterrence effectiveness might change given different information availability circumstances, as that seems the most practical of the theoretical analysis options we have introduced: a practitioner naturally wants to know whether they should make their security measures overt or covert to deter! The other two analysis options are areas of interest but we do not give examples of how deterrence effectiveness might change if those options are varied; we save this for future research. We discuss the theories behind those analysis options briefly in our analysis of the need for these options and our literature review.

WHY THE THEORETICAL ANALYSIS OPTIONS? DIFFERENT DETERRENCE EFFECTIVENESS DATA MAY RESULT

Equation 1 can be applied to quantify deterrence subject to specific theoretical analysis options for how a CI defender and an attacker perceive the utility of different COA. These analysis options include assumptions about information availability circumstances, or the extent to which an attacker knows what a defender is doing to defend their CI. In game theory modeling, whether actors have perfect (known) or imperfect (unknown) information about their opponents is critical to understanding the game's outcomes. Since we will be leveraging principles from game theory in our approach, we focus on this analysis option. These theoretical analysis options also include different utility theory assumptions, or how assumptions about people assess information and make decisions about the value of an outcome. Nikhil Dighe et. al. urge that future work on analyzing deterrence investment should consider alternatives to traditional rationality, or subjective expected utility (SEU).[19] Utility functions can assume either (SEU) or an alternative, prospect utility theory (PU).

Finally, these analysis options include assumptions about how risk is reduced with defender investment. In theory, probabilities of attack success may change linearly with investment, but in reality, the relationship may be an exponential function of investment; it becomes increasingly difficult to reduce risk as more risk is reduced. Dighe et. al. advocate exploration of the effects of security investments when those investments do not reduce the success probability of attacks to zero.[20] This nonlinearity can be observed in engineering problems; for CI security measures, both engineering factors and social factors such as public resistance to increased security may contribute to nonlinearity.

The practical implications of experimenting with the different theoretical analysis options are numerous. Uncertainties facing an attacker are critical to understanding deterrence, but are rarely leveraged in game-theoretic analysis of counterterrorism.[21] We have done initial work on this and have produced evidence suggesting whether it is more advantageous for a CI defender to make defensive investments unknown, or imperfect, to would-be attackers rather than perfectly known. The defender advantage in such cases would be lower average unconditional risk after deterrence under specific information availability circumstances.

If evidence supports an advantage in deterrence effectiveness under imperfect information circumstances, implications are that a government should consider classifying its CI defensive investments and any analysis of how those investments reduce infrastructure vulnerability and consequence. Conversely, if evidence supports an advantage of perfect information, a government should instead actively communicate to would-be CI attackers not only its defensive investments but also detailed calculations of how those investments reduce vulnerability and consequence. This approach assumes the attacker believes the defender's disclosure of investment and effectiveness information is accurate.

In order to generalize these findings, any advantage of a specific information availability circumstance must be robust given different utility theory assumptions, and also given different probability-investment relationship assumptions. Behavioral science and historical research support the need to analyze decision making under nonlinear or biased circumstances and thus examine different assumptions about utility when we analyze deterrence. Nonlinear engineering phenomena support the need to analyze how system performance, or CI attacker capability, vulnerability, and consequence mitigation in CIP research, changes nonlinearly as a function of effort or investment expended. We do not give examples of sustained defender advantage from a specific information availability circumstance across these different analysis option assumptions in this paper, but this is an area for future research.

DISCUSSION OF RELEVANT WORK

We now examine relevant work to create context for our general methodology to quantify CI deterrence, and for our specific theoretical analysis options. We integrate different aspects of this work with our own techniques into a methodology to quantify CI deterrence and create a deterrence portfolio. There is historical work to measure deterrence effects in other areas, including fishery enforcement, drug smuggling, and aviation terrorism, but DHS currently lacks a methodology to explicitly quantify CI terrorism deterrence. However, there is plenty of research on quantitative risk measurement, analysis, and management, which suggests we should leverage quantitative risk to quantify deterrence.

RISK, RETURN ON INVESTMENT, AND DETERRENCE

DHS defines risk as a function of threat, vulnerability, and consequence.[22] For terrorism, threat is defined as the likelihood of an attack being initiated, accounting for both adversary intent and capability. Intent is the probability an attack is desired, and capability is the probability it can be executed. Vulnerability is the likelihood an attack is successful given it is attempted. Finally, consequences are the effects a terrorist attack. Thus, this approach to analyzing terrorism risk is probabilistic in nature. The probabilistic approach to risk analysis has been widely accepted for some time.[23]

Since DHS is using probabilistic risk analysis techniques for counterterrorism,[24] the relationship between probabilistic risk analysis and general deterrence theory must be examined before we attempt to quantify deterrence effectiveness of CI protection strategies. In general, there is much precedent for integrating risk analysis with deterrence. Richard Lebow and Janet Stein propose that deterrence effectiveness depends upon an actor's attitude toward risk.[25] Elaine Bunn advocates that deterrence policies must account for adversary attitudes toward risk.[26] However, these arguments are not specifically focused on CI attack deterrence, and contain no explicit quantitative calculations of probabilistic risk that would support deterrence quantification.

Other literature focuses on possible relationships between CI risk analysis and deterrence. For example, William McGill and Bilal Ayyub claim that deterrence is a function of infrastructure vulnerability, a component of infrastructure risk.[27] In fact, the probability an attack would be initiated, or threat, is claimed to be a function of the perceived target attractiveness, and the target attractiveness is a function of the perceived probability of success, or vulnerability. Thus, a lower vulnerability target means a less attractive target; a less attractive target then means a lower probability of attack initiation; and a lower attack initiation probability then means the attacker has been deterred. This suggests a fundamental relationship between probabilistic risk and deterrence. We must now examine deterrence theory more closely and put our proposed definition of CI deterrence measurement in context of this historical discussion.

DETERRENCE THEORY

Deterrence theory has been vigorously discussed and debated for decades. Patrick and Audrey Cronin propose that deterrence occurs when an actor discourages aggression toward another actor, with the intended outcome that the former never has to respond to aggressive action by the latter.[28] In other words, the goal of deterrence is to convince an adversary to refrain from aggressive action. Lebow claims that deterring involves manipulating an adversary's assessment of a situation by convincing them that costs of acting exceed benefits.[29] Finally, putting theory into practical guidance the Department of Defense Deterrence Operations Joint Operating Concept specifies that deterrence affects adversary decision making in three ways: (1) imposing costs on an undesirable adversary course of action; (2) denying the perceived benefits of such a COA; and (3) encouraging restraint by making it seem more attractive than action.[30]

These approaches to deterrence suggest that the only way to deter is to convince an adversary not to attack. But, applying this approach to deterring terrorist attacks on CI is problematic for two reasons. One, the United States has so many CI that spending resources to prevent attacks on all seems inefficient. Two, these deterrence theories were developed during the Cold War when nation-states could be deterred by threat of mutually assured destruction. With asymmetrical warfare and terrorism, convincing our adversaries we can eliminate them is a challenge since they are difficult to find. Thus, a modified definition of deterrence is needed for CI protection. An all-or-nothing approach applied to DOD tactic (1) is impractical because it is difficult to impose costs. For (2), denying benefits would mean defending all infrastructures, which we attempt to do at great expense. Finally, for (3), our adversaries have stated their intent to continue to attempt to inflict damage; restraint will likely never be acceptable to organizations such as Al-Qaeda whose raison d'être is to inflict damage upon the United States and its allies. Thus, our form of deterrence for CI protection – influencing an adversary's decision making process such that their expected utility from attacking a CI changes after we deter by investing – is appropriate.

The idea of accepting losses is not without precedent. For example, John Major proposes the concept of equilibrium expected loss (EEL): a defender moves resources to various high-risk infrastructures, making them equally unattractive to an attacker and creating an EEL. Any unprotected infrastructure, if attacked, will yield a gain to the attacker that is less than the EEL.[31]

Two critical components of deterrence are credibility and signaling. Bunn claims that credibility is critical: to deter, we must be able to signal to the adversary we have both the credible capability and the will to impose costs or deny benefits.[32] Thus, for CI deterrence, a terrorist must believe that we can deny them benefits by defending our infrastructure. That belief can be reinforced by the terrorist's own observations of our defenses, and our explicit signaling of those defenses. The former type of reinforcement begs the question of whether the terrorist accurately perceives our defenses, and the latter begs the question of whether the terrorist believes our disclosures. These issues have been researched; for example Erik Jenelius et. al. analyze the effects of adversary perception on their decision making.[33] The present research assumes that prior to deterrent investments at the start of a deterrence game, all information about the defender's original "pre-deterrence" defenses is known to an attacker, and subsequent deterrence game outcomes determine how deterrence is quantified depending on whether the attacker has perfect or imperfect information about the defender's deterrence COA. Thus, our signaling of information should be a key determinant of how quantifiably effective our deterrence investments are. There is the possibility that an attacker could infer weaknesses if we signal prospective defenses before they are implemented. Importantly, we do not advocate actually signaling the existence of security or response measures before they are implemented. We only argue that when we are considering what deterrence investments to implement, we should determine whether making those investments transparent or obscured, once actually implemented, would influence the attacker more. Also, we can run our deterrence game simulation with as many CIKR targets as desired, using computer programs if necessary, to account for the possibility that signaling defenses at targets will focus the attacker attention elsewhere. We now know there is a relationship between probabilistic risk analysis and deterrence, and we have a working definition of CI attack deterrence as a change in attacker intent. Thus, we will use a probabilistic risk approach to quantify deterrence. Next, we examine previous and ongoing efforts to analyze and quantify deterrence.

ANALYZING DETERRENCE

There is a considerable repository of literature offering insights into how deterrence might be quantified, without ever explicitly stating how to quantify it for CI protection. In general, the literature shows that analysis of drug smuggling, fisheries law enforcement, critical infrastructure security, and aviation security, reveals that lowering the chances of successful attack or violation of a law, and communicating the imposition of penalties if caught in violation, tended to deter would-be violators. One study gleaned that the subjects interviewed would be deterred if they knew the high probabilities of failure if they attempted to violate a law, but they had difficulty estimating those probabilities with any rigor. Other studies more explicitly focused on critical infrastructure security, postulating that the change in expected utility or outcome of a CI attack would influence the probability an attacker would want to attack in the first place. Much of the more recent literature leverages ideas from utility theory and game theory, and thus we leverage the ideas of influencing probabilities and consequences, along with game theory and utility theory, in our approach to quantify deterrence. We now expound on these theories.

GAME THEORY AND THEORETICAL ANALYSIS OPTIONS

Game theory is an analytical approach designed to help understand interactions between competing or cooperating decision makers. It treats decision makers as "players" participating in a "game" with certain rules and assumptions.[34] Two of these assumptions are that decision makers are rational and that they simultaneously account for their opponents' anticipated behavior when choosing a course of action. Game theory is an appropriate tool for studying the strategic interaction between governments and terrorists.[35]

Game theory assigns an expected "payoff" value or expected utility from certain courses of action to each player. A payoff is the subjective value of the outcome of a course of action. It is assumed that each player wants to maximize his own payoff or utility.[36] But, game theory often predicts an outcome in which each player has a payoff which may be less than that which they could gain without the influence of their opponent. Our approach will create payoffs or utility functions to be used in a game, from traditional risk components: capability, vulnerability, and consequence. Ultimately, utility is a function of probabilistic risk. We treat the defender's probabilistic loss as the attacker's gain or utility in the present research. We also derive detailed expected utility functions for an attacker, reflecting the belief that terrorists perform cost-benefit analyses.[37]

We also examine how perfection or imperfection of attacker information, assumptions which we call information availability circumstances, influences the quantifiable effectiveness of deterrence and other components of our deterrence portfolio. In traditional game theoretical approaches, information imperfection, or lack of information about all opponent moves, lends to a pure strategy Nash Equilibrium (NE) as the predicted outcome of the game, whereas perfect information generally lends to a mixed strategy NE as the predicted game outcome in a Stackelburg leader-follower game.

However, we do not follow the traditional game theoretical approach in our research. In essence, we eschew the traditional game theoretical approach of leveraging a pure Nash Equilibrium (NE) or a mixed strategy, to account for the likelihood that adversary decision making is more complicated that what a simple equilibrium predicts. So, for distinguishing between the effects of information availability upon deterrence effectiveness, we create utility functions in our deterrence game that reflect either perfect attacker information, or one of what we propose are three possible "information imperfection biases" that the attacker is assumed to have: neutral, conservative, and overconfident. These biases are part of assessing the adversary in a design-basis threat approach.

A neutral bias means an "average attacker", who will assume the defender deters by reducing target consequence and vulnerability by 50% of what they were prior to deterrence investment. A conservatively biased attacker will assume the defender deters by reducing target consequence and vulnerability by 95% of what they were prior to deterrence investments,[38] and an overconfidently biased attacker will assume the defender invests nothing to deter; thus post-deterrence attacker expected utilities are the same as pre-deterrence expected utilities. We offer these biases as a start to exploring what we call a "theoretical analysis option" for quantifying deterrence; more research into modeling attacker decision making under imperfect information conditions is an area for future research. There is some literature on this subject, for example see Jenelius et. al.[39]

There are additional "theoretical analysis options" to apply to our methodology. These options will allow us to vary the composition of the utility functions that we will use in our methodology to quantify deterrence. Game theory leverages utility functions and we use a basic game in our approach, and so we expound upon utility theory and explain different views on how people assess subjective value in our extended literature review. In brief, subjective expected utility claims people make decisions linearly; whereas prospect utility claims people make decisions nonlinearly, subject to biases. There is experimental evidence to support nonlinear decision-making; hence a well-informed approach to quantifying deterrence must consider these different ideas as a sensitivity analysis. Prospect utility would require us to modify components of our utility functions with weights and other modifiers to reflect nonlinear biased decision-making.

The other theoretical analysis option is probability-investment relationships. We model utility function components such as capability and vulnerability probabilities as functions of investments, to capture cost information for return on investment calculations. That said, evidence from engineering sciences supports a nonlinear relationship between effort applied to solve a problem and the extent to which the problem is solved. There are linear and nonlinear (exponential) ways to model how these probabilities change as we invest more money to reduce them.

Now that we have examined literature on how deterrence might be quantified, and introduced theories that are relevant to creating the utility functions in a game theoretical approach to quantifying deterrence, we will explain our general methodology to quantify deterrence in context of a case study. Also, we will show how to quantify defender risk as result of the deterrence strategies in the case study, and we will show how to determine deterrence investment efficiency or ROI. Before we move to our case study, we summarize the applicability of our different theoretical analysis options to CI attack deterrence in Table 1:

Information Availability Circumstance	Perfect Information	Applied if we believe the attacker knows everything about the defender's deterrence investments
	Imperfect Information	Applied if we believe the attacker does not know everything about the defender's deterrence investments; the attacker biases their estimates based on whether they are overconfident, conservative, or neutral decision makers
Utility Theory	Subjective Expected Utility	Applied if we believe the attacker analyzes probabilities of attack success and probabilities of producing consequences, in a linear fashion
	Prospect Utility	Applied if we believe the attacker analyzes probabilities of attack success and probabilities of producing consequences in a nonlinear fashion. They bias toward overweighing expected utility when in the domain of loss, or behind a reference point. Thus, an attacker may be more likely to prefer attacking a target that objectively is very hard to attack or would suffer minimal consequence, if a successful attack will put them closer to their goals or reference point.
Probability-Investment Relationship	Linear	Applied if we believe the attacker's and defender's investments change attack success and consequence mitigation probabilities linearly (e.g. it is possible to attain 100% vulnerability elimination)
	Exponential	Applied if we believe the attacker's and defender's investments change attack success and consequence mitigation probabilities nonlinearly, or more specifically exponentially (e.g. it is impossible to attain 100% vulnerability elimination; and it becomes progressively harder to reduce the same amount of vulnerability with the same dollar amount, as we continue to reduce vulnerability)

Table 1: Applicability of Theoretical Analysis Options to CIP

DETERRENCE QUANTIFICATION METHODOLOGY - A REAL WORLD CASE STUDY

Since we are proposing a methodology to support risk analysis with deterrence measurement, and risk analysis is largely a quantitative process, we develop a technical approach to quantify deterrence. We summarize the results of this technical approach here in the context of a case study, focused on the potential deterrence effectiveness of notional FEMA Port Security Grant Program (PSGP) infrastructure security grants.[40] In general, we must compare attacker expected utilities from both before and after deterrence investments, those investments being PSGP grants improve CIKR security, in order to quantify deterrence.

In our example we create deterrence portfolios showing the quantification of deterrence given certain deterrence COA, how defender unconditional risk changes as a result of that quantifiable deterrence, and defender investment efficiency or ROI of those deterrence COA. We also can show how the results of these portfolios vary based on information availability circumstances: either the attacker has perfect information and defender deterrence COA perform a certain way, or the attacker has imperfect information and the defender deterrence effectiveness may vary depending on the attacker's information imperfection biases. We assume subjective expected utility and a linear probability-investment relationship throughout our example in this paper; future work would analyze deterrence effectiveness under different utility theory and probability-investment relationship assumptions.

In our example we leverage notional CI risk data from an existing DHS CI risk analysis tool, the US Coast Guard's Maritime Security Risk Analysis Model (MSRAM), though any CI risk analysis tool that leverages threat, vulnerability, consequence data, and mitigation measures could be used. Our methodology incorporates MSRAM risk data and extends the interpretation of that data. We will demonstrate how our methodology can be applied across multiple DHS CI sectors, as our notional CIKR competing for grant allocations to deter attacks are in different CIKR sectors[41].

Importantly, in our example we assume the federal government allocates the deterrence funding. Obviously, this assumption can be changed depending on the scenario modeled. We also assume an attacker is not motivated to attack a specific target type (e.g., an eco-terrorist only wanting to attack chemical refineries). Rather, the attacker is expected to want to maximize its possible utility, irrespective of target type. But, the methodology could be applied to include only targets of a specific type in future research.

CASE STUDY: QUANTIFYING DETERRENCE EFFECTIVENESS OF DEFENDER INVESTMENTS, AND RESULTING TERRORIST ATTACK RISK REDUCTION FOR A CHEMICAL FACILITY AND A FERRY TERMINAL

Suppose we want to determine the quantifiable deterrence effectiveness of different CI investments at two different targets: a maritime chemical facility which we denote target A, and a ferry terminal which we denote target B. The "defender" is the government and target owner/operators working in collaboration, and the design-basis threat for the "attacker" is a non-specific terrorist organization that attacks these targets in one specific way (e.g., boat bomb). Our two notional targets are in different DHS CI sectors; yet they often exist together in a local operational environment such as a port. We note that we only use two targets for simplicity in illustrating the methodology; however our methodology can be applied to as many targets as desired. The deterrence game can be expanded as necessary; computer programming may be required to capture CIKR in a large game.[42]

The defender has already invested to reduce vulnerability and consequence at these targets, and thus we can estimate how we have reduced risk in the traditional sense. However, given DHS encouragement to examine deterrence, and given that agencies currently report risk reduction performance effectiveness metrics to Congress and thus conceivably may be required to report deterrence effectiveness and resulting risk reduction in the future, we now want to analyze the potential deterrent effects of hypothetical future deterrence investments, in a quantifiable way.

There are eleven basic steps to quantifying deterrence and creating our deterrence portfolio to address all three deterrence strategy questions introduced at the beginning of this paper. In sum, there are three phases of analysis: pre-deterrence analysis, post-deterrence analysis, and comparison to quantify deterrence and create the deterrence portfolio. This portfolio would add data to the existing library of CIKR security plans and supporting decision-making. We discuss these three analysis phases in context of our case study.

PRE-DETERRENCE ANALYSIS FINDINGS

Pre-Deterrence Attacker Expected Utilities

The notional expected utility values in Table 2 include capability to attack, vulnerability to attack as a function of security investments, and consequence mitigated by recovery/ response measures. For example, suppose the intelligence community estimates a terrorist organization's capability to initiate an attack with a boat bomb against either type of facility. Also, prior to implementation of PSGP grants intended to deter, suppose both the chemical facility and ferry terminal have measures to detect an attack in progress (e.g., cameras); training and equipment to engage an attack before it reaches the focal point of the infrastructure for detonation (such as armed guards stationed within a reasonable response time from where an attack would initiate); and has the means to defeat that attack (guards have proper weapons and training). Finally, suppose each facility has business continuity plans and coordinates with stakeholders to ensure recoverability if attacked. These are all pre-deterrence mitigation capabilities. These data are assumed to represent (1) the attacker's perceptions of their attack capabilities and (2) target vulnerability and consequence based on attacker "scoping" reflecting extensive research and observations of the above-listed target attributes.

Attacker *i*th attack option	Notional expected utility value ($) $U_e T_i^{pre}$
Attack chemical facility	$257,142.86
Attack ferry terminal	$763,636.36
Attack both	$1,020,779.22
Refrain from attack	$0

Table 2: Notional attacker pre-deterrence expected utilities

Table 2 shows that given our notional pre-deterrence vulnerability, and consequence data for the chemical facility and ferry terminal, and attacker capabilities to attack each respectively, attacking both targets simultaneously yields the greatest expected utility to the attacker. This is intuitive, but we quantify that intuition here in dollars, and in comparison to the utility of other attack options.

Pre-Deterrence Attacker Intent Values

Given the above expected utilities, we can create intent values:

Attacker *i*th attack option	**Notional intent value (%)** $Intent_i^{pre}\Big\vert_{\$AB} = \dfrac{U_e T_i^{pre}\Big\vert_{\$AB}}{\sum\limits_{i=4}(U_e T_i^{pre}\Big\vert_{\$AB})}$
Attack chemical facility	12.60%
Attack ferry terminal	37.40%
Attack both	50.00%
Refrain from attack	0.00%

Table 3: Notional attacker pre-deterrence intent values

Table 3 shows that attacker intent to attack both targets simultaneously is greatest. Since Table 2 showed the greatest expected utility to results from this attacker option, it is no surprise that this option results in the greatest attacker intent value.

Pre-Deterrence Defender Unconditional Risk

We then apply these intent values in Table 3 to conditional risk values, to get unconditional risk values:

Attacker *i*th attack option	**Notional unconditional risk value ($)** $Risk_i^{pre}\Big\vert_{\$AB} = \dfrac{(Cap_i^{pre} V_i^{pre} Con_i^{pre}\Big\vert_{\$AB})^2}{\sum\limits_{i=4}(U_e T_i^{pre}\Big\vert_{\$AB})}$
Attack chemical facility	$32,388.22
Attack ferry terminal	$285,634.98
Attack both	$510,389.61
Refrain from attack	$0.00

Table 4: Notional defender pre-deterrence unconditional risk values

Not surprisingly, Table 4 shows that the unconditional risk of having both targets attacked simultaneously is greatest. So, going forward we can expect that the attacker's greatest expected utility before we deter would result from attacking both targets simultaneously; thus their intent is greatest. And, our unconditional expected loss, or risk, from having both targets attacked simultaneously, is our greatest risk.

POST-DETERRENCE ANALYSIS FINDINGS

Post-Deterrence Attacker Expected Utilities

We note that there are sixteen attacker expected utilities post-deterrence, as opposed to four pre-deterrence, because the defender has four different deterrence investment options. These options are (1) to invest PSGP grant money at only the chemical facility (for instance to increase the CCTV capabilities to detect); (2) invest grant money at only the ferry terminal (for instance to provide more training and equipment for terminal security personnel); (3) award grants to both; or (4) award grants to neither.[43] The attacker post-deterrence expected utilities are as a function of these investment options are:

Attacker *k*th attack option	**Notional expected utility value ($), defender invests at chemical facility** $U_eT_k^{post}\big\vert_{\$A}$	**Notional expected utility value ($), defender invests at ferry terminal** $U_eT_k^{post}\big\vert_{\$B}$	**Notional expected utility value ($), defender invests at both** $U_eT_k^{post}\big\vert_{\$AB}$	**Notional expected utility value ($), defender refrains from investment** $U_eT_k^{post}\big\vert_{\$0}$
Attack chemical facility	$57,142.88	$257,142.86	$57,142.86	$257,142.86
Attack ferry terminal	$763,636.36	$190,000.00	$190,000.00	$763,636.36
Attack both	$820,779.22	$447,142.86	$247,142.86	$1,020,779.22
Refrain from attack	$0.00	$0.00	$0.00	$0.00

Table 5: Notional attacker post-deterrence expected utilities, defender deters by investing in four different ways

Table 5 shows that given our notional post-deterrence vulnerability, and consequence data for the chemical facility and ferry terminal, and attacker post-deterrence capabilities to attack each respectively, attacking both targets simultaneously yields the greatest expected utility to the attacker, for each of the four defender deterrence COA. This notional data may reflect defender efforts to reduce capability (prevent), reduce vulnerability (protect), and reduce consequence (response and recover), though in this example the specific PSGP grants are intended to reduce vulnerability. Overall the attacker would gain the greatest expected utility if they attacked both and the defender invested nothing in deterrence, as seen in Table 5. This seems intuitive, but here we quantify this intuition.

Post-Deterrence Intent Values

Given the expected utilities described above, we can create sixteen intent values.

Attacker *k*th attack option	**Notional intent value (%), defender invests at chemical facility** $Intent_k^{post}\Big\|_{\$A}$	**Notional intent value (%), defender invests at ferry terminal** $Intent_k^{post}\Big\|_{\$B}$	**Notional intent value (%), defender invests at both** $Intent_k^{post}\Big\|_{\$AB}$	**Notional intent value (%), defender refrains from investment** $Intent_k^{post}\Big\|_{\$0}$
Attack chemical facility	3.48%	28.75%	11.56%	13.62%
Attack ferry terminal	46.52%	21.25%	38.44%	37.55%
Attack both	50.00%	50.00%	50.00%	50.00%
Refrain from attack	0.00%	0.00%	0.00%	0.00%

Table 6: Notional attacker post-deterrence intent values, defender deters by investing in four different ways

Table 6 shows that attacker intent to attack both targets simultaneously always dominates, as would be expected since attacker expected utility is always greatest.

Post-Deterrence Defender Unconditional Risk

We then apply these intent values in Table 6 to conditional risk values, to get unconditional risk values.

Attacker *k*th attack option	**Notional unconditional risk value ($), defender invests at chem. facility** $Risk_k^{post}\Big\|_{\$A}$	**Notional unconditional risk value ($), defender invests at ferry terminal** $Risk_k^{post}\Big\|_{\$B}$	**Notional unconditional risk value ($), defender invests at both** $Risk_k^{post}\Big\|_{\$AB}$	**Notional unconditional risk value ($), defender invests from investment** $Risk_k^{post}\Big\|_{\$0}$
Attack chemical facility	$1,989.15	$73,938.84	$6,606.11	$32,388.22
Attack ferry terminal	$355,235.90	$40,367.41	$73,034.68	$285,634.98
Attack both	$410,389.61	$223,571.43	$123,571.43	$510,389.61
Refrain from attack	$0.00	$0.00	$0.00	$0.00

Table 7: Notional defender post-deterrence unconditional risk values, defender deters by investing in four different ways

Table 7 shows that regardless of what the defender does to deter, the greatest risk is of having both targets simultaneously attacked; but intuitively – by investing at both targets – this risk is minimized. These unconditional risk scores reflect the vulnerability reduction effects of the notional grant investments, but also reflect the deterrence effects (changing attacker intent) of the investments. Now, we compare this pre-deterrence and post-deterrence data to develop our deterrence portfolio.

COMPARING PRE-DETERRENCE AND POST-DETERRENCE DATA

Quantification of Deterrence

We compare intent values from Tables 3 and 6, using Equation 1 to quantify the deterrence effectiveness of each defender COA, given all possible attacker options.

Attacker *k*th attack option	**Deterrence Effectiveness (%), defender invests at chem. facility** $E_{\$A}\Big\|_k$	**Deterrence Effectiveness (%), defender invests at ferry terminal** $E_{\$B}\Big\|_k$	**Deterrence Effectiveness (%), defender invests at both** $E_{\$AB}\Big\|_k$	**Deterrence Effectiveness (%), defender invests from investment** $E_{\$0}\Big\|_k$
Attack chemical facility	72.36%	-128.29%	8.22%	0.00%
Attack ferry terminal	-24.37%	43.23%	-2.77%	0.00%
Attack both	0.00%	0.00%	0.00%	0.00%
Refrain from attack	N/A	N/A	N/A	0.00%

Table 8: Notional quantification of deterrence effectiveness of four different defender deterrence COA

Notice that investing at target A reduced attacker intent to attack target A nearly 72%, whereas the negative sign for attacking B means the attacker is actually incentivized to attack target B. Thus, this deterrence COA is quantified as 72% effective at deterring an attack against target A, whereas it is 24% *ineffective* at deterring an attack against target B. It is intuitive that we would transfer intent to B by deterring attacks on A, and we can quantify this intuition here. Now, we want to know the deterrence effectiveness of investing at target B.

As expected, investing at target B incentivizes the attacker to attack target A, by increasing intent 128%, whereas it reduces intent to attack B and thus deters by 43%. Thus, this transfers intent to attack target A. Given that we have to make a decision on where to invest to deter, would we rather invest at A or B? To support this decision, we must know how defender unconditional risk changes given each possible deterrence COA. Transfer of intent and quantifiable deterrence effectiveness are useful metrics but are also means to an end: determining the resulting change in unconditional risk.

Finally, we examine the deterrence effectiveness of investing at both targets. The attacker is marginally incentivized to attack target B, perhaps because of the large consequence as compared to that of target A, and the attacker is slightly deterred from attacking target A in response. As with the previous two deterrence investment options, we must determine the resulting defender conditional risk across all attacker options, if the defender deters by investing at both targets. We note that the deterrence effectiveness of refraining from investment is 0% for all attacker options; attacker post-deterrence intent does not change from pre-deterrence intent if the defender does nothing.

New Defender Unconditional Risk

Given this deterrence quantification data, we now analyze the second component of a deterrence portfolio, or how defender unconditional risk changes as a result of quantifiable deterrence. We leverage Table 7 data and aggregate defender unconditional risk given an *l*th deterrence COA.

Deterrence COA	**Aggregate defender risk ($)** $\sum_{k=4}(Risk_k^{post}\big|_{\$l}) = Risk_A^{post}\big|_{\$l} + Risk_B^{post}\big|_{\$l} + Risk_{AB}^{post}\big|_{\$l}$
Invest at chemical facility	$767,704.66
Invest at ferry terminal	$337,877.67
Invest at both	$203,212.22
Invest at neither	$828,412.81

Table 9: Comparing aggregate defender unconditional risk given certain defender deterrence investments

Overall, the best quantifiably effective defender deterrence COA is to deter by investing at both the chemical facility and the ferry terminal, as the resulting unconditional risk is lowest. The least quantifiably effective deterrence COA is to do nothing, as the risk is highest. From a deterrence effectiveness and unconditional risk reduction perspective, these seem intuitive decisions, but with our methodology we can quantify the deterrent effect of these COA and quantify the impacts on risk reduction, relative to the impacts of other COA. Thus we see relative strengths and weakness of each option, and a decision maker can improve their risk reduction and deterrence performance measurement reporting beyond just reporting "we've invested to deter and reduce risk." Also, this technique of aggregate defender unconditional risk across all attacker options assuages the concern with transferring intent or risk from one target to another;[44] if we simply compare all possible outcomes before deterrence and after deterrence, we now have a metric that acknowledges possible transfer of intent between targets, but looks at the big picture. We believe a "transfer of intent" does not necessarily equate to a transfer of risk. We must consider the capability to execute an attack, vulnerability to attack, and consequence of an attack on all targets in the game. If the resulting intent to attack target B, when combined with these other risk factors, renders an overall lower unconditional risk to B than the unconditional risk of an attack on target A, in fact we have not transferred risk.

Return on Investment

Suppose the grant to defend the chemical facility with additional CCTV is for $1 million; the grant to defend the ferry terminal by training and equipping security personnel costs $2 million, and $3 million is the cost to defend both simultaneously. We show the ROI of each deterrence COA:

Deterrence COA	**ROI (unitless)** $ROI_l = \frac{\frac{Risk_{AB}^{pre}\Big\vert_{\$AB} - \frac{\sum_{k=4}(Risk_k^{post}\big\vert_l)}{4}}{Risk_{AB}^{pre}\Big\vert_{\$AB}}}{\$_l}$
Invest at target A	2.66*10^-7
Invest at target B	3.01*10^-7
Invest at both	3.12*10^-7
Invest at neither	N/A

Table 10: ROI of certain defender deterrence investments (ROI unit-less since risk reduction in $)

Notice that there is no ROI of refraining from investment, as we cannot divide by $0. Since we invested the same total dollar amount for each *l*th deterrence COA in this example, we get the best ROI from investing in both targets, as that created the lowest aggregate defender unconditional risk as seen in Table 9. This is an additional performance measurement for a decision maker to report. However, if we were to use different dollar values for the different *l*th COA, the ROI rankings might be different.

Using Tables 2 through 10, we can now create deterrence portfolios for each *l*th deterrence COA. For example the deterrence portfolio for *l*th COA=investing at target A, with respect all possible *k*th attacker post-deterrence options, is:

$$\left[\begin{array}{c} \begin{bmatrix} E_{\$A}\Big|_{A} = 72.36\% & E_{\$A}\Big|_{B} = -24.37\% \\ E_{\$A}\Big|_{AB} = 0\% & E_{\$A}\Big|_{0} = n/a \end{bmatrix} \\ \\ \overline{Risk_k^{post}}\Big|_{\$A} = \$191{,}903.67 \\ \\ ROI_{\$A} = 2.66 * 10^{-7} \\ \\ \end{array}\right]$$

Figure 1: Values for deterrence portfolio, defender deters by investing at chemical facility alone (target A), perfect information

In Figure 1, $\overline{Risk_k^{post}}\Big|_{\$A}$ represents average unconditional risk resulting from deterring at target A. We can repeat this consolidation of Tables 2 through 10 for the remaining three deterrence COA, which will show that investing at both targets is the best option for reducing risk and ROI. These deterrence portfolios are a succinct way for government agencies and CI owners/operators to report the deterrence effectiveness of deterrence measures, and resulting CI risk reduction and ROI. Decision makers may value different components of these portfolios differently; if both the government and CI operators will invest to deter, then they must decide what the most important component of the deterrence portfolios are. If industry is the only contributor of deterrence funding, then perhaps ROI may be the most important component. If government is the only contributor, the public may expect risk to be reduced as much as possible, regardless of ROI. Our methodology simply supports the need to analyze the data; decision makers must then make the decisions.

COMPARING CASE STUDY RESULTS TO THOSE OF A CASE STUDY WITH IMPERFECT INFORMATION ASSUMPTIONS

Since we assumed perfect information for the results of our case study, we want now to show how these deterrence portfolios change for different attacker information imperfection biases. We show this analysis, using the same notional data as before but including each of the three attacker information imperfection biases: neutral, overconfident, and conservative. These biases are part of the scenario development, often referred to as “design-basis threat.” This analysis is available from the authors, and has practical implications for decision makers. We summarize our limited analyses:

1. We get counterintuitive results for when a neutral attacker is assumed; we can actually show that when we have little to invest at the chemical facility (target A) it actually makes sense to make that information known to the attacker, because the average unconditional risk to the defender is less if the deterrence investment is known given this deterrence COA. In contrast, if we have a lot to invest at the chemical facility, we can show that it makes sense to withhold that information. These findings are due to the disparity in the notional maximum economic consequence values of the chemical facility and the ferry terminal, and may not hold if we change the notional target data. Thus, in circumstances similar to those of our notional example, a decision maker may prefer to make a counterintuitive decision and ensure perfect information if they deter by investing at the chemical facility, even if they do not invest much.
2. For when we assume an overconfident attacker, it makes sense for the defender to withhold information if their deterrence COA is to invest at a target with lower maximum economic value. This seems counterintuitive as in general if we have an overconfident attacker, making our deterrence information known to the attacker might reduce their confidence, but we show that this is not necessarily true. Thus, a decision maker in circumstances similar to those of our notional examples may prefer to invest at the chemical facility but ensure that the details of that investment are withheld from an overconfident attacker.
3. Finally, for when we assume a conservative attacker, it makes sense for the defender to communicate deterrence information if their deterrence COA is to invest at a target with a lower maximum economic value. This also seems counterintuitive as in general if we have a conservative attacker they might be emboldened by information about what the defender is doing to deter, but we show this is not necessarily true. Thus, a decision maker in circumstances similar to those of our notional examples may prefer to invest at the chemical facility, and should ensure the details of that investment are communicated to a conservative attacker.

DETERRENCE METHODOLOGY AND APPLICATION: SUMMARY

We have shown how to quantify deterrence effectiveness of different defender courses of action, based on the changes in attacker intent values after the defender deters, as compared to the attacker intent values before the defender deters. These intent values are functions of attacker pre-deterrence and post-deterrence expected utilities, and those utilities are functions of attacker capability to attack, CI target vulnerability, and CI consequence if attacked. Together, capability, vulnerability, and consequence make up defender conditional risk, or the risk given an attack is already desired (intent=100%).

We have also shown how to quantify the average defender unconditional risk, or conditional risk multiplied by the attacker intent probability, both before and after measurable deterrence. Finally, we have shown how to calculate return on investment of the deterrence COA, and we have shown how to incorporate these data into a deterrence portfolio which CI officials and government agencies can use to report both quantifiable deterrence and resulting CI risk reduction/ROI. This approach allows a practitioner to implement DHS policy, which advocates both deterrence and risk reduction, in an integrated fashion. Our methodology is general enough to be applied to CI in multiple sectors, as our example demonstrated, and the general concepts of capability, vulnerability, and consequence can be applied to any type of intentional act.

Our example of how our methodology would be applied produced deterrence portfolios of four different deterrence COA. Overall, we found that investing in both targets in a two-target game yielded the best quantifiable deterrence effectiveness, the lowest unconditional risk to the defender, and the best ROI. Our brief analysis of a deterrence portfolio under imperfect information circumstances revealed some counterintuitive results with implications for decision makers.

DIRECTION FOR FUTURE WORK

Our methodology and findings are based on many assumptions. Future work on quantifying CI deterrence must consider these assumptions and vary them in order to discover new insights into this challenge. We offer the following specific suggestions.

CIKR SECTOR SPECIFICS

While we believe our proposed methodology is general enough to be applied to all DHS sectors, future work should examine ways to tailor this broad methodology to reflect nuances of CIKR in different CIKR sectors.

OWNER/OPERATOR ROLE

Future work may be necessary to modify this methodology for scenarios where owners/ operators implement security measures intended to deter, but bear the costs themselves.

INITIAL TARGET AND ACTOR DATA ASSUMPTIONS

We have assumed specific CI target characteristics, defender pre-deterrence investments in vulnerability and consequence reduction, attacker investment in capability enhancement, and other factors. Future work should modify this data before drawing conclusions from this methodology to quantify deterrence. Also, future work could change the complexity of various risk equation components as desired; for example, the consequence equations could be expanded to include factors such as loss of public trust.

UNCERTAINTY ANALYSIS

There is inherent uncertainty and subjectivity in probabilistic risk analysis. Since our methodology uses probabilistic capability, vulnerability, and consequence mitigation estimates, future work must address the uncertainty associated with these estimates and determine the implications for quantifiable deterrence effectiveness and resulting risk reduction. Tools such as MSRAM, which inform and are improved by our deterrence quantification methodology, have built-in subjectivity reduction measures that could be leveraged.

GAME THEORY APPROACH ASSUMPTIONS

We have used an unusual approach to a game theoretical analysis by discarding the notion of using either a pure or mixed strategy Nash Equilibrium. We also have taken the unusual approach of using attacker information imperfection biases as a proxy for information when perfect information about defender investments is not available. Our approach is suited for our assumptions that neither pure nor mixed equilibrium strategies are appropriate for our present focus on quantifying the deterrence effects of defender investments. Nonetheless, future work should leverage more traditional game theoretical approaches to quantify deterrence. In support of such future work, we have already hypothesized how defender expected utility functions should be formulated for a more traditional game theoretical approach. The defender expected utility at equilibrium would be converted to loss, and then combined with attacker intent to attack (presumably with a pure NE that intent is 100%) to show the new defender unconditional risk.

Also, we assumed a two-target game; future work should examine how deterrence can be quantified in a game with more than two CI targets. This may require computer programs. Also, there are many repositories of CIKR data and risk information, such as MSRAM, ACAMS, and other databases to assist with this endeavor.

ADVERSARY TARGET PREFERENCES

Our example assumes the attacker has no preference for a certain target type; they simply want to maximize their overall expected utility from attacks. Future research can apply this methodology for issue-specific groups such as eco-terrorists who only want to target specific target types. This can support decision making on where to allocate defensive resources amongst similar target types in a defined geographic area.

OTHER RISK METHODOLOGIES

Our example has leveraged notional data from one risk analysis tool, the Coast Guard's MSRAM. Future work may use risk data from different tools and techniques to get the CIKR vulnerability and consequence data, and attacker capability data our approaches requires.

NETWORK ANALYSIS AND CIKR INTERDEPENDENCIES

Future work will examine how to quantify the deterrence effects of investments to secure networks of infrastructure from attack, including cyber networks. Network analysis uses different mathematical techniques to calculate network risk, but our methodology could still be applied. Future research could also consider the deterrent effects of investments in CIKR when those CIKR are co-located in or close proximity to other CIKR.

UTILITY FUNCTION COMPOSITION ASSUMPTIONS

We assumed attacker utilities and defender expected losses (risk) did not subtract their respective expenditures from the CI target values; the utility functions were only capability, vulnerability, and consequence modeled as functions of those expenditures. However, one could easily claim that attacker utility and defender risk must account for the respective lost expenditures, and subtract them from the expected benefit or add them onto the defender loss if an attack on a CI succeeds. Future work could create utility functions that subtract expenditures from target value. Also, we assumed the attacker's expected utility as a function of capability, vulnerability, and consequence, was the same as the defender's expected loss, thus creating a "symmetry" between attacker and defender functions. Introducing costs and budgets to the utility functions for each actor could create an asymmetrical game.

INFORMATION AVAILABILITY CIRCUMSTANCE ASSUMPTIONS

We have assumed that CI information was known to the attacker prior to defender deterrence efforts; future analysis could assume that the defender wants to quantify the deterrent effect of additional defensive investments even when their original CI investments and resulting security were not clear to the attacker. We have also assumed that attacker information is perfectly known to the defender; in reality this is likely not true and should be examined.

In situations where perfect information is determined to be more advantageous, the CI operator must work actively with law enforcement and other government entities to aggressively communicate security and consequence mitigation deterrence measure information, in great detail to potential attackers. Obviously, this will be a challenging and imperfect endeavor, but the effort is still warranted if the methodology shows it is to our advantage to make information known to prospective attackers. The intelligence community could contribute to this effort as they glean information on what the adversary knows and believes about our deterrence efforts, but much work must be done by CI operators and law enforcement authorities, as well as public figures that communicate effectively and reach a broad audience.

However, if the effort required to develop and execute a public affairs deterrence campaign exceeds the expected loss to the government if deterrence fails, perhaps the campaign is not worthwhile, and the quantitative result supporting a perfect information advantage may be discarded.

In situations where imperfect information is determined to be more advantageous, the CI operator and law enforcement, as well as public figures, must collaborate to ensure security measures are not obvious to any observer. This could entail conduct "red cell" exercises where someone simulates a terrorist planning cell and visit a CI to see what security they can identify, and/or researches open source documentation on the security and economic value of CI. Also, the government may even have to classify security measures, such as federal grant funding.

UTILITY THEORY ASSUMPTIONS

We assumed subjective expected utility in our case study. In our initial work to expand this methodology leveraging prospect utility theory, we have used a linear scale for the PU utility and probability modifiers, but the work of Daniel Kahneman and Amos Tversky suggests a nonlinear scale.[45] Future work should leverage a nonlinear scale, and can exercise that scale by assuming different deterrence investment amounts in the deterrence game and determining which modifier applies. Kahneman and Tversky have done other work on decision analysis that could be applied to the utility functions. Also, Lisa Carlson and Raymond Dacey suggest that the PU probability weighting functions may determine the difference between behavior of attackers under SEU assumptions and under PU assumptions;[46] thus the extent to which deterrence might be more quantifiably effective under one assumption as compared to the other could be analyzed by explicitly observing the effects of different probability weights.

PROBABILITY-INVESTMENT RELATIONSHIP ASSUMPTIONS

Future work should also change the slope of the exponential probability-investment relationship curves that support capability, vulnerability, and economic consequence estimates. We assumed elimination fractions of 5% vulnerability and 5% capability to attack in our initial work to create deterrence portfolios leveraging exponential probability-investment relationship assumptions; these fractions could be modified to yield different curves and possibly yield different deterrence portfolio results. Also, we have assumed a linear relationship between target economic value retention and defender consequence mitigation investment in this paper's example. In reality, this relationship may be nonlinear. We only assumed that CI target vulnerability and attacker capability probability-investment relationships could be nonlinear on theoretical grounds; future work should analyze relationships between probability reduction and investments to develop rough cost curves. Finally, we have assumed attacker capability is reduced by defender offensive counterterrorism investment, but the attacker does not counteract this effect by increasing their investments. Future work should model the simultaneous effects of attacker and defender investment upon attacker capability.

ABOUT THE AUTHORS

Eric F. Taquechel *is a lieutenant commander in the United States Coast Guard and currently serves as chief of Contingency Planning and Force Readiness, Sector Boston. He has ten years of active duty commissioned Coast Guard service and previously served in USCGC GALLATIN, Marine Safety Office and Sector Houston-Galveston, and the Domestic Port Security Evaluation Division at USCG Headquarters. He most recently authored "Layered Defense: Modeling Terrorist Transfer Threat Networks and Optimizing Network Risk Reduction," in* IEEE Network Magazine*. LCDR Taquechel earned a master's degree in security studies from the Naval Postgraduate School and prior to that earned his undergraduate degree at the US Coast Guard Academy. LCDR Taquechel may be contacted at eric.taquechel@uscg.mil.*

Ted G. Lewis *is a professor of computer science and executive director of the Center for Homeland Defense and Security at the Naval Postgraduate School. He has forty years experience in academic, industrial, and advisory capacities, ranging from academic appointments at the University of Missouri-Rolla, University of Louisiana, and Oregon State University, to senior vice president of Eastman Kodak Company, to CEO and president of DaimlerChrysler Research and Technology, North America. Dr. Lewis has published over thirty books and 100 research papers. He is the author of* Critical Infrastructure Protection in Homeland Security: Defending a Networked Nation *(2006) and, most recently,* Network Science: Theory and Applications *(2009). He received his PhD in computer science from Washington State University. Dr. Lewis may be contacted at tlewis@nps.edu.*

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the contributions of Dr. Richard Adler.

The views expressed herein are those of the authors and do not reflect the views of the United States Coast Guard or the Department of Homeland Security.

[1] U.S. Department of Homeland Security, *National Infrastructure Protection Plan* (2009), http://www.dhs.gov/xlibrary/assets/NIPP_Plan.pdf

[2] http://www.dhs.gov/xabout/laws/gc_1214597989952.shtm#1

[3] Andrew R. Morral and Brian A. Jackson, "Understanding the Role of Deterrence in Counterterrorism Security," RAND Occasional Paper (2009), http://www.rand.org/pubs/occasional_papers/2009/RAND_OP281.pdf

[4] See http://teamcore.usc.edu/projects/coastguard/

[5] Information about this initiative is available from the authors.

[6] http://birenheide.com/sra/2011AM/program/singlesession.php3?sessid=M4-H.

[7] In our example, we will show that deterrence measurement by itself may not be conclusive; it must be incorporated into a risk equation to have value and show changes in risk.

[8] DHS agencies currently are reporting risk reduction effectiveness of various CIKR security initiatives. For example, the Coast Guard reports port, waterways, and coastal security (PWCS) risk reduction efforts to Congress annually. Further, if DHS funds security investments for CIKR o/o, then they will necessarily be "joined at the hip" with those o/o to monitor implementation. Thus, DHS would be able to explain how the CIKR o/o's are deterring and thus reducing risk. For scenarios where DHS does not provide funding, there is no guarantee they could effectively report to Congress because they would not be coordinating with their CIKR stakeholders.

[9] Deterrence is the study of influencing an intelligent actor's decision making; this does not apply to unintentional threats where there is no human element or deliberative process.

[10] http://www.dhs.gov/xabout/laws/gc_1215444247124.shtm.

[11] Morral and Jackson, "Understanding the Role of Deterrence."

[12] Some DHS risk analysis tools, such as the US Coast Guard's Maritime Security Risk Analysis Model (MSRAM), have the ability to monetize human lives using a technique known as Value of Statistical Life (VSL), and draw equivalencies between different types of consequence using what is referred to as a Consequence Equivalency Matrix.

[13] The Coast Guard's MSRAM model uses a strategic threat value, without considering specific target attributes.

[14] Opportunistic attackers still face the challenge of overcoming security. If deterrence is influencing the decision making of an adversary, an opportunistic attacker could still be deterred by what they observe as they commence their attack, but that amounts to more of a tactical deterrence or last minute change in the attack planning cycle. With respect to those same security measures, more deliberate attackers would be deterred but earlier in their planning cycle, and have more time to adapt.

[15] Louis A. Cox, "Some Limitations of 'Risk=Threat x Vulnerability x Consequence' for Risk Analysis of Terrorist Attacks," *Risk Analysis 28* (2008): 1749-1761.

[16] Recoverability and resiliency measures may be too expensive for certain CIKR owners/operators to implement, but the risk equation still includes a placeholder for a consequence mitigation factor. In application, this mitigation factor may be 0, meaning expected consequence =100% of maximum potential consequence. Also, federal grants could help offset the costs of resilience measures.

[17] For a detailed explanation of MSRAM, see Brady C. Downs, "The Maritime Security Risk Analysis Model," *CG Proceedings 64* (2007): 36-38.

[18] DHS, *National Infrastructure Protection Plan.*

[19] Nikhil S. Dighe, Jun Zhuang, and Vicki M. Bier, "Secrecy in Defensive Allocations as a Strategy for Achieving more Cost Effective Deterrence," *International Journal of Performability Engineering* 5 (2009): 31- 43.

[20] Ibid.

[21]Morral and Jackson, "Understanding the Role of Deterrence."

[22] U.S. Department of Homeland Security, *DHS Risk Lexicon* (2010), http://www.dhs.gov/xlibrary/assets/dhs-risk-lexicon-2010.pdf.

[23] See, for example: Stanley Kaplan and John B. Garrick, "On the Quantitative Definition of Risk," Risk Analysis 1 (1981): 11-27; Tim Bedford and Roger Cooke, Probabilistic Risk Analysis: Foundations and Methods (Cambridge: Cambridge University Press, 2001); and Elisabeth Paté-Cornell and Seth Guikema, "Probabilistic Modeling of Terrorist Threats: A Systems Analysis Approach to Setting Priorities Among Countermeasures," *Military Operations Research* 7 (2002): 5-20.

[24] For examples of probabilistic risk tools being used in DHS, see Table 2-1 on pages 25-26 of the National Academy of Sciences Review of the Department of Homeland Security's Approach to Risk Analysis", 2010, available at http://www.nap.edu/openbook.php?record_id=12972&page=R1.

[25] Richard N. Lebow and Janet G. Stein, "Rational Deterrence Theory: I Think, Therefore I Deter," *World Politics* 41 (1989): 208–224.

[26] Elaine M. Bunn, "Can Deterrence Be Tailored?" *Strategic Forum*, No. 225 (Institute for National Strategic Studies, January 2007), http://www.hsdl.org/?view&did=481759.

[27] William M. McGill and Bilal M. Ayyub, "The Meaning of Vulnerability in the Context of Critical Infrastructure Protection," in *Critical Infrastructure Protection: Elements of Risk* (GMU School of Law Critical Infrastructure Protection Program, December 2007): 25-48, http://www.steelcityre.com/documents/RiskMonograph_1207.pdf.

[28] Patrick M. Cronin and Audrey K. Cronin, "Challenging Deterrence: Strategic Stability in the Twenty-first Century," *Special Joint Report of the International Institute for Strategic Studies and the University of Oxford Character of War Programme* (2007), http://ccw.modhist.ox.ac.uk/events/archives/mt06_deterrence/deterrence_report_mt2006.pdf.

[29] Richard N. Lebow, "The Cuban Missile Crisis: Reading the Lessons Correctly," *Political Science Quarterly* 98 (1983): 431–458.

[30] U.S. Department of Defense, *Deterrence Operations Joint Operating Concept*, version 2.0 (Washington, DC: Department of Defense, 2006), www.dtic.mil/futurejointwarfare/concepts/do_joc_v20.doc.

[31] John A. Major, "Advanced Techniques for Modeling Terrorism Risk," *Journal of Risk Finance* 4 (2002): 15-24.

[32] Bunn, "Can Deterrence by Tailored?".

[33] Erik Jenelius, Jonas Westin, and Åke J. Holmgren, "Critical Infrastructure Protection Under Imperfect Attacker Perception," *International Journal of Critical Infrastructure Protection* 3 (2010): 16-26.

[34] Martin J. Osborne and Ariel Rubenstein, *A Course in Game Theory* (Cambridge: MIT Press, 1994).

[35] Todd Sandler and Daniel G. Arce, "Terrorism & Game Theory," *Simulating and Gaming* 34 (2003): 319–337.

[36] Roger B. Myerson, *Force and Restraint in Strategic Deterrence: a Game Theorist's Perspective* (Strategic Studies Institute, US Army War College, 2007). http://www.strategicstudiesinstitute.army.mil/pubs/display.cfm?pubID=823.

[37] Morral and Jackson, "Understanding the Role of Deterrence."

[38] We use 95% elimination instead of 100% elimination to accommodate the theory that probability-investment relationships are exponential and thus probability of loss cannot be 100% eliminated, although in our case study examples our utility functions assume a linear relationship.

[39] Jenelius et. al., "Critical Infrastructure Protection Under Imperfect Attacker Perception."

[40] See http://www.fema.gov/pdf/government/grant/2011/fy11_psgp_kit.pdf for more information about Port Security Grants.

[41] The US Coast Guard's MSRAM tool includes maritime CIKR in all eighteen DHS CIKR sectors.

[42] Computing power would likely also be necessary for quantifying deterrence effects of investments, when multiple attack modes for the targets under consideration are modeled. Our example only assumes one attack mode (boat bomb).

[43] Port Security Grants could be used to reduce consequences as well, to add deterrent value. Overseas military operations and counterterrorism actions are examples of actions that mitigate attacker capabilities.

[44] Across all attacker options means the attacker uses the same attack mode (boat bomb), but could attack different combinations of CIKR. Since we don't use the traditional Nash Equilibrium to find the optimal attacker choice after we invest to deter, we need to aggregate resulting risk across all attacker choices.

[45] Daniel Kahneman and Amos Tversky, "The Framing of Decisions and the Psychology of Choice," *Science* 211 (1981): 453-458.

[46] Lisa J. Carlson and Ramond Dacey, "Sequential Analysis of Deterrence Games with a Declining Status Quo," *Conflict Management and Peace Science* 23 (2006): 181-198.