Risk-based guard allocation for threatened individuals

Nachtegael, Mirthe C.E.

Award date:
2024

Department of Industrial Engineering & Innovation Sciences
Operations, Planning, Accounting & Control

Risk-based guard allocation for threatened individuals

Master thesis Operations Management & Logistics

Author:
M.C.E. (Mirthe) Nachtegael

First supervisor:
L.P.J. (Loe) Schlicher

Second supervisor:
M. (Marco) Slikker

Third supervisor:
B. (Bart) Smeulders

Organization & organizational supervisor:
Dutch police
R. (Rob) Zandbergen

Eindhoven, 2024-02-21

Abstract

This study investigates the potential of quantitative models to enhance guard allocation to threatened individuals in the Dutch Bewaken & Beveiligen system. The system is currently under significant pressure due to the growing threat from organized crime. Three types of quantitative models were employed. First, risk factors for activities planned by threatened individuals were identified and used to quantify their level of risk exposure using discrete choice modeling. Secondly, two optimization models were used to allocate guards based on risk: one representing a semi-flexible allocation method and the other a flexible one. Finally, an attacker-defender game was developed to anticipate strategic behavior in guard allocation. Numerical analyses demonstrated the performance of the proposed allocation methods. Despite relying on several assumptions that could be relaxed in future studies, each proposed method showed notable improvements over the current allocation method.

Keywords: Resource allocation, Discrete choice modeling, Convex optimization, Attacker-defender games

Executive summary

Introduction

Crime and terrorism present a constant threat to individuals, particularly public figures. In the Netherlands, the Bewaken & Beveiligen (B&B) system protects threatened individuals. However, the B&B system is facing significant pressure due to the increasing threat from organized crime. Consequently, research efforts have emerged to strengthen the system structurally and ensure its future viability.

Problem statement

Currently, the use of quantitative models for guard allocation is limited. The decision on how many guards to assign to a threatened individual is based solely on a qualitative assessment. This means that the exact risk levels associated with the activities planned by the threatened individuals are not calculated or considered when determining the number of guards required. This results in inefficient use of available guard capacity and a suboptimal level of protection for all threatened individuals.

Research questions

The overarching research question (RQ) is:

RQ. How can the B&B system's guard allocation to threatened individuals inside the Netherlands be improved?

To answer this, the following subquestions (SQs) are addressed:

SQ1. How can the risk exposure of threatened individuals during the day be quantified?

SQ2. How can guard allocation be improved, assuming the number of guards assigned to each threatened individual does not vary during shifts?

SQ3. How can guard allocation be improved, assuming the number of guards assigned to each threatened individual can vary during shifts?

SQ4. How can guard allocation to threatened individuals be improved by anticipating strategic behavior?

Quantifying risk exposure

SQ1 aimed to quantify the risk exposure of threatened individuals throughout the day. Risk exposure is defined as the probability of an attack, which varies based on the situation and is influenced by the attacker's choices.

To model attacker decision-making, discrete choice modeling was used. A conditional logit model was formulated where the choice set for an attacker represents the daily agenda of the threatened individual. The choice set included the planned activities and the option to refrain from attacking. Activities were categorized as visits to locations or trips between locations. Risk factors for each activity type were identified based on an internal study and incorporated into the model. Risk factors were binary and included at the level of activity type.

A stated choice experiment was conducted to collect data from experts, as data from actual attackers was not available. An efficient design with 20 choice situations was generated and used to collect the data. Using this data, the conditional logit model was estimated.

The risk exposure of an activity was quantified using the conditional logit model. This was achieved by calculating the probability of the activity being selected by an attacker while taking into account the estimated coefficients for the risk factors.

Semi-flexible risk-based allocation

To answer SQ2, a semi-flexible risk-based guard allocation model was developed. The objective was to improve guard allocation while assuming the number of guards assigned to a threatened individual could vary between shifts but not during shifts.

A conceptual model was presented in which guards were assigned to threatened individuals based on the risk scores of their planned activities and their threat type. The model was mathematically formulated as a nonlinear integer program that minimized the total expected damage on a day subject to capacity constraints. The expected damage from an attack on a threatened individual was determined using the Fine-Kinney method and an exponential success function dependent on the threat level of the individual.

To solve the semi-flexible model, 1000 random situations were generated, involving three threatened individuals. The semi-flexible model was compared to a base model representing the current allocation method. This was done for a scenario where the threatened individuals to be protected all had the same threat level and a scenario where they all had different threat levels.

Flexible risk-based allocation

To answer SQ3, the semi-flexible risk-based model of SQ2 was extended. This resulted in a flexible risk-based allocation method that allowed for variation in the number of guards assigned to a threatened individual not only between shifts but also during shifts.

A conceptual model was presented in which guards could be reallocated between threatened individuals during shifts based on their planned activities' risk levels. The model was mathematically formulated as a nonlinear integer program. The objective of this model was again to minimize the total expected damage to all threatened individuals by optimally allocating the available guards.

The flexible risk-based model was solved and compared to the base model and SQ2's semi-flexible model using the 1000 randomly generated situations from SQ2. This was done again for two scenarios: one with identical threat levels for the individuals being protected, and another with different threat levels.

Attacker-defender game

SQ4 investigated how guard allocation could be improved by anticipating both defensive and offensive strategic behavior in the allocation of guards to threatened individuals. To answer the SQ, a single-period, simultaneous-move, two-player, zero-sum attacker-defender game was formulated.

The conceptual model described how the defender could assign guards to threatened individuals, and the attacker could select one of those individuals to attack. The mathematical model formalized this using a payoff matrix that represented the opposing payoffs for each combination of strategies.

It was explained how to identify one or more Nash equilibria (NEs) in the proposed game by finding the strategies where no player benefits from unilateral deviation.

To solve the game, the generated situations from SQ2 and SQ3 were used. The model was compared to a base game representing the current allocation method. This was also done for

the two scenarios: one with identical threat levels and one with different threat levels.

Results

Model	Type of threat levels	Decrease in expected damage	Improved instances
Semi-flexible model	Identical	19.6%	63.5%
	Different	21.6%	66.5%
Flexible model	Identical	46.3%	99.2%
	Different	45.5%	98.9%
Attacker-defender game	Identical	36.0%	100%
	Different	34.1%	100%

Table 1: Performance of different models.

Conclusions, limitations, and recommendations

The study concludes by summarizing the answers to the subquestions and overall research question, limitations, recommendations, and scientific contributions.

A mathematical model was developed to quantify the risk exposure of threatened individuals throughout the day, focusing on the probability of an attack for planned activities. Recommendations included revising the list of risk factors, considering attributes related to attackers and threatened individuals, and incorporating preference heterogeneity.

Two allocation methods, a semi-flexible and a flexible, were proposed to optimize guard allocation based on the total expected damage. Recommendations involved reconsidering the objective and its calculation, evaluating success functions, addressing variable travel times, and prioritizing the development of the flexible model.

A game-theoretic model was developed to anticipate strategic behavior in guard allocation, showing Nash equilibria. Recommendations include exploring more complex games, reconsidering payoffs and threat level distinctions, and studying the effects of varying guard numbers.

Overall, the study aimed to improve guard allocation in the B&B system, offering quantitative methods that show potential to enhance current practices. Recommendations include establishing a clear allocation objective through stakeholder collaboration, deciding on the method for determining attacker intent, and utilizing actual data for analysis.

The scientific contributions of this study lie in a novel application in predicting attacks on threatened individuals, by introducing discrete choice modeling to quantify risk exposure, a novel application in predicting attacks on threatened individuals; proposing nonlinear integer programming models coupled with the Fine-Kinney method, providing a unique approach to optimize guard allocation; and developing a game-theoretic model, specifically an attacker-defender game, with a unique application using the Fine-Kinney method for payoff determination.

Preface

This thesis completes my Master's program in Operations Management and Logistics at the Eindhoven University of Technology. I found the thesis project very rewarding as I was able to contribute to the improvement of the Dutch Bewaken & Beveiligen system. Although the topic was unusual for my Master's field, I am grateful for the opportunity. I believe that contributing to the safety and well-being of others is one of the noblest things in life. Therefore, this project has intrinsically motivated me.

I would like to thank Loe Schlicher for being my first assessor and mentor during this project. Loe provided detailed feedback, useful insights, and a sympathetic ear for challenges and problems. Loe mentored me not only during the project but also throughout the rest of my Master's program, helping me to achieve my academic goals. As an underdog with a Bachelor's degree in Industrial Design, I am grateful that Loe saw my potential. Second, I would like to thank Marco Slikker for his extensive feedback, which helped me improve the project. I would also like to thank Rob Zandbergen for facilitating this project and giving me the freedom to define my own research project. Last but not least, I would like to thank my friends and family, especially my parents, for supporting me throughout my time as a student. The journey has not always been easy, but the supportive people around me have helped me to stay positive.

Now, not only my Master's program but also my time as a student at the Eindhoven University of Technology is coming to an end. I am extremely satisfied and proud of my academic achievements over the past years, as well as the enjoyable experiences I have had outside of my studies. Although it is difficult to see my student days come to an end, I am glad that this project is coming to an end. I look forward to my next challenge.

Mirthe Nachtegael

Contents

List of Figures

List of Tables

1 Introduction

1.1 Research motivation

In today's world, crime and terrorism pose a constant threat to individuals, especially public figures. For instance, by the end of November 2022, Dutch politicians had reported 1072 threats, a significant increase from the 588 reports in 2020 and the previous record of 620 in 2018 (Jonker, 2022). However, public figures beyond politicians also face threats. In 2022, Dutch journalists reported 198 incidents, including threats, physical violence, stalking, and intimidation (PersVeilig, 2023). Although the number of incidents was lower than the previous year, it is presumed that these numbers only represent the tip of the iceberg. Lawyers are another group of public figures who are threatened. According to a study by I&O Research, half of the lawyers surveyed reported experiencing at least one incident of aggression in the preceding twelve months (van Miltenburg, van Straaten, & Bouwmeester, 2022). Of those, a quarter experienced multiple incidents during that period. These examples highlight the importance of implementing effective and efficient security measures to safeguard these threatened individuals.

1.2 Organization background

In the Netherlands, the protection of threatened individuals is arranged in the Bewaken en Beveiligen (B&B) system. The purpose of this system is to prevent attacks on persons, objects, and services (Ministerie van Justitie en Veiligheid, 2023). The general principle of the system is that individuals are responsible for their own safety. In addition, employers, companies, and institutions must take measures when a threat arises from work tasks. When the nature and magnitude of the threat are such that the individual and the employer can no longer offer resistance, the government takes additional measures.

In principle, taking security measures to protect threatened individuals is the responsibility of the local competent authority (Ministerie van Justitie en Veiligheid, 2023). In the B&B system, this type of security is referred to as the decentralized domain. In the decentralized domain, personal security is carried out by the Dienst Koninklijke en Diplomatieke Beveiliging (DKDB) of the Dutch police. In addition to the decentralized domain, there is also a national domain, that includes the protection of a limited group of individuals whose safe and undisturbed functioning is of national interest (Ministerie van Justitie en Veiligheid, 2023). The DKDB and the Brigade Speciale Beveiligingsopdrachten (BSB) of the Koninklijke Marechaussee (KMar) are responsible for the implementation of personal security measures in the national domain. While the DKDB is responsible for the interior of the Netherlands, the BSB is dedicated to performing its tasks in high-risk environments abroad.

Since the B&B system is under great pressure, a multidisciplinary partnership was founded to structurally strengthen the system and keep it future-proof. This partnership includes a core team of officials from the Dutch police, the Openbaar Ministerie (OM), the NCTV, and the KMar and is called the Kenniscentrum Bewaken & Beveiligen (Kenniscentrum B&B) (Ministerie van Justitie en Veiligheid, 2023). By combining its partners' knowledge, skills, and expertise, the Kenniscentrum investigates how the B&B system can deal more effectively and efficiently with the increasing threat from organized crime. The Kenniscentrum B&B works on quality improvement, knowledge assurance, and innovation.

1.3 Problem definition

As described in the previous section, the Kenniscentrum B&B is committed to improving the B&B system in terms of effectiveness and efficiency. One of the main potential points of improvement concerns the allocation of guards to threatened individuals inside the Netherlands.

1.3.1 Problem statement

Protecting a group of individuals who are under threat can be a challenging task. In the Netherlands, individuals with the highest threat level receive protection from guards. These individuals have busy agendas that can change at any moment, making short-term planning a significant challenge. The availability constraints of the guards add to the complexity of this challenge. Therefore, it is essential to efficiently utilize the available guard capacity.

Currently, the use of quantitative models for guard allocation is limited. The decision on how many guards to assign to a threatened individual is based solely on a qualitative assessment. This means that the exact risk levels associated with the activities planned by the threatened individuals are not calculated or taken into account when determining the number of guards required. This results in inefficient use of available guard capacity and a suboptimal level of protection for all threatened individuals.

1.3.2 Research questions

Following the objective of the Kenniscentrum B&B to improve the B&B system in terms of effectiveness and efficiency, the guiding research question (RQ) in this study is consequently formulated as:

RQ. How can the B&B system's guard allocation to threatened individuals inside the Netherlands be improved?

In this RQ, "improving" refers to employing quantitative methods to find allocation methods that perform better than the current way of working. The research directions required to answer this RQ are threefold. The first direction addresses the quantification of the risk exposure of threatened individuals, resulting in the first subquestion (SQ):

SQ1. How can the risk exposure of threatened individuals during the day be quantified?

The second research direction focuses on a risk-based allocation of guards, both semi-flexibly and flexibly, resulting in SQ2 and SQ3:

SQ2. How can guard allocation be improved, assuming the number of guards assigned to each threatened individual does not vary during shifts?

SQ3. How can guard allocation be improved, assuming the number of guards assigned to each threatened individual can vary during shifts?

Unlike SQ2 and SQ3, which focus solely on defensive decisions, SQ4 also analyzes the offensive decisions that the threat poses, resulting in SQ4:

SQ4. How can guard allocation to threatened individuals be improved by anticipating strategic behavior?

The output of SQ1 serves as input for the other SQs. Furthermore, SQ2 proposes a semi-flexible risk-based allocation method that permits variation in the number of guards assigned to each threatened individual, but not during shifts. SQ3 proposes a flexible risk-based allocation method that extends SQ2 by allowing the number of guards to vary not only between but also during shifts. SQ4 is similar to SQ2 in that it allows variation in the number of guards between shifts, but not during shifts. However, SQ2 only incorporates defensive strategic behavior, while SQ4 also considers strategic offensive behavior.

1.4 Outline

This report describes how the B&B system's allocation of guards to threatened individuals within the Netherlands can be improved. To accomplish this, this report first provides a literature review of background information and current practices of relevant literature in Section 2. Then, Section 3 describes how to quantify the risk exposure of threatened individuals during the day, thus answering SQ1. Section 4 and Section 5 then answer SQ2 and SQ3, respectively, and describe how the guard allocation can be improved by using a risk-based approach. Section 4 examines a semi-flexible risk-based allocation method where the number of guards is allowed to vary between, but not during, shifts. Section 5 examines a flexible risk-based allocation method where the number of guards is allowed to vary not only between shifts but also during shifts. Finally, Section 6 explores how strategic behavior can be anticipated in guard allocation to threatened individuals, answering SQ4. This report concludes with answers to the SQs and RQ, along with limitations and recommendations in Section 7.

1.5 Confidentiality

The allocation of guards in the B&B system is a highly confidential matter. Therefore, this report treats sensitive information with utmost confidentiality. Information is only shared if it is necessary for the comprehensiveness of this report. This also applies to confidential results. Furthermore, some of the information presented in this report has been intentionally altered and does not accurately reflect reality. These modifications aim to maintain confidentiality and ensure the secure handling of sensitive information. The report does not specify where the information has been altered.

2 Literature review

The literature review in this section is aimed at providing an overview of relevant literature and the development of research within the fields of quantifying risk exposure of threatened individuals and capacity allocation applicable to the protection of threatened individuals. The topic most applicable to quantifying risk exposure is discrete choice modeling for crime location choice. The topic most applicable to guard allocation to threatened individuals is attacker-defender game theory. Both these topics are therefore discussed through general information, relevant modeling methodologies, and previous work from the perspective of this study.

Before providing an overview of the existing literature on discrete choice models and attacker-defender games, an understanding of the related background is needed. Therefore, the following section elaborates on the concepts of threat and risk.

2.1 Background

Threat According to the Oxford Learner's Dictionary of Academic English (n.d.-b) the word "threat" has three definitions:

1. *"a statement in which you tell somebody that you will punish or harm them, especially if they do not do what you want"*
2. *"the possibility of trouble, danger or disaster"*
3. *"a person or thing that is likely to cause trouble, danger, etc."*

Note from these definitions that a threat can be either concrete or abstract. A concrete threat is specific and identifiable, as seen in definitions 1 and 3. In contrast, an abstract threat is more general and intangible, as seen in definition 2. For this study, threatened individuals are considered. Being threatened here does not necessarily imply that the individual has received an actual threat (i.e., definition 1), but rather that there exists a potential source of harm (i.e., definition 3) and a possibility of being harmed by that source (i.e., definition 2). Therefore, in this study, a threat refers to a possibility or source of harm.

Although it is difficult to measure threats, their strength can be evaluated based on the degree of potential harm they pose. This is often referred to as the threat level.

Risk According to the Oxford Learner's Dictionary of Academic English (n.d.-a), a definition of the word "risk" includes the following:

"the possibility of something bad happening at some time in the future; a situation that could be dangerous or have a bad result"

According to this definition, risk could be interpreted as the probability of an unwanted event occurring. In this study, risk refers to the chance of harm, where chance is the mathematical probability of harm occurring.

Threat & Risk Threat and risk are related concepts, but they have distinct meanings. In this study, a threat is defined as a possibility or source of harm, while being threatened means there is a possibility of being harmed by others. Risk, on the other hand, refers to the probability of harm occurring as a result of the present threat(s). In other words, a threat represents a possibility or source of harm, while its associated risk measures the probability of it occurring. Additionally, different threats pose varying levels of risk.

2.2 Discrete choice models

To understand how attackers make decisions about where and when to attack, discrete choice modeling for crime location choice can be used. This modeling approach offers a framework to analyze the decision-making process of individuals involved in criminal activities, providing insights into influential factors. For an overview of related work, see Appendix A.

2.2.1 Basic concepts

Discrete choice models describe, explain, and predict decision-maker choices between two or more discrete alternatives. The discrete choice model defines four elements of a choice situation (Bernasco & Ruiter, 2014):

- The **decision-maker** is the person who chooses from a set of alternatives.
- The set of **alternatives** includes a finite number of available alternatives from which the decision-maker chooses. These alternatives are mutually exclusive and collectively exhaustive.
- Alternatives have **attributes** that affect the utility to the decision-maker when being chosen. The decision-maker also has attributes that may affect the utility when selecting an alternative.
- The **decision rule** describes the process of the decision-maker to choose an alternative. The decision-maker chooses the alternative from the set that maximizes the (expected) utility.

Most of the assumptions in discrete choice modeling are based on the random utility maximization (RUM) theory (Bernasco & Ruiter, 2014). RUM theory is the micro-economic theory of behavior that allows for fluctuations in the utility to the decision-maker by adding a random component to the utility function. This random component represents information unknown to the analyst, not the decision-maker. By making assumptions about the distribution of this random component, probabilistic statements can be formulated and tested using a statistical model.

A discrete choice model typically considers a set of decision-makers $N \subseteq \mathbb{N}$ and a choice set $J_n \subseteq \mathbb{N}$ containing alternatives (i.e., the set of alternatives) faced by decision-maker $n \in N$. Decision-makers obtain a certain level of utility from each alternative, e.g., decision-maker n experiences a level of utility $U_{n,j} \in \mathbb{R}$ when choosing alternative $j \in J_n$. The decision-maker's utility is assumed to be incompletely observable by the researcher. Therefore, the utility $U_{n,j}$ includes two components: an observed term $V_{n,j} \in \mathbb{R}$ and an unobserved term $\varepsilon_{n,j} \in \mathbb{R}$, such that $U_{n,j} = V_{n,j} + \varepsilon_{n,j}$.

The probability $P_{n,j} \in [0, 1]$ that decision-maker n chooses alternative j from the choice set J_n is the probability that the utility associated with choosing this alternative is greater than the utility associated with any other alternative in the choice set, i.e., $P_{n,j} = Pr(U_{n,j} > U_{n,i} \forall i \in J_n \backslash \{j\})$. By substituting $U_{n,j} = V_{n,j} + \varepsilon_{n,j}$ and rewriting the equation, this becomes $P_{n,j} = Pr(\varepsilon_{n,i} - \varepsilon_{n,j} < V_{n,j} - V_{n,i} \forall i \in J_n \backslash \{j\})$, which is the most general formulation of the discrete choice model. The unobserved term $\varepsilon_{n,j}$ can be of different distributions, and therefore, $P_{n,j}$ can be determined differently.

While $V_{n,j}$ is assumed not to be directly observable, it can be observed through the chosen alternative j, the set of alternatives J_n, and the attributes $x_{n,i} \in \mathbb{R}^L$ of each alternative $i \in J_n$ as faced by decision-maker n (where $L \subseteq \mathbb{N}$ represents the set of attributes). From these observations, the researcher can specify the observed utility using the so-called representative or systematic function $V : \mathbb{R}^L \to \mathbb{R}$. The representative function takes the vectors $x_{n,j}$ and

β_n as input, where $\beta_n \in \mathbb{R}^L$ is the vector of coefficients to be estimated for decision maker n. Although the function V can take different forms, the most common form is linear.

2.2.2 Related models

This subsection discusses the discrete choice models used in crime location choice research. The conditional logit model is the simplest model used in most crime location choice research. Other discrete choice models used in crime location choice studies include the mixed logit model and the latent class logit model. These three models are discussed in the following sections.

Conditional logit The conditional logit model, also known as the multinomial logit model, is a type of discrete choice model. It assumes that decision-makers perceive the attributes of alternatives differently. The coefficient estimates for the observed attributes are identical across all decision-makers (i.e., $\beta_n = \beta \forall n \in N$). Furthermore, this model assumes that the unobserved utility term $\varepsilon_{n,j}$ follows an extreme value type I distribution (i.e., $\varepsilon_{n,j} \sim EV(0,1)$).[1] Assuming $\varepsilon_{n,i} = \varepsilon_{n,j} + V_{n,j} - V_{n,i}$, the cumulative distribution gives $P_{n,j}|\varepsilon_{n,j} = \prod_{i \in J_n \setminus \{j\}} e^{-e^{-(\varepsilon_{n,j}+V_{n,j}-V_{n,i})}}$. Because $\varepsilon_{n,j}$ is unknown, $P_{n,j} = \int (\prod_{i \in J_n \setminus \{j\}} e^{-e^{-(\varepsilon_{n,j}+V_{n,j}-V_{n,i})}}) e^{-\varepsilon_{nj}} e^{-e^{-\varepsilon_{n,j}}} d\varepsilon_{n,j}$. Algebraic manipulation of this integral results in the closed-form expression $P_{n,j} = \frac{e^{V_{n,j}(\beta)}}{\sum_{i \in J_n} e^{V_{n,i}(\beta)}}$.[2] Conditional logit models are typically estimated using maximum likelihood estimation (Frith, 2019).

Although the conditional logit model has been widely adopted in crime location selection, it is not without limitations (Frith, 2019). First, it assumes that all decision-makers share the same preferences (i.e., β). However, it may be that decision-makers substantially vary in their decision-making. This decision-maker heterogeneity can significantly impact the model's results. Second, the conditional logit model assumes independence of irrelevant alternatives (IIA), which may be too restrictive. IIA implies that the relative probability of choosing any alternative over another is independent of any other alternatives. While this seems reasonable, it ignores the potential similarity and, thus, substitutability of the alternatives.

Mixed logit The mixed logit model is a more sophisticated conditional logit model that allows for variations in preferences among the decision-maker population. This model accommodates unobserved preference heterogeneity by allowing variation in the population coefficient estimates β, such that $\beta_n \sim f(\beta|\theta)$, where θ is the distribution of β_n's over the population. Since β_n is random and unknown, the probability that decision-maker n selects alternative j is calculated as $P_{n,j} = \int L_{n,j}(\beta) f(\beta|\theta) d\beta$, where $L_{n,j}(\beta) = \frac{e^{V_{n,j}(\beta)}}{\sum_{i \in J_n} e^{V_{n,i}(\beta)}}$.[3] Mixed logit models are typically estimated through simulation with either maximum simulated likelihood or hierarchical Bayes (Frith, 2019).

Although the mixed logit model has the advantage of accommodating unobserved preference heterogeneity, meaning that it does not assume IIA, it also has its limitations (Frith, 2019). One limitation is that the mixed logit model requires the specification of the preference distribution $f(\beta|\theta)$, which can be challenging. Another limitation is that, unlike the conditional logit model, the mixed logit model has no closed form and cannot be solved analytically. As a result, it must be estimated through simulation, which is computationally expensive.

[1] The probability density function and cumulative function for $\varepsilon_{n,j} \sim EV(0,1)$ are $f(\varepsilon_{n,j}) = e^{-\varepsilon_{n,j}} e^{e^{-\varepsilon_{n,j}}}$ and $F(\varepsilon_{n,j}) = e^{e^{-\varepsilon_{n,j}}}$, respectively.

[2] If $V_{n,j}$ is assumed to be linear in β, i.e., $V_{n,j}(\beta) = \beta x_{n,j}$, then $P_{n,j} = \frac{e^{\beta x_{n,j}}}{\sum_{i \in J_n} e^{\beta x_{n,i}}}$.

[3] If $V_{n,j}$ is assumed to be linear in β, i.e., $V_{n,j}(\beta) = \beta x_{n,j}$, then $P_{n,j} = \int \frac{e^{\beta x_{n,j}}}{\sum_{i \in J_n} e^{\beta x_{n,i}}} f(\beta|\theta) d\beta$.

Latent class logit The latent class logit is a special case of the mixed logit model. Similar to the mixed logit model, this model assumes unobserved preference heterogeneity. The difference, however, is that the latent class logit model assumes that each decision-maker can be divided into a class $c \in C$ with a certain probability. The coefficient estimates of all decision-makers within a class c are identical and can be denoted as $b_c \in \mathbb{R}^L$. The probability that decision-maker n belongs to class c equals $s_c \in [0, 1]$, which is dependent on the observed decision-maker characteristics $d_n \in \mathbb{R}^M$ (where $M \subseteq \mathbb{N}$ represents the set of attributes). d_n is consistent between decision-makers within each class c. Therefore, the probability that decision-maker n chooses alternative j is $P_{n,j} = \sum_{c \in C} s_c \frac{e^{V_{n,j}(b_c)}}{\sum_{i \in J_n} e^{V_{n,j}(b_c)}}$.[4] Here s_c can be estimated within the model along with b_c for each class. Latent class models are typically estimated through maximum likelihood estimation or expectation-maximization (Frith, 2019). However, because the latter is significantly faster than the first, it is often preferred.

The latent class logit model, like the conditional and mixed logit models, has limitations (Frith, 2019). Firstly, the model requires a specification of the number of distinct classes. This is typically done by repeatedly estimating models with varying amounts of classes and comparing information criteria values. This process can be time-consuming. Secondly, the latent class logit model is computationally expensive to estimate, similar to the mixed logit model, due to the large number of parameters.

2.2.3 Maximum likelihood estimation

Most discrete choice models are estimated by maximizing a function, such as the likelihood function, the simulated likelihood function, or squared moment conditions (Train, 2009). This section described the numerical procedure to maximize a likelihood function.

The goal of maximum likelihood estimation is to determine the β that maximizes the likelihood function. Typically, the log-likelihood function is used for this purpose because it is more manageable and produces equivalent results. The log-likelihood function takes the form $LL(\beta) = \sum_{n \in N} \ln(P_n(\beta))$, where $P_n(\beta) \in [0, 1]^{J_n}$ is the probability vector of the observed outcomes for decision maker n. Note that $LL(\beta)$ always returns a negative value since the likelihood is a probability between 0 and 1, and the natural logarithm of any number between 0 and 1 is always negative.

The numerical optimization then works as follows (Train, 2009). First, β_0 is specified. The β attained after t iterations from β_0 is denoted by β_t. Each iteration updates β such that the value of $LL(\beta)$ is higher than at the previous iteration, i.e., $LL(\beta_{t+1}) > LL(\beta_t)$. Updating β is repeated until convergence, i.e., until the difference between the old and new values is significantly small. The difficulty here is determining the best value for β at the next iteration β_{t+1}, therefore, maximization algorithms are used. The most prominent maximization algorithms developed over the years are Newton–Raphson and Berndt–Hall–Hall–Hausman (BHHH).

[4] If $V_{n,j}$ is assumed to be linear in b_c, i.e., $V_{n,j}(b_c) = b_c x_{n,j}$, then $P_{n,j} = \sum_{c \in C} s_c \frac{e^{b_c x_{n,j}}}{\sum_{i \in J_n} e^{b_c x_{n,i}}}$.

2.3 Attacker-defender games

Attacker-defender games can be used to study the strategic interaction between defenders and attackers. They are derived using game theory, a widely deployed mathematical modeling tool proven to enable important operational insights. In attacker-defender games for resource allocation, the defender and attacker seek to optimally allocate defensive resources among a set of targets and attack one or more of these targets, respectively. For an overview of related work, see Appendix A.

Problem settings Hunt and Zhuang (2023) conducted a systematic review of journal articles on attacker-defender games. Their literature review presents an overview of the problem settings covered in the attacker-defender game literature, including infrastructure and asset protection games. These games involve the allocation of defensive resources by a defender among multiple targets and the target selection decisions made by one or more attackers after observing the defender's allocation. Infrastructure/asset protection games are considered for this study because individuals who are considered threatened can be viewed as assets to be protected.

2.3.1 Modeling approaches

Hunt and Zhuang (2023) conducted a detailed analysis of the game formulations proposed in the literature. The authors discussed the differences in these formulations concerning the sequence of moves, number of players, players' decision variables, players' objective functions, and time horizon. Additionally, Guan and Zhuang (2016) described the variation in contest success functions among different formulations.

Sequence of moves The sequence of moves describes how an attacker-defender game is played and affects the solving method employed. A game can be either a sequential-move or a simultaneous-move game (Hunt & Zhuang, 2023). In a sequential-move game, one player moves first, and the next player moves second after observing the strategy executed by the first mover. In a two-player attacker-defender game, the first mover is the defender and the second mover is the attacker. In this case, the defender moves first by allocating resources among targets, and the attacker moves second by selecting one or more targets to attack based on observing the defender's allocation strategy. In contrast, simultaneous-move games involve both players moving concurrently without knowledge of their opponent's strategy. The defender allocates resources among the targets, and the attacker executes their attack strategy simultaneously. This means that players select their strategy without knowing what the other player has chosen. It is important to note that players in simultaneous games could move sequentially in reality.

Resource allocation to protect threatened individuals can be modeled as a sequential-move or simultaneous-move attacker-defender game, depending on the observability of the players' strategies. It is commonly assumed that the defender allocates resources to specific locations at certain times, while the attacker decides whether, where, and when to attack. However, the sequence of moves is contingent on the attacker's observability. A game is considered sequential-move when the attacker can observe the defensive strategy, and simultaneous-move when the attacker cannot.

Number of players According to Hunt and Zhuang (2023), most studies consider two-player attacker-defender games with one defender and one attacker. However, in some cases, it is more realistic to consider multiple attackers. This is especially true when the defender needs to protect multiple individuals from multiple attackers. In such cases, the attackers are assumed

to have specific preferences towards different threatened individuals, and therefore the model must consider a heterogeneous population of attackers.

Players' decision variables Hunt and Zhuang (2023) discovered that defenders commonly base their decisions on resource allocation, selecting the number of resources to distribute among multiple targets. This number can be continuous or discrete, depending on the nature of the defender's resources. Meanwhile, the attacker's most common decision variable is target selection. This selection can include a discrete or continuous choice for the attacker to select their target(s). However, another interesting decision variable is a binary attack choice where the attacker decides whether or not to attack a target.

Players' objective functions Hunt and Zhuang (2023) discovered that the most frequent objective of defenders is to maximize expected utility or minimize expected disutility. Utility functions for defenders are typically defined as the difference between expected benefits and losses. The most common attacker objective is to maximize utility. Utility functions for attackers are broadly defined.

Keeney (2007) explained various methods for developing objective functions for attackers and defenders. In certain attacker-defender games, the sum of both players' objective functions is zero, resulting in a zero-sum game. In these types of games, the attacker seeks to find a strategy that maximizes their utility while operating within the constraints of the defender's defenses. Conversely, the defender's objective is to minimize the attacker's utility by allocating defensive resources to prevent successful attacks. The utilities of the attacker and defender are directly opposed. In the attacker-defender game literature, utility is also commonly referred to as payoff or reward.

Time horizon Hunt and Zhuang (2023) observed that the most common approach is to model a single-period game, while another approach considers multi-period games. A single-period game is analyzed after one round of play, while a multi-period game involves a defender selecting a defensive strategy in each period and the attacker selecting the attack strategy after observing this defensive strategy.

Contest success function A contest success function (CSF) determines the probability of winning or losing a game as a function of defense and attack efforts. According to Guan and Zhuang (2016), CSFs in the attacker-defender game literature either have an exponential or ratio form. Exponential-form CSFs assume binary attack and continuous defense efforts and let the probability of a successful attack decrease exponentially in the defender's effort and be independent of the attack effort. Ratio-form CSFs assume continuous attack and defense efforts and let the probability of a successful attack decrease convexly in the defender's efforts (and the inherent defense level) and increase concavely in the attacker's efforts. Additionally, the exponential and ratio forms can be combined in a CSF. Table 2 summarizes the CSFs in the attacker-defender game literature. $A, D, C \in \mathbb{R}_+$ represent the attack effort, defense effort, and inherent defense level, respectively, and $k, m \in \mathbb{R}$ are effectiveness coefficients.

Form	Function	Source(s)
Exponential	e^{-kD}	Bier, Haphuriwat, Menoyo, Zimmerman, and Culpen (2008); Hao, Jin, and Zhuang (2009); Wang and Bier (2011); Shan and Zhuang (2013a)
Ratio	$\frac{A}{k(A+D+C)}$	Zhuang and Bier (2007)
	$\frac{A^m}{A^m+D^m}$	Hausken (2008)
	$\frac{A}{A+D+C}$	Hausken and Zhuang (2012)
	$\frac{k_1 A}{k_1 A+k_2 D+C}$	Guan, He, Zhuang, and Hora (2017)
	$1-e^{-kA/D}$	Nikoofal and Zhuang (2012)

Table 2: CSFs in the attacker-defender game literature. From: Guan and Zhuang (2016).

2.3.2 Common assumptions

Hunt and Zhuang (2023) elaborate on the three most common assumptions enforced in attacker-defender games. These include perfect rationality, risk neutrality, and complete information.

Perfect rationality Perfect rationality assumes that all players in a game act rationally by identifying and selecting optimal decisions. However, in practice, this assumption may not always hold. In such cases, the assumption of perfect rationality is relaxed by introducing bounded rationality. Unlike the assumption of perfect rationality, bounded rationality recognizes that a player's ability to make fully rational decisions is constrained.

Risk neutrality Risk neutrality assumes that all players in a game are neutral to risk. However, in reality, players may not always be risk-neutral; they can also be risk-seeking or risk-averse. In such cases, the assumption of risk neutrality is relaxed, and risk preference parameters need to be incorporated. The most common way to do so is by the power utility function $u(y) = y^\beta$ with $y \in \mathbb{R}_+$, where $\beta \in \mathbb{R}_+$ represents the risk preference value. Depending on the risk preference of the strategic player, the power utility function is adjusted to cover risk-averse ($0 < \beta < 1$), risk-neutral ($\beta = 1$), and risk-seeking ($\beta > 1$) behaviors.

Complete information The assumption of complete information assumes that all players know their own and other players' parameters, objectives, and decision options, as well as the sequence of moves. However, this assumption may not hold in reality, resulting in incomplete information. Incomplete information does not necessarily imply that players have no information at all, but rather that they have limited or uncertain information. Incomplete information can take different forms, depending on the player to whom the information is unknown and the type of information. Incomplete information can affect the sequence of moves in sequential-move games, which may need to be modeled as simultaneous-move games.

2.3.3 Solving methods

Solving an attacker-defender game involves identifying the optimal strategies for the players involved, typically resulting in the identification of equilibrium points. This solution provides insights into how rational players should behave in the given strategic interaction. The type of

solution and method for solving a game depend on the problem setting, modeling approaches, and (relaxed) common assumptions.

According to Hunt and Zhuang (2023), closed-form solutions are the most common in the attacker-defender game literature. A solution is considered closed-form if it can be evaluated in a finite number of generally accepted functions and mathematical operations. Due to the linear nature of their objective functions and constraints, and the compactness of their strategy spaces, games with closed-form solutions are often tractable, meaning that solutions can be obtained with relative ease. Backward induction is an example of an approach used to generate closed-form solutions for sequential-move games, while best response analysis is used for simultaneous-move games.

Alternatively, algorithmic and heuristic methods can be used to solve attacker-defender games. The need for such methods arises from the complexities of games based on real-world problems. According to Hunt and Zhuang (2023), new heuristics and algorithms are frequently proposed to solve games, in addition to traditional mathematical programming approaches. Moreover, many games are formulated as mixed integer linear programs (MILPs).

3 Quantifying risk exposure

SQ1. How can the risk exposure of threatened individuals during the day be quantified?

SQ1 aimed to develop a quantitative model capable of calculating a risk score for any planned activity on any given day for any individual at risk, based on identified risk factors. The objective was to quantify the risk exposure of threatened individuals throughout the day.

3.1 Risk identification

Risk exposure can be interpreted as the probability of an attack on a threatened individual, which varies depending on the situation and is influenced by the attacker's choices. Attackers make decisions to maximize their utility or satisfaction based on a rational assessment of available options. To identify the risk factors that attackers use to make rational assessments, it is necessary to specify the options available to them. This study assumes that the attacker has a finite number of options available per day, which include the activities planned by the threatened individual that the attacker is interested in.

Activities Attackers may strike at any time, but they are more likely to target a threatened individual during certain activities. In this study, attackers are assumed to decide when and where to attack based on a rational assessment of the threatened individual's planned activities. Additionally, attackers are assumed to make trade-offs between activities based on their attributes (i.e., risk factors).

This study considers an activity planned by a threatened individual as either a visit to a location or a trip from one location to another. Therefore, activities are classified into two categories: locations and routes. Each type of activity has a distinct set of risk factors that influence attackers' decisions.

Risk factors An attacker's decision-making process is influenced by the risk factors associated with the activities planned by the threatened individual they are interested in. To identify the risk factors for the two types of activities, namely locations and routes, experts were consulted. The experts referred to an internal confidential study conducted by the Dutch police. Using this study, nine risk factors were identified: five for locations and four for routes. Table 3 provides an overview of the identified risk factors per activity type. The risk factors are numerically labeled to keep their actual meaning confidential. All risk factors are binary variables, indicating their presence or absence. Additionally, all risk factors are assumed to have a positive impact on risk levels, meaning that the presence of a risk factor increases the level of risk.

Activity type	
Location	**Route**
x_1	x_6
x_2	x_7
x_3	x_8
x_4	x_9
x_5	

Table 3: Risk factors per activity type.

None of the identified risk factors included the number of guards assigned to a threatened individual during an activity. Therefore, the presence of guards would not affect the probability of an attack during an activity if these risk factors were used to calculate it. However, it is important to note that in reality, the presence of guards may still influence the probability of an attack. Additionally, the identified risk factors did not consider the duration of the activity. However, activities with the same risk factors may have varying attack probabilities depending

on their duration. The probability of an attack increases as the activity duration increases. The set of risk factors was not expanded to include the number of guards or activity duration to maintain consistency with the internal confidential study that was consulted.

3.2 Modeling

After identifying the risk factors, the next step was to develop a model that quantifies the impact of each risk factor on an attacker's decision to attack a threatened individual during a specific activity on a particular day. The modeling approach, including the type of model, was determined first, followed by the development of the model itself.

3.2.1 Modeling approach

Discrete choice modeling (DCM) was utilized to model the decision-making process of attackers and quantify the impact of each identified risk factor. DCM is a quantitative technique that analyzes the decision-making process of individuals when presented with a finite set of discrete alternatives. See Subsection 2.2 for a general explanation of discrete choice models. In this study, the set of alternatives available to an attacker on a given day included the activities planned by the threatened individual that the attacker is interested in. DCM was considered an appropriate approach, given that this set is finite and discrete.

3.2.2 Model development

Discrete choice models come in many forms. In this case, attacker preferences were assumed to be the same across all attackers, and therefore, a conditional logit model was developed. The developed conditional model follows a specific structure. The choice set of an attacker includes the activities planned by the threatened individual they are interested in on a given day and the option not to attack. The attacker chooses which activity in the choice set to attack the threatened individual, or not to attack at all. The choice set therefore includes three alternative types labeled as 'none', 'location', and 'route'. The choice set contains one and only one 'none' alternative that represents the option not to attack the threatened individual on the given day and has no attributes. This option may, for instance, represent the choice to attack the threatened individual on another day or not to attack. Furthermore, the choice set contains one or more alternatives of the remaining two types. These alternatives include the activities planned by the threatened individual on the given day. The 'location' alternatives represent the options to attack the threatened individual at a particular location. The 'route' alternatives are options for attacking the threatened individual while traveling from one location to another. The unique attributes of these alternatives include the identified risk factors.

The conditional logit model was formulated using the following notations.

A	Set of activity types[5]
$J \in \mathbb{N}$	Set of activities planned by the threatened individual
$J_a \subseteq J$	Set of activities of type $a \in A$ planned by the threatened individual
$P_{none} \in [0,1]$	The probability that the threatened individual is not attacked
$P_j \in [0,1]$	The probability that the threatened individual is attacked at activity $j \in J$
$V_{none} \in \mathbb{R}$	The observed utility for the attacker when not attacking
$V_j \in \mathbb{R}$	The observed utility for the attacker when attacking at activity $j \in J$

[5] Activities can be of the type location or route, therefore, the set of activity types is $A = \{location, route\}$.

Choice probabilities The probability of selecting each alternative in a given choice set can be calculated based on the attacker's observed utility. To determine the probability of choosing an alternative, divide the exponential observed utility of the considered alternative by the sum of the exponential observed utilities of all alternatives in the choice set. The choice set includes one and only one 'none' alternative representing the option not to attack. The other alternatives in the choice set represent the planned activities. These alternatives are of the type 'location' or 'route'. To model an attacker's decision-making behavior towards a specific individual on a particular day, the following conditional logit was used.

$$P_{none} = \frac{e^{V_{none}}}{e^{V_{none}} + \sum_{a \in A} \sum_{j \in J_a} e^{V_j}} \tag{1}$$

$$P_{j'} = \frac{e^{V_{j'}}}{e^{V_{none}} + \sum_{a \in A} \sum_{j \in J_a} e^{V_j}} \forall j' \in J \tag{2}$$

Observed utility functions A distinct linear utility function was used for each alternative type, namely 'none', 'location', or 'route'. Only a linear form was considered due to the binary nature of the risk factors. It was assumed that the option not to attack results in no observed utility to the attacker. Therefore, the utility function of the 'none' alternative is $V_{none} = 0$. To represent activities, the observed utility to the attacker was assumed to depend on an alternative-specific constant and the identified risk factors. The linear utility function for activity $j \in J_a$ with type $a \in A$ was defined as $V_j = \alpha_a + \sum_{n \in N} \beta_n x_{n,j}$. Here, $\alpha_a \in \mathbb{R}$ is the alternative-specific constant to be estimated for activities of type a. This constant is used to account for unobserved risk factors that may affect the choice probabilities of the alternatives but are not explicitly included in the model. $\beta_n \in \mathbb{R}_+$ is the coefficient to be estimated for risk factor $n \in N$, where $N \in \mathbb{N}$ is the set of identified risk factors from Table 3. $N_a \subseteq N$ is the set of risk factors identified for activities of type a, such that $|N| = \sum_{a \in A} |N_a|$. $x_{n,j} \in \{0, 1\}$ is the value of risk factor $n \in N$ for activity $j \in J_a$ with type $a \in A$, where $x_{n',j} = 0$ for all $n' \in N \backslash N_a$. Bringing this altogether, Equation 1 and Equation 2 were further rewritten to:

$$P_{none} = \frac{1}{1 + \sum_{a \in A} \sum_{j \in J_a} e^{\alpha_a + \sum_{n \in N} \beta_n x_{n,j}}} \tag{3}$$

$$P_{j'} = \frac{e^{\alpha_{a'} + \sum_{n \in N} \beta_n x_{n,j'}}}{1 + \sum_{a \in A} \sum_{j \in J_a} e^{\alpha_a + \sum_{n \in N} \beta_n x_{n,j}}} \forall j' \in J_{a'}, a' \in A \tag{4}$$

3.3 Stated choice experiment

After formulating the model, the next step was to collect data on the choice preferences of attackers to estimate the alternative-specific constant α_a for each activity type $a \in A$ and the coefficient β_n for each $n \in N$. Unfortunately, such data was not readily available, and collecting data from actual attackers was not an option. Therefore, a stated choice experiment was set up to collect data from experts. For a general explanation of stated choice experiments, steps to create a stated choice experiment, and design types, see Appendix B. According to ChoiceMetrics (2021), creating a stated choice experiment involves three primary steps: model specification, experimental design generation, and questionnaire construction.

3.3.1 Model specification

The first step is to specify the model, which was already done in Subsection 3.2. A conditional logit model was formulated with three alternative types: 'none', 'location', and 'route'. The 'none' alternative has no attributes, and the attributes for locations and routes include the identified risk factors. It was decided to measure only the main effects and not the interaction effects of the risk factors. Based on the information provided in the internal confidential study by the Dutch police, it was reasonable to assume that the identified risk factors acted independently of each other. Therefore, interaction effects were not expected to play a significant role. Furthermore, the inclusion of interaction effects increases complexity and the number of coeffi- cients to be estimated. Due to time constraints, it was deemed more practical and efficient to focus on estimating the main effects of the risk factors. Finally, nonlinear effects were not considered since all risk factors are binary variables. The use of alternative forms for the utility function would not provide added value.

3.3.2 Experimental design generation

Experimental design Generating the design was the second step. The following decisions were made to generate the experimental design.

- Three sets of unlabeled alternatives were decided to be included in each choice situation (i.e., a situation where a respondent must select one option among the set of alternatives). These sets included the three alternative types: 'none', 'location', and 'route'. One 'none' alternative and two alternatives of each activity type (i.e. 'location' or 'route') were included to ensure that effects could be measured between and within the activity types. All alternatives were unlabeled because the risk factors for the alternatives representing activities are generic for the respective activity type and the 'none' alternative has no attributes.
- The number of risk factor levels was set to two for each risk factor because the risk factors are binary variables. Similarly, the risk factor levels included 0 and 1 for each risk factor.
- An efficient design was chosen for the stated choice experiment. This design aims to be statistically efficient in terms of the predicted standard errors of the coefficient estimates. In other words, it attempts to maximize the information obtained from each choice situation. For further information on efficient designs and other design types in stated choice experiments, refer to Appendix B. In this case, a full factorial design that tests all possible combinations was deemed inappropriate due to the large number of choice situations required. Therefore, a fractional factorial design that selects choice situations from the full factorial design was considered. Among the design types falling under this category, the efficient design was deemed the most suitable. This is because an efficient design is always superior to an orthogonal design when coefficient information is available. Efficient designs use prior coefficient knowledge to optimize the design by gaining the most information from each choice situation. Although limited information was available on the prior coefficients for the risk factors, the signs of the coefficients were known. Specifically, it was known that the coefficients were non-negative because all risk factors were known to increase the probability of an attack.
- It was decided not to maintain attribute-level balance in the design. Attribute-level balance requires that each risk factor level (i.e., 0 or 1) appears an equal number of times in the design for every risk factor. Ensuring balance at the risk factor level helps to effectively estimate coefficients across a range of levels, thereby avoiding data points at only a few selected risk factor levels. Although attribute-level balance is often considered a desired property, ignoring this property typically produces more efficient designs (ChoiceMetrics,

2021). To generate a more efficient design, it was decided not the balance the risk factor levels.

- A total of 20 choice situations were decided to be generated for the experimental design. The number of choice situations is typically constrained from below by the number of coefficients to be estimated (including constants) and the number of choice situations needed to maintain balance at the attribute level. However, in this case, attribute-level balance was not maintained. Therefore, the number of choice situations was constrained by the number of coefficients to be estimated, which was 11, given the use of 9 risk factors and 2 constants in the conditional logit model. To ensure adequate variation in the data, the number of choice situations in the design was set higher than the number of coefficients to be estimated. It was unnecessary to generate a larger design. According to Rose and Bliemer (2013), smaller optimal designs are more effective in retrieving information from best-choice tasks compared to larger optimal designs, which tend to contain inferior-choice tasks.

Choice situations The construction of choice situations for the efficient design was done using the Ngene software. Ngene is a software distributed by ChoiceMetrics for generating experimental designs for stated choice experiments to estimate choice models, particularly of the logit type. Although most decisions were already made during the generation of the experimental design, a few remained for the construction of the choice situations using Ngene. These included the following.

- To reduce the task effort per expert, choice situations were generated and divided into two blocks of 10. A block is a subset of an experimental design presented to respondents. Experimental designs are divided into blocks based on the minimum correlation principle. This principle aims to minimize the correlation between attributes (in this case, risk factors) and blocks, ensuring diversity in the combinations of choice situations presented to respondents.

- The D-error measure was chosen to determine an efficient design. This measure assumes a single respondent and is derived from the determinant of the asymptotic variance-covariance (AVC) matrix. For more information on the D-error and AVC matrix, see Appendix B.

- The Modified Federov algorithm was selected to generate the design. See Appendix B, for more information on algorithms for generating efficient designs. The Modified Federov algorithm was chosen because it allows flexible constraints to be imposed on a design, which helps to avoid dominance among alternatives. It was decided to allow the algorithm to run for 30 minutes without further improvements.

- Constraints were used during design generation to prevent dominance in the generated choice situations. An alternative is considered dominant if it is preferred over another alternative. Experimental designs without dominant alternatives provide valuable information by forcing experts to make clear trade-offs between risk factors. In this case, the constraints ensured that the Modified Federov algorithm did not include choice situations with dominated alternatives in the efficient design.

- The prior coefficients in the efficient were non-informative priors. Non-informative priors are based on knowledge of the sign of the coefficients. For more information on setting priors, see Appendix B. In this case, the sign of the risk factors was known to be non-negative. Therefore, the non-informative prior coefficients in the efficient design were set to a positive, near-zero value of 0.001.

The syntax used to generate the 20 choice situations divided into two blocks using the

efficient design is available in . The resulting choice situations can also be found there. The D-error of the efficient design was 0.518. also contains the syntax for an orthogonal design and its resulting choice situations. The D-error of this orthogonal design was 0.625, which demonstrates that, as expected, the efficient design is better than the orthogonal design (as $0.518 < 0.625$).

3.3.3 Questionnaire construction

The final step in creating the stated choice experiment was to develop two questionnaires (i.e., one for each block of the design) for experts to fill out. Translating the experimental design into questionnaires in a way that makes sense to the experts is essential. Therefore, first, a pilot questionnaire was developed, tested, and reviewed.

Pilot questionnaire The pilot questionnaire was developed in Microsoft Forms. Its content was presented in Dutch to ensure the experts understood all the included information. First, a short introduction and a clear explanation of what to expect was provided. This included a table detailing brief descriptions and examples of each risk factor. Next, the 10 choice situations of the first block of the experimental design generated by the Ngene software were presented, each representing the daily plan of a threatened individual with four activities. Each choice situation consisted of five alternatives. These alternatives included four activities (i.e., two locations and two routes), and the option of no attack. For each activity, the presence of each risk factor was indicated in a table. An example of such a table is presented in Figure 1. Note that in this example, the headers are in English, and the brief descriptions of the risk factors are replaced by x_1 through x_9 (refer to Table 3) due to confidentiality reasons. In the pilot questionnaire, Dutch headers and brief descriptions of the risk factors were used.

Location 1	Location 2	Route 1	Route 2
x1	x1	x6	
	x2		x7
	x3	x8	
	x4	x9	
x5			

Figure 1: Example of a tabular choice situation presentation.

Each choice situation was presented on a separate page with two questions. The goal of the first question was to identify the activity most likely to be attacked. Only one of the four activities could be selected. The choice of not attacking was presented in the second question for each choice situation. The goal of this question was to indicate whether or not an attack was expected to occur more or less likely. A separate question was used to avoid experts selecting the no-attack alternative solely because they found it challenging to choose from the available activities. It would have become impossible to estimate the model if the option of no attack was chosen too often, as this would have resulted in insufficient data for the other alternatives.

Final questionnaires The pilot questionnaire was distributed to four employees of the Dutch police who provided feedback. Their overall opinion was positive. The main comment was related to the choice of words. They suggested replacing some words with specialized jargon to minimize misunderstandings among experts. The pilot questionnaire was adapted to this feedback, resulting in the following questions and response options (translated into English). The original Dutch questions used in the questionnaire can be found in Appendix D.

1. *In which of the following activities is the threatened individual most likely to be attacked? Assume no security guards are present.*
   (a) *Location 1*
   (b) *Location 2*
   (c) *Route 1*
   (d) *Route 2*
2. *Is it more likely that an attack will or will not occur in the above situation?*
   (a) *It is more likely that an attack **will** occur.*
   (b) *It is more likely that an attack **will not** occur.*

After evaluating the pilot questionnaire, the two final questionnaires (one for each block) were developed in Microsoft Forms.

Distribution The next step was to distribute the two questionnaires to the experts. Experts could complete one or both questionnaires as long as the total number of responses for the questionnaires remained approximately the same. However, they were not allowed to complete any of the questionnaires more than once. Before distributing the questionnaires, the minimum sample size was determined.

Several researchers have proposed rules of thumb for estimating sample size requirements in stated choice experiments. The most frequently cited rule was suggested by Orme (1998), who proposed the following equation to estimate the minimum sample size for stated choice experiments that involve only main effect estimation.

$$N \geq 500 \cdot \frac{L^{\max}}{J \cdot S} \tag{5}$$

Here N is the minimum sample size, $L^{\max}$ is the largest number of levels for any of the attributes, J is the number of alternatives considered, and S is the number of choice tasks each respondent faces. This method of determining the sample size needed has, however, received some criticism. It for example does not relate the experimental design to the sample size calculation, only some dimensions of the design (Rose & Bliemer, 2013). However, using other methods was not possible in this case because more specific calculations can only be made when informative coefficient priors are available.

Thus, despite the criticism, Orme's rule of thumb was used to calculate the minimum sample size. In the case of this study, $L^{\max} = 2$, because all identified risk factors are binary variables. Furthermore, five alternatives and a total of 20 choice situations were used in the stated choice experiment, therefore $J = 5$ and $S = 20$. Substituting these values into Equation 5 results in $N = 10$, and thus a minimum sample size of 10. However, since two blocks were used in the experimental design, the actual minimum sample size in this case was 20 ($= 2 \cdot 10$). Consequently, the goal was to distribute the questionnaires to as many experts as possible and subsequently collect at least 10 responses for each block.

Responses Thirteen responses were collected for the first block, and ten for the second block. Although the difference in the number of responses per block may introduce some form of bias, such bias is washed out because the sample sizes (i.e., the number of responses for the blocks) are large enough and the discrepancy in sample sizes is not very large.

Appendix D includes a figure with pie charts displaying the selected choices for each choice situation. The figure shows that there is a great variability in the choices made by the experts.

3.4 Model estimation

The following step involved estimating the conditional logit model using the data. This entailed estimating the alternative-specific constant α_a for each activity type $a \in A$ and the coefficients β_n for each $n \in N$. The Biogeme package in Python was utilized for this purpose. Biogeme is an open-source Python package that is specifically designed for maximum likelihood estimation of parametric models, with a particular focus on discrete choice models. Subsubsection 2.2.3 provides a general explanation of maximum likelihood estimation (MLE) for discrete choice models. The Python code used to estimate the model was written following the guidelines outlined by Bierlaire (2023).

The estimation report and the estimated coefficients can be found in Appendix E. The estimated values for α_a for each $a \in A$ and β_n for each $n \in N$ range from -2.500 for α_{route} to 81.595 for β_1. The alternative-specific constant and the estimated coefficients for route alternatives are much smaller than those for location alternatives. This indicates that, according to the experts who completed the questionnaires, threatened individuals are generally more likely to be attacked at a location than at a route. The highest coefficient estimates for location alternatives are β_3 and β_4, meaning that x_3 and x_4 are the most influential risk factors for locations. The highest coefficient estimates for route alternatives are β_6 and β_7, meaning that x_6 and x_7 are the most influential risk factors for routes. Most coefficients show a p-value of 0, indicating statistical significance. The p-values for α_{route} and β_9, however, are both 1, indicating that the estimates for these coefficients are not considered reliable or meaningful based on the data. In other words, there is not enough evidence to conclude that the observed effects are not due to variability in the data.

3.4.1 Model evaluation

The final stage involved evaluating the developed conditional logit model. The estimation report was used to assess McFadden's R-squared, which is a widely used metric for measuring the explanatory power of a conditional logit model. McFadden's R-squared compares the log-likelihood of the fitted model to that of the null model (a model where all coefficients to be estimated are set to zero). It is calculated as follows.

$$R^2_{McFadden} = 1 - \frac{LL_{fitted}}{LL_{null}} \tag{6}$$

In this equation, $R^2_{McFadden} \in \mathbb{R}$ is the McFadden's R-squared. Furthermore, $LL_{fitted} \in \mathbb{R}$ is the log-likelihood for the model, and $LL_{null} \in \mathbb{R}$ is the log-likelihood for the null model. A higher log-likelihood value indicates a better fit. In practice, log-likelihood values are usually negative, and as a result, $R^2_{McFadden}$ typically ranges from 0 to 1, with a higher value indicating a better fit. Values between 0.2 and 0.4 indicate a good fit, while values above 0.4 indicate an excellent fit (McFadden, 1979).

From the estimation report in Appendix E, it can be seen that the $R^2_{McFadden} = 1.36$ for the conditional logit model, which is atypical. Although the LL_{fitted} of 132, compared to the LL_{null} of -370, indicates a significant improvement in model fit, the $R^2_{McFadden}$ cannot be interpreted because its value is greater than 1. One possible reason for the atypical value could be the variability in the data, which can be observed from the pie charts in Appendix D.

3.5 Conclusion

Risk exposure can be interpreted as the probability of an attack on a threatened individual during an activity that the threatened individual has planned on a given day. This probability can be determined by analyzing attacker decision-making, which is influenced by the presence or absence of risk factors during these activities. In this study, five risk factors were identified for visits to locations, while four were identified for trips between locations. A conditional logit model was developed to estimate the impact of each risk factor on the probability of an attack. The model was trained using data collected through a stated choice experiment. It was shown how to estimate and evaluate the conditional logit model. Due to the variability in the collected data, some estimated coefficients did not show significance.

4 Semi-flexible risk-based allocation

SQ2. How can guard allocation be improved, assuming the number of guards assigned to each threatened individual does not vary during shifts?

SQ2 aimed to create a quantitative model that improves guard allocation to threatened individuals. A requirement was that the number of guards assigned to each threatened individual could vary between but not during shifts.

Base allocation The base allocation is a simplified representation of the current method of assigning guards to threatened individuals. It assumes that guards are assigned based solely on the threatened individual's threat level, which is determined by the individual's characteristics and the potential threat(s) they face. Three distinct threat levels are assumed: low, general, and high. The required level of protection is determined based on the level of threat. Each threat level corresponds to a predetermined number of guards. Low-threat individuals are assigned one guard, general-threat individuals are assigned two guards, and high-threat individuals are assigned three guards. The base allocation exists only when the number of required guards is equal to the number of available guards.

The guards are assumed to work in two shifts: an early shift and a late shift. A shift is assumed to have seven working hours to protect threatened individuals. Additionally, the shifts on a day are assumed to be consecutive. Guards are assigned to threatened individuals on a shift-to-shift basis, depending on their planned activities. If no activities are planned, then no guards are assigned to the threatened individual. If one or more activities are planned during a shift, the number of guards associated with the threat level is assigned to protect the threatened individual during that shift. This means protection may be provided during the early shift, the late shift, or both. The number of guards assigned to the threatened individuals is not affected by the specific characteristics of the planned activities.

Consider a scenario in which six guards are available, and three threatened individuals are to be protected, each with a different threat level and agenda. Figure 2 shows the distribution of guards over the threatened individual in this scenario, following the base allocation. The allocation of guards to threatened individuals is based solely on their threat level and not on the characteristics of their planned activities.

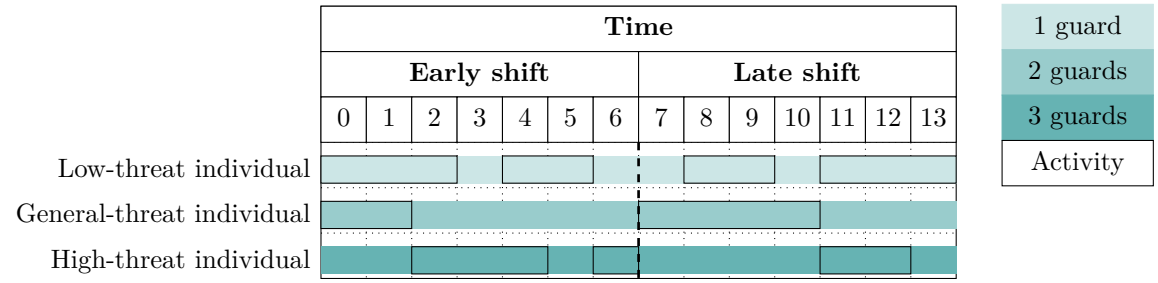


Figure 2: Example of base allocation.

Semi-flexible allocation This SQ analyzes the effect of varying the number of guards assigned to a threatened individual from shift to shift on the efficiency of guard allocation. A semi-flexible allocation method is proposed to determine the number of guards based on the threat levels of the threatened individuals and the risk levels of their planned activities, rather than using a predetermined number for all individuals at a given threat level.

Consider again the scenario from Figure 2, where six guards are to be divided over three

threatened individuals. Suppose now that the risk level of the general-threat individual's planned activity during the late shift is higher than that of the high-threat individual's planned activity during this shift. In this case, it is appropriate to deviate from the base allocation. Figure 3 shows the distribution of guards over the threatened individual in this scenario, following the proposed semi-flexible allocation. In contrast to the base allocation, the assignment of guards to threatened individuals is based on both their threat levels and the risk levels associated with their planned activities. It is important to note that this is just an example, and any deviation from the base allocation will depend on the specific scenario.

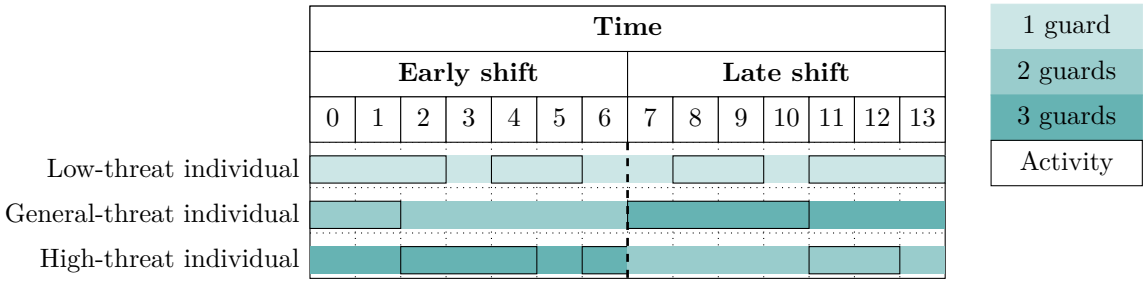


Figure 3: Example of proposed semi-flexible risk-based allocation.

4.1 Modeling

This subsection presents the conceptual and mathematical model for the semi-flexible allocation method to improve the allocation of guards to threatened individuals. The conceptual model provides an abstract representation of the key concepts, relationships, and structures within the semi-flexible model. The mathematical model offers a formal representation of the conceptual model using mathematical language and expressions to describe the semi-flexible model's behavior quantitatively, enabling a detailed and precise analysis.

4.1.1 Conceptual model

Objective To develop a model, it is essential to establish a clear and concise objective that includes the aim of the model. Input from the Dutch police, experts, academics, and stakeholders is necessary to set the appropriate objective. A complete exploration of the process of setting the objective was outside the scope of this study. Therefore, only a small group of employees from the Dutch police were consulted to formulate an objective. Based on the collected input, the model's objective was determined as follows:

Minimization of the total expected damage on a given day

The total expected damage on a given day is calculated as the sum of the expected damage over all activities planned by the threatened individuals on that day. The expected damage for a planned activity was calculated using the Fine-Kinney method (Fine, 1971; Kinney & Wiruth, 1976). This method employs a mathematical formula to calculate the risk score associated with a given threat. The risk score is calculated by multiplying the exposure, probability, and consequence. Exposure refers to the likelihood of a threat occurring, while probability refers to the likelihood of a threat resulting in an expected consequence. Consequence refers to the most likely effect of a potential threat. The risk score is a quantification of the potential threat's severity, which takes into account the probability of the threat causing a specific consequence. The Fine-Kinney risk score is a useful tool for estimating the expected damage of a planned activity. Thus, calculating the expected damage for an activity on a given day involves multiplying the following factors.

1. The probability of an attacker having the intention to attack the threatened individual.
2. The probability of an attack on the threatened individual at the considered activity, given that an attacker intends to attack the individual.
3. The probability of success of an attack on the threatened individual during the considered activity, given the assigned guards and the individual's threat level, and assuming an attacker attacks.
4. The value of the threatened individual.

Factor 1 considers the probability that a potential attacker intends to attack the threatened individual on the day in question. Factor 2 includes the conditional probability that an attack will occur during the specific activity planned by the threatened individual, given that an attacker intends to attack the individual. This probability depends on the activity's characteristics and can be calculated using the conditional logit model from SQ1. Multiplying these two factors results in the 'exposure' used in the Fine-Kinney method. Although this method assumes a range of 0 to 10, this study uses a range of 0 to 1 for factors 1 and 2 and their product.

Factor 3 represents the conditional probability of a successful attack if one were to occur. The probability of success can be determined by using a function that takes into account the number of guards assigned to the threatened individual at the time. The success function considers the threat level of the threatened individual to account for the nature of potential attacks. A high threat level is assumed to be associated with a potential attacker with better means to attack a potential attacker associated with a lower threat level. Therefore, the success function used for a threatened individual should be based on their level of threat. For individuals with the same level of threat, the same success function can be used. However, for individuals with different threat levels, a different success function is required. This factor can be interpreted as the 'probability' used in the Fine-Kinney method. Although this method assumes a range of 0 to 10, this study uses a range of 0 to 1 for factor 3.

Factor 4 concerns the value of the threatened individual to the Dutch police, which may include the potential impact of an attack on the individual as well as organizational or political interests. The organization is responsible for determining an appropriate value for a threatened individual. This factor can be interpreted as the 'consequence' used in the Fine-Kinney method. Although the Fine-Kinney method assumes a range of 0 to 100, this study uses a range of 0 to 10 for factor 4.

Success functions To calculate the probability of success for an attack based on the number of guards assigned, a success function is utilized, as previously explained. Each threatened individual's success function is assumed to depend on their threat level, meaning that the success function may differ from one threatened individual to another. The success function for threatened individual $k \in K$ can be denoted by $f_k : \mathbb{Z}_+ \to [0, 1]$, where K is the set of threatened individuals.

Inspired by the CSFs in the attacker-defender literature (see Bier et al. (2008); Hao et al. (2009); Wang and Bier (2011); Shan and Zhuang (2013a)), the success function is assumed to have an exponential form to ensure that a probability of success of 1 is achieved when no guards are assigned, and a lower probability of success when more guards are assigned. This approach also considers the varying degrees by which the probability of success decreases with each additional guard. For instance, adding the first guard results in a greater decrease in the probability of success than adding the second guard, etc. Assume that $d \in \mathbb{Z}_+$ is the number of guards assigned and $\lambda_k \in \mathbb{R}_+$ is a constant reflecting the effectiveness of defense for threatened

individual $k \in K$. Then, the success function for threatened individual $k \in K$ can be described as follows.

$$f_k(d) = e^{-\lambda_k d} \tag{7}$$

Note that other formulations for threatened individuals' success function are possible. To estimate the success function f_k (i.e., to determine λ_k) for each threatened individual $k \in K$, the average number of typically required guards $\overline{d}_k \in \mathbb{R}_+$ and a threshold τ_k were used. This threshold reflects the acceptable probability of success and includes a realistic value, not a desired one. A value of $\tau_k = 0.1$, for instance, means that out of 100 attacks on threatened individual $k \in K$, 10 are expected to be successful. The value of the threshold influences the importance of additional guards. The higher the value, the more weight is placed on the first assigned guard(s). Determining λ_k for a threatened individual $k \in K$ includes solving the following equation.

$$\begin{aligned} \tau_k &= e^{-\lambda_k \overline{d}_k} \\ \Rightarrow \lambda_k &= -\frac{\ln(\tau_k)}{\overline{d}_k} \end{aligned} \tag{8}$$

It is assumed that each threatened individual has a low, general, or high threat level, thus distinguishing three types of success functions. To estimate the success functions for these three types, data was needed. An additional section was added to the Microsoft Forms questionnaire for SQ1 to collect this data. This section presented respondents with three questions and instructed them to assume that threatened individuals either have a low, general, or high threat level. The following questions were posed. The original Dutch questions used in the questionnaire can be found in Appendix D.

1. *What is the minimal number of guards needed to fully resist an attack on a **low-threat** threatened individual?*
2. *What is the minimal number of guards needed to fully resist an attack on a **general-threat** threatened individual?*
3. *What is the minimal number of guards needed to fully resist an attack on a **high-threat** threatened individual?*

Although these questions suggest that assigning a specific number of guards can reduce the probability of success, it is important to note that there always exists a small probability of success, even if it is extremely small. Additionally, the probability of success can never be eliminated, as the negative exponential function can never reach zero. Nonetheless, the questions were deemed appropriate because a small positive near-zero value for the threshold τ_k for each $k \in K$ could be used. The threshold should be determined carefully by experts. However, for illustrative purposes, τ_k was set at 0.01 for each $k \in K$. This implies that out of 100 attacks, only one is expected to succeed. The responses to the questions for each type of threatened individual were averaged, resulting in $\overline{d}_k$. These averages were substituted in Equation 8 to estimate the success functions.

Appendix F contains Python code to estimate the success functions. To generate these plots it was assumed that $\tau_k = 0.01$ for each $k \in K$ and that $\overline{d}_k = 1$ if $k \in K$ is a low-threatened individual, $\overline{d}_k = 2$ if $k \in K$ is a general-threatened individual, and $\overline{d}_k = 3$ if $k \in K$ is a high-threatened individual. The resulting success functions were derived using Equation 7 and Equation 8. For threatened individual $k \in K$ these functions are $f_k(d) = e^{-4.605d}$, $f_k(d) = e^{-2.303d}$, and $f_k(d) = e^{-1.535d}$, for a low, general, and high threat level, respectively.

Assumptions The semi-flexible model was formulated for a daily period to minimize the total expected damage on a given day. This decision was reinforced by the fact that guard allocation planning is done per day, making it impractical to consider any other period.

Although the working hours of the shifts are not fixed in practice and may vary daily based on the planning of the threatened individuals, the semi-flexible model assumed non-overlapping early and late shifts to ensure a clear structure with logical progression. Additionally, it assumed that seven hours were reserved for protecting threatened individuals during each shift, without overlap. It is important to note that this duration is not fixed in reality.

For the sake of simplicity, the semi-flexible model assumed that an activity of a threatened individual is planned during either the early or late shift. In reality, the start time, end time, and duration of shifts may be adjusted for some guards to accommodate the planned activities of all threatened individuals in either shift. However, each activity in the semi-flexible model was simplified to take place during either the early or late shift.

Finally, it was assumed that activities would start and end on the hour. The minimum duration of an activity was set at one hour, while the maximum was set at seven hours. The latter is the assumed number of hours available for protection during each shift.

4.1.2 Mathematical model

The subsequent step involved translating the conceptual model into a mathematical model. The mathematical model for a given day was formulated using the following notations.

$K \subseteq \mathbb{N}$	Set of threatened individuals
S	Set of shifts
$J_k \subseteq \mathbb{N}$	Set of activities planned by threatened individual $k \in K$
$T \subseteq \mathbb{Z}_+$	Set of time in hours
$T_s \subset T$	Set of hours reserved for protecting threatened individuals during shift $s \in S$
$T'_s \subset T_s$	T_s for shift $s \in S$ excluding the first hour of the shift
$T_{k,j} \subseteq T_s$	Set of hours for activity $j \in J_k$ for threatened individual $k \in K$
$d_{k,j} \in \mathbb{Z}_+$	Number of guards assigned to threatened individual $k \in K$ for activity $j \in J_k$
$d'_{k,t} \in \mathbb{Z}_+$	Number of guards assigned to threatened individual $k \in K$ at time $t \in T$
$f_k : \mathbb{Z}_+ \to [0,1]$	Success function for attacks on threatened individual $k \in K$
$P_k \in [0,1]$	Probability that someone intends to attack threatened individual $k \in K$
$P_{k,j} \in [0,1]$	Probability that threatened individual $k \in K$ is attacked at activity $j \in J_k$
$v_k \in [0,10]$	Value of threatened individual $k \in K$
$B_s \in \mathbb{N}$	Number of available guards during shift $s \in S$

Recalling from the conceptual model, the set defining the shifts included $S = \{early, late\}$. Additionally, the set of time was $T = \{0, 1, \ldots, 12, 13\}$, where $T_{early} = \{0, 1, \ldots, 5, 6\}$, and $T_{late} = \{7, 8, \ldots, 12, 13\}$. The shifts do not overlap, meaning that T_{early} and T_{late} are mutually exclusive. Furthermore, since activities planned by a threatened individual cannot overlap, $T_{k,j'} \cap T_{k,j} = \emptyset$ for each $k \in K$, $j' \in J_k$, and $j \in J_k \backslash \{j'\}$. Additionally, since the attack probabilities for the activities planned by a threatened individual on a day are assumed to be determined by the conditional logit model from SQ1, $\sum_{j \in J_k} P_{k,j} = 1$ for each $k \in K$.

$$
\begin{aligned}
\min_{(d_{k,j})_{k\in K, j\in J_k}, (d'_{k,t})_{k\in K, t\in T}} \quad & \textstyle\sum_{k\in K} P_k v_k \sum_{j\in J_k} P_{k,j} f_k(d_{k,j}) & & \\
\text{s.t.} \quad & d_{k,j} = d'_{k,t} & & \forall k \in K, j \in J_k, t \in T_{k,j} \\
& d'_{k,t} = d'_{k,t-1} & & \forall k \in K, t \in T'_s, s \in S \\
& \textstyle\sum_{k\in K} d'_{k,t} \leq B_s & & \forall t \in T_s, s \in S \\
& d_{k,j} \in \mathbb{Z}_+ & & \forall k \in K, j \in J_k \\
& d'_{k,t} \in \mathbb{Z}_+ & & \forall k \in K, t \in T
\end{aligned}
\tag{9}
$$

Equation 9 represents the semi-flexible model. This model is a nonlinear integer programming (NLIP) problem. NLIP is a method for optimizing a mathematical model that includes a system of constraints and a nonlinear objective function, with variables constrained to take integer values. The objective function determines the quantity to be optimized, and the goal is to find the values of the decision variables that maximize or minimize the objective function under the given constraints.

The objective function (i.e., $\sum_{k\in K} P_k v_k \sum_{j\in J_k} P_{k,j} f_k(d_{k,j})$) represents the total expected damage on a given day, which should be minimized. This objective function is nonlinear and convex. First, the objective function is nonlinear due to the term $f_k(d_{k,j})$. This term represents an exponential function of the decision variable $d_{k,j}$ within the nested summation. Because of this, the objective function is a sum of exponential functions, making it nonlinear. Second, the objective function is convex, because the term $f_k(d_{k,j})$ is a convex function. Because a sum of convex functions is also convex, the objective function is convex.

The decision variables include the number of guards assigned to each threatened individual $k \in K$ during each activity $j \in J_k$ and at each time of the day $t \in T$, denoted by $d_{k,j}$ and $d'_{k,t}$, respectively. The constraints in Equation 9 reflect the restrictions on the decision variables. The first constraint ensures a constant number of guards assigned to a threatened individual during an activity. The second constraint ensures a constant number of guards assigned to a threatened individual during each shift. The third constraint ensures that the total number of guards assigned during each hour of a shift does not exceed the number of available guards. The fourth and fifth constraints ensure that the decision variables are non-negative integers.

The semi-flexible model in Equation 9 has not been formulated more simply on purpose. Instead of using two decision variables, one variable representing the number of guards assigned to a threatened individual during a shift could have been used. However, this simpler formulation was not employed to ensure alignment with the flexible model presented in SQ3.

4.2 Model evaluation

To solve the semi-flexible model in Equation 9, a variety of situations were randomly generated and solved. This included finding the optimal allocation and associated objective value for each generated situation. The results were compared to those of a model that represents the base allocation. The sensitivity of the semi-flexible model was also evaluated.

The interpretation of the model evaluation results should be done with care as they highly depend on the model's assumptions.

4.2.1 Solving method

Because the semi-flexible model in Equation 9 includes a nonlinear and convex optimization problem, CVXPY was used to solve it. CVXPY is a Python library that can be used for convex

optimization, including NLIP problems. This library automatically translates the problem into a standard form that various underlying solvers can solve, calls a solver, and unpacks the results.

The semi-flexible model was solved using CVXPY for a set of generated situations, with the MOSEK solver. MOSEK is a commercial optimization solver that supports a wide range of problem types, including linear programming, quadratic programming, conic optimization, and mixed-integer optimization.

4.2.2 Situations generation

To produce a diverse set of situations, 1000 random situations were generated. Although generating more situations would have been preferable, 1000 was chosen to limit computational time. To save computational time, only three threatened individuals were considered, i.e., $K = \{1, 2, 3\}$. For each situation, random values for the parameters P_k, v_k, J_k, $P_{k,j}$, $T_{k,j}$ for $k \in K$ and $j \in J_k$ were generated. This process involved executing the following steps:

1. Draw a random probability P_k for each threatened individual $k \in K$ from the uniform distribution on the interval $[0, 1]$, rounded to three decimal places.
2. Draw a random value for v_k for each threatened individual $k \in K$ from the uniform distribution on the interval $[0, 10]$, rounded to three decimal places.
3. Draw a random integer from the range $[1, 3]$ for the number of activities planned by each threatened individual $k \in K$ for each shift $s \in S$, and store those into J_k.
4. Generate a random attack probability $P_{k,j}$ for each activity $j \in J_k$ and probability of no attack $P_{k,none}$ for each threatened individual $k \in K$, summing up to 1[6], rounded to three decimal places.
5. Generate a random set of hours for each activity $j \in J_k$ planned by each threatened individual $k \in K$, denoted by $T_{k,j}$, without overlap in time between the activities in the set J_k.[7]

In Appendix G, the Python code for generating the situations is included. Additionally, a table with the descriptive statistics of the randomly generated parameters in the situations is included. These descriptives were rounded to three decimals if necessary. Appendix H contains the Python code used to solve the semi-flexible model for each generated situation.

4.2.3 Model comparison

The semi-flexible allocation model was compared to a base and an optimized base model using the generated situations. The base model assigns a predetermined number of guards to threatened individuals depending on their threat level. In the optimized base model, the number of guards assigned to each threatened individual may vary from these predetermined numbers but must remain constant throughout the day. The semi-flexible model includes all possible assignments available in the base and optimized base models, making it superior by definition. Similarly, by definition, the optimized base model is superior to the base model.

The models were compared for two scenarios. These included a scenario where the threatened individuals to be protected all have the same threat level and a scenario where they all have

[6] First, $|J_k|+1$ random numbers are drawn from a uniform distribution on the interval $[0, 1]$. These numbers are then normalized by dividing each number by the sum of the $|J_k| + 1$ generated numbers, ensuring they add up to 1. The resulting normalized numbers represent the generated probabilities.

[7] The start times for the activities in shift $s \in S$ are determined by drawing and ordering a uniform random sample from T_s without replacement with a size of the number of activities planned during the shift. The end times are determined by randomly drawing an integer between the start time of the considered activity and the minimum of the start time of the next activity (if there is one) and the end time of the shift.

different threat levels. For each of these two scenarios, the three models were solved for all generated situations, and the objective values were calculated using the objective function in Equation 9.

Identical threat levels To compare the base, optimized base, and semi-flexible models when considering threatened individuals with the same threat level, it was assumed that individuals 1, 2, and 3 were assigned general threat levels, such that $\overline{d}_1 = \overline{d}_2 = \overline{d}_3 = 2$ and $f_1(d) = f_2(d) = f_3(d) = e^{-2.303d}$, as derived from Equation 8. Furthermore, it was assumed that $B_{early} = B_{late} = \overline{d}_1 + \overline{d}_2 + \overline{d}_3 = 6$.

Figure 4 displays three box plots of objective values: one for the base model, one for the optimized base model, and one for the semi-flexible model. Table 4 shows the improvements in the mean objective value and the number of improved instances for each combination of models.

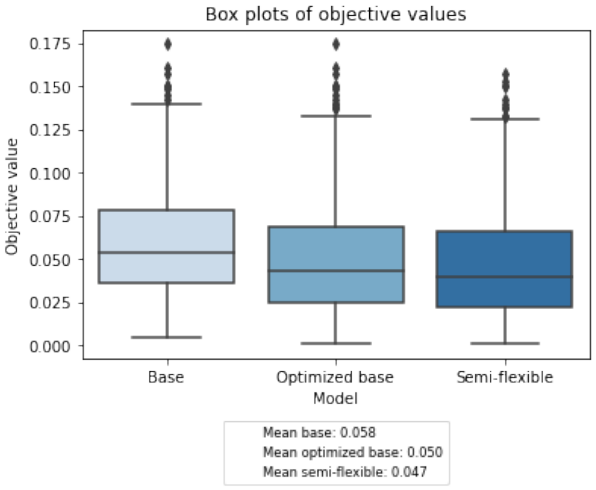


Figure 4: Box plot of objective values for the base, optimized base, and semi-flexible models, assuming identical threat levels.

Model		Improvement	
Reference	**Alternative**	**Mean objective value**	**Improved instances**
Base	Optimized base	13.8%	40%
Base	Semi-flexible	19.6%	63.5%
Optimized base	Semi-flexible	6.7%	44%

Table 4: Comparison of the base, optimized base, and semi-flexible models, assuming identical threat levels.

Different threat levels To compare the base, optimized base, and semi-flexible models when considering threatened individuals with different threat levels, it was assumed that individuals 1, 2, and 3 were assigned threat levels low, general, and high, respectively, such that $\overline{d}_1 = 1$, $\overline{d}_2 = 2$, $\overline{d}_3 = 3$, $f_1(d) = e^{-4.605d}$, $f_2(d) = e^{-2.303d}$, and $f_3(d) = e^{-1.535d}$, as derived from Equation 8. Furthermore, it was assumed that $B_{early} = B_{late} = \overline{d}_1 + \overline{d}_2 + \overline{d}_3 = 6$.

Figure 5 displays the box plots of objective values for each model, and Table 5 shows the improvements in the mean objective value and the number of improved instances for each combination of models.

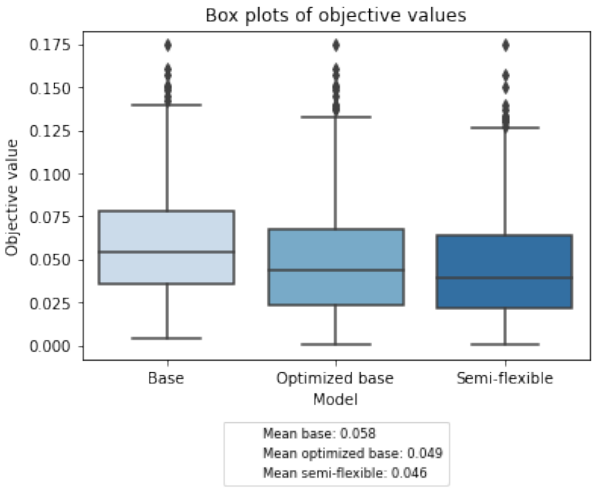


Figure 5: Box plot of objective values for the base, optimized base, and semi-flexible models, assuming different threat levels.

Model		Improvement	
Reference	**Alternative**	**Mean objective value**	**Improved instances**
Base	Optimized base	15.9%	43%
Base	Semi-flexible	21.6%	66.5%
Optimized base	Semi-flexible	6.8%	47.5%

Table 5: Comparison of the base, optimized base, and semi-flexible models, assuming different threat levels.

Scenario comparison The figures and tables for both scenarios show that the optimized base model, and even more so the semi-flexible model, outperforms the base model. Although Figure 4 and Figure 5 are very similar, Table 4 and Table 5 reveal some interesting differences. The improvement in mean objective value and number of improved instances for each combination of models is slightly greater when assuming different threat levels than when assuming identical threat levels. This means that using the optimized base model and the semi-flexible model can lead to slightly greater improvements when the threat levels of the individuals to be protected are different than when the threat levels are similar.

4.2.4 Sensitivity analysis

A sensitivity analysis was conducted to evaluate the effects on the optimal solution and its associated objective value as parameters within the semi-flexible model change. The goal of a sensitivity analysis is to explore the range in which a parameter can be altered while still maintaining the value of decision variables within the optimal solution. Additionally, it is used to investigate the impact of changes in a parameter on the objective values. To accomplish these goals, two questions were to be answered:

1. How does a change in a parameter affect the optimal solution?
2. How does a change in a parameter affect the objective value?

Method To answer these questions, it was assumed that the threatened individuals were $K = \{1, 2, 3\}$ and that each threatened individual $k \in K$ has two activities planned on a day (i.e., $|J_k| = 4$), two for each shift. The timing of these activities is not relevant because the model assumes that activities are planned during the early or late shift and therefore the times do not affect the model.

To analyze parameter sensitivity, each parameter was individually tested while holding all other parameters at a chosen base value. This allowed for the evaluation of each parameter's impact on the semi-flexible model individually, aiding in the comprehension of which parameters have the greatest impact on the optimal solution and objective value. The parameters that were tested included the product of P_k, v_k, and $\sum_{j \in J_k} P_{k,j}$ for each $k \in K$, λ_k for each $k \in K$, and B_s for each $s \in S$. When testing the parameter $P_k v_k \sum_{j \in J_k} P_{k,j}$ for each $k \in K$, the value of each $P_k v_k P_{k,j}$ for each $j \in J_k$ was set to $\frac{P_k v_k \sum_{j \in J_k} P_{k,j}}{|J_k|}$. The parameter J_k for each $k \in K$ was not tested because a change in J_k directly influences the times ($T_{k,j}$) and attack probability ($P_{k,j}$) of each activity in $j \in J_k$, making it impossible to independently asses the sensitivity for this parameter.

To assess how a change in a parameter affects the optimal solution, stability ranges were calculated, indicating the sensitivity of each parameter. The stability range refers to the range of parameter values in which the optimal solution remains unchanged when the parameter value is altered. Stability ranges are important indicators of the robustness of the optimal solution for changes in parameters. Parameters with narrow stability ranges are more sensitive, meaning that small changes in these parameters could lead to a different optimal solution. Conversely, parameters with wide stability ranges impact the optimal solution's stability less.

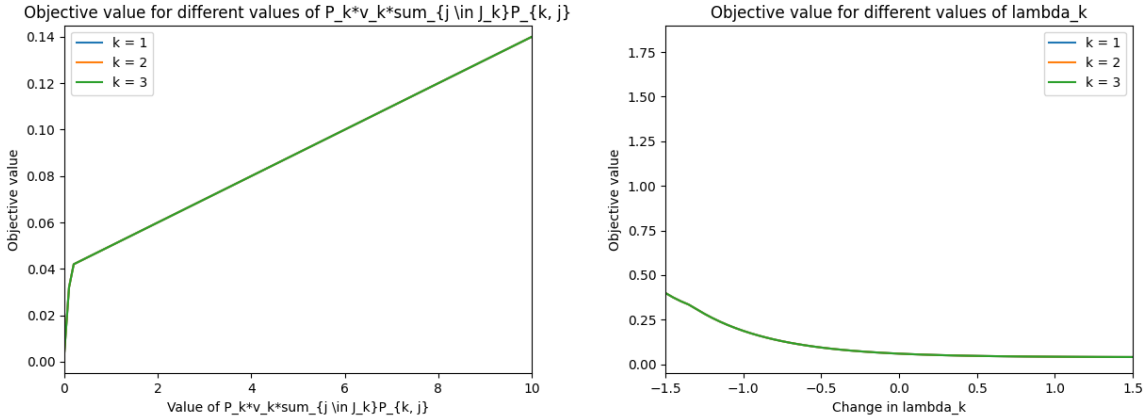
To assess how changing a parameter affects the optimal solution, each parameter was tested individually while holding all other parameters at their base values. For each combination tested, the objective value was calculated. Finally, a line plot of the objective values was created for each parameter.

The sensitivity analysis was conducted for the two scenarios used in the model comparison from Subsubsection 4.2.3: one where all threatened individuals to be protected have the same threat level and another where they have different threat levels.

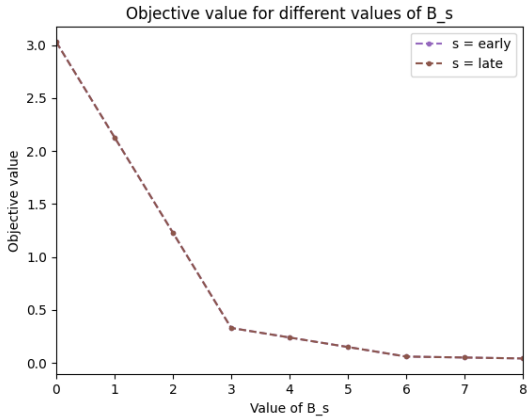
Identical threat levels For each parameter to test, a base value was determined. The base value for $P_k v_k \sum_{j \in J_k} P_{k,j}$ for each $K \in K$ was set to the product of the mean possible values of its independent terms. For P_k and v_k for each $K \in K$, these mean values included 0.5 and 5, respectively. For $\sum_{j \in J_k} P_{k,j}$, the mean possible value for each $K \in K$ included 0.8 (meaning that $P_{k,j} = P_{k,none} = 0.2$ for each $j \in J_k$ and $k \in K$). Therefore, the base value for $P_k v_k \sum_{j \in J_k} P_{k,j}$ for each $K \in K$ was set to $0.5 \cdot 5 \cdot 0.8 = 2$. It was assumed that all individuals were assigned general threat levels, such that $\overline{d}_1 = \overline{d}_2 = \overline{d}_3 = 2$ and the base values for λ_k included 2.303 for each $k \in K$, as derived from Equation 8. Finally, the base values for B_{early} and B_{late} were both set to $\overline{d}_1 + \overline{d}_2 + \overline{d}_3 = 6$.

Parameter	Possible values	Base value	Step size	Stability range
$P_1 v_1 \sum_{j \in J_1} P_{1,j}$	[0, 10)	2	0.001	(0.200, 10)
$P_2 v_2 \sum_{j \in J_2} P_{2,j}$	[0, 10)	2	0.001	(0.200, 10)
$P_3 v_3 \sum_{j \in J_3} P_{3,j}$	[0, 10)	2	0.001	(0.200, 10)
λ_1	$(0, \infty)$	2.303	0.001	(0.963, 4.702)
λ_2	$(0, \infty)$	2.303	0.001	(0.963, 4.702)
λ_3	$(0, \infty)$	2.303	0.001	(0.963, 4.702)
B_{early}	$\{1, 2, \ldots, \infty\}$	6	1	{6}
B_{late}	$\{1, 2, \ldots, \infty\}$	6	1	{6}

Table 6: Results of parameter effects on the optimal solution for the semi-flexible model, assuming identical threat levels.



(a) $P_k v_k \sum_{j \in J_k} P_{k,j}$ **(b)** λ_k



(c) B_s

Figure 6: Line plots of parameter effects on the objective value for the semi-flexible model, assuming identical threat levels.

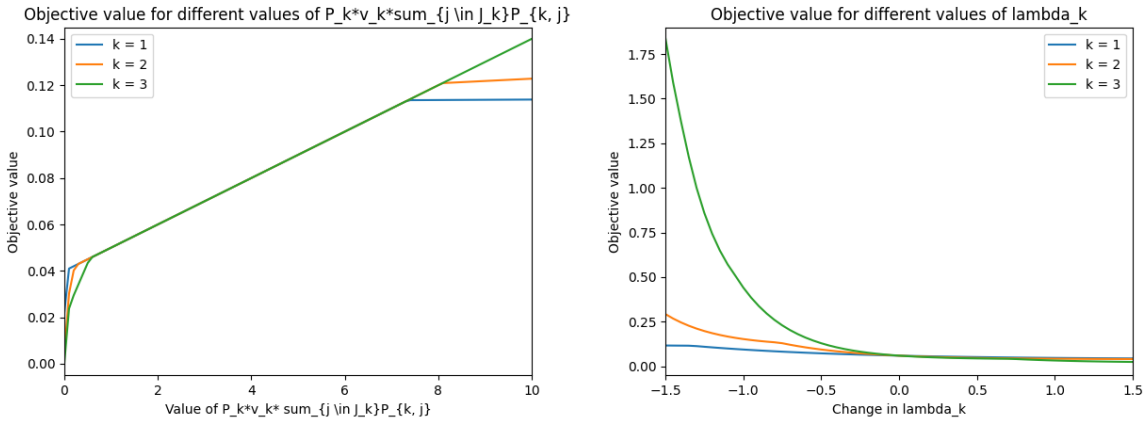
Table 6 and Figure 6 contain the results for the sensitivity analysis for the semi-flexible model, assuming identical threat levels. The stability ranges for the product of the parameters P_k, v_k, and $\sum_{j \in J_k} P_{k,j}$ for all $k \in K$ are identical and are observed to encompass nearly the entire possible range, denoting a limited effect of alterations in these parameters on the optimal solution. The products have an identical positive linear relationship with the objective value inside and outside their stability range. Alternatively, λ_k for each $k \in K$ exhibits a negative relationship with the objective value and a narrow stability range relative to its possible range. The latter implies that the optimal solution is highly sensitive to changes in λ_k. The stability range for B_s for each $s \in S$ remains constant at its base value, signaling extreme sensitivity to alterations in B_s. This is supported by the plot of objective values for different values of B_s. The objective value decreases sharply as B_s increases, but the rate of decrease slows down as B_s gets larger. This suggests that B_s has a strong negative impact on the objective value, especially at lower values of B_s. This high level of sensitivity stems from the fact that the optimal solution requires a consistent assignment of all available guards. Assigning fewer guards than available will always result in a suboptimal solution. As a result, any change in the number of available guards directly impacts the optimal solution.

Different threat levels For each parameter to test, a base value was determined. The base value for $P_k v_k \sum_{j \in J_k} P_{k,j}$ for each $K \in K$ again was set to 2. It was assumed that individuals 1, 2, and 3 were assigned low, general, and high threat levels, respectively, such that $\overline{d}_1 = 1$, $\overline{d}_2 = 2$, $\overline{d}_3 = 3$. The base values for λ_k for threatened individuals 1, 2, and 3 therefore included 4.605, 2.303, and 1.535, respectively, as derived from Equation 8. Finally, the base values for B_{early} and B_{late} were both set to $\overline{d}_1 + \overline{d}_2 + \overline{d}_3 = 6$.

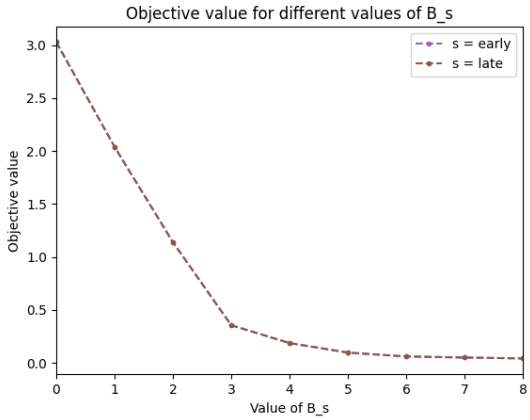
Table 7 and Figure 7 contain the results for the sensitivity analysis for the semi-flexible model, assuming different threat levels. The stability range for the product of the parameters P_k, v_k, and $\sum_{j \in J_k} P_{k,j}$ for all $k \in K$ are again observed to contain a substantial part of the possible range of values, indicating high sensitivity. The differences in the slopes of the effect on the objective value outside the stability range seem to depend on the threat level of each threatened individual $k \in K$ and thus λ_k. However, in contrast to the scenario where identical threat levels were assumed, the stability range for the individual under general threat is narrower, indicating higher sensitivity. Furthermore, the stability range of λ_k is wide for threatened individual 1, but narrow for the others. This implies a low sensitivity of the optimal solution to changes in λ_1, but a high sensitivity to changes in λ_2 and λ_3. It can be seen that λ_k has a negative relationship with the objective value for each $k \in K$, with the steepest decrease occurring for threatened individual 3, which has the highest threat level. Considering both the stability ranges and the effects on the objective value, this means that the sensitivity for λ_k for each $k \in K$ increases with the threat level. Finally, the stability range of B_s for each $s \in S$ remains again constant at its base value, signaling extreme sensitivity to alterations in B_s. No differences are observed compared to the scenario with identical threat levels for the individuals.

Parameter	Possible values	Base value	Step size	Stability range
$P_1 v_1 \sum_{j \in J_1} P_{1,j}$	[0, 10)	2	0.001	(0.018, 7.357)
$P_2 v_2 \sum_{j \in J_2} P_{2,j}$	[0, 10)	2	0.001	(0.219, 8.091)
$P_3 v_3 \sum_{j \in J_3} P_{3,j}$	[0, 10)	2	0.001	(0.543, 10)
λ_1	$(0, \infty)$	4.605	0.001[8]	$(3.274, 0.189 \cdot 10^7)$
λ_2	$(0, \infty)$	2.303	0.001	(1.535, 4.605)
λ_3	$(0, \infty)$	1.535	0.001	(0.010, 2.253)
B_{early}	$\{1, 2, \ldots, \infty\}$	6	1	{6}
B_{late}	$\{1, 2, \ldots, \infty\}$	6	1	{6}

Table 7: Results of parameter effects on the optimal solution for the semi-flexible model, assuming different threat levels.



(a) $P_k v_k \sum_{j \in J_k} P_{k,j}$ **(b)** λ_k



(c) B_s

Figure 7: Line plots of parameter effects on the objective value for the semi-flexible model, assuming different threat levels.

[8] To find the upper value of the stability range for λ_1, a step size of $0.001 \cdot 10^7$ was used.

4.3 Conclusion

Currently, a fixed number of guards is assigned to threatened individuals solely based on their threat level. This SQ proposed an alternative semi-flexible allocation method that allows for variation in the number of guards assigned to each individual between shifts. To implement this method, an NLIP model was formulated that minimizes the total expected damage on any given day. The total expected damage for a given day was calculated by summing the expected damage for all threatened individuals on that day. The expected damage for an activity was calculated using the Fine-Kinney method. This included multiplying the probability of an attacker intending to attack the threatened individual; the probability of an attack on the threatened individual at the considered activity, given that an attacker intends to attack the individual; the probability of success of an attack on the threatened individual during the considered activity, given the assigned guards and the individual's threat level, and assuming an attacker attacks; and the value of the threatened individual. In this calculation, an exponential success function was used to calculate the success probability depending on the threat level of an individual. 1000 situations were randomly generated and solved to evaluate the semi-flexible model objectively. The results showed that assuming identical threat levels for the individuals to be protected, the semi-flexible model decreased the expected damage on a day by 13.8% on average, and reduced the expected damage for 40% of the generated situations. When assuming different threat levels, the expected damage on a day decreased by 15.9% on average, and the expected damage was reduced for 43% of the generated situations. The model evaluation revealed the need for caution when determining the success functions.

5 Flexible risk-based allocation

SQ3. How can guard allocation be improved, assuming the number of guards assigned to each threatened individual can vary during shifts?

SQ3 aimed to create a quantitative model that improves guard allocation to threatened individuals. A requirement was that the number of guards assigned to each threatened individual could vary between and during shifts, in contrast to SQ2.

Flexible allocation This SQ analyzes the effect of changing the number of guards assigned to a threatened individual from activity to activity on guard allocation efficiency. A flexible allocation method is proposed to determine the number of guards based on the threat levels of the threatened individuals and the level of risk levels and their planned activities, rather than using a predetermined number for all individuals at a particular threat level. Unlike the semi-flexible method used in SQ2, the number of guards assigned to a threatened individual may vary during a shift because guards are allowed to travel to another threatened individual during a shift. They may travel alone or in groups. However, the number of available guards may be affected by the time needed to travel between the threatened individuals.

Consider again the scenario from Figure 2 and Figure 3, where six guards are to be divided over three threatened individuals. Now assume that the travel time between all threatened individuals includes one hour and the number of guards has to remain constant during an activity. Figure 8 shows the distribution of guards over the threatened individual in this scenario, following the proposed flexible allocation. In contrast to the base and semi-flexible allocation, the number of guards may be adjusted during a shift as needed. It can be seen that the number of guards assigned during each activity under the flexible allocation is always at least the number assigned under the semi-flexible allocation in Figure 3. It is important to note that this is just an example, and any deviation from the base and semi-flexible allocation will depend on the specific scenario.

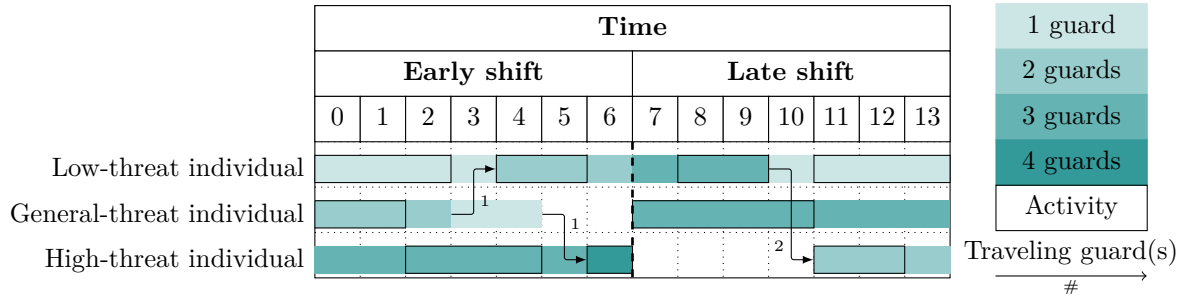


Figure 8: Example of proposed flexible risk-based allocation.

5.1 Modeling

This subsection presents the conceptual and mathematical model for the flexible allocation method to improve the allocation of guards to threatened individuals.

5.1.1 Conceptual model

Similar to the semi-flexible model in SQ2, the objective was to minimize the total expected damage on a given day. The same success functions and assumptions were employed for the flexible model. However, an additional assumption was introduced: the travel time for guards to move from one threatened individual to another during a shift was assumed one hour. The travel times before and after a shift were excluded because they are not part of the hours

available for protecting threatened individuals. Furthermore, the number of guards assigned to a threatened individual was assumed to remain constant during an activity.

5.1.2 Mathematical model

Building upon the conceptual model and using the notations from Subsubsection 4.1.2, the mathematical model was formulated. In this mathematical model, a new decision variable $y_t \in \mathbb{Z}_+$ was introduced, which contains the number of unassigned guards at time $t \in T$.

$$
\begin{aligned}
\min_{(d_{k,j})_{k\in K, j\in J_k}, (d'_{k,t})_{k\in K, t\in T}} \quad & \textstyle\sum_{k\in K} P_k v_k \sum_{j\in J_k} P_{k,j} f_k(d_{k,j}) & & \\
\text{s.t.} \quad d_{k,j} &= d'_{k,t} & & \forall k \in K, j \in J_k, t \in T_{k,j} \\
y_t &= B_s - \textstyle\sum_{k\in K} d'_{k,t} & & \forall t \in T_s, s \in S \\
y_t &= y_{t-1} - \textstyle\sum_{k\in K} (d'_{k,t} - d'_{k,t-1}) & & \forall t \in T'_s, s \in S \\
y_{t-1} &\geq \textstyle\sum_{k\in K} \max(0, d'_{k,t} - d'_{k,t-1}) & & \forall t \in T'_s, s \in S \\
d_{k,j} &\in \mathbb{Z}_+ & & \forall k \in K, j \in J_k \\
d'_{k,t} &\in \mathbb{Z}_+ & & \forall k \in K, t \in T \\
y_t &\in \mathbb{Z}_+ & & \forall t \in T
\end{aligned}
\tag{10}
$$

Equation 10 represents the flexible model. This model also includes a nonlinear integer programming (NLIP) problem and has the same convex nonlinear objective function (i.e., $\sum_{k\in K} P_k v_k \sum_{j\in J_k} P_{k,j} f_k(d_{k,j})$) as the semi-flexible model from SQ2.

The decision variables are $d_{k,j}$, $d'_{k,t}$, and y_t. $d_{k,j}$ and $d'_{k,t}$ denote the number of guards assigned to each threatened individual $k \in K$ during each activity $j \in J_k$ and at each time of the day $t \in T$, respectively. Additionally, y_t denotes the number of unassigned guards at time $t \in T$. A guard may be intentionally unassigned or become unassigned while traveling between threatened individuals.

The constraints in Equation 10 reflect the restrictions on the decision variables. The first constraint ensures that a constant number of guards are assigned to a threatened individual during an activity. The second constraint ensures that the total number of assigned plus unassigned guards equals the number of available guards at any given time. The third constraint ensures that traveling guards are unavailable for at least one hour by adjusting the number of unassigned guards when guards travel from one threatened individual to another. This is done by increasing and decreasing the number of unassigned guards when guards leave and join threatened individuals, respectively. The fourth constraint ensures that the total number of newly assigned guards does not exceed the number of unassigned guards from the previous hour. Constraints three and four do not apply to the first hour of a shift since travel times prior to the start of a shift can be disregarded. The fifth, sixth, and seventh constraints ensure that the decision variables are non-negative integers.

While the flexible model in Equation 10 assumes travel times between threatened individuals to be one hour, it is important to note that in reality, travel times may be longer. To address this, an extension of the flexible model is presented in Appendix I, considering travel times longer than one hour. Similar to the model in Equation 10, this extension does not account for varying travel times among threatened individuals.

5.2 Model evaluation

To evaluate the flexible model in Equation 10, the solving method and generated situations from SQ2 were used. For each generated situation, the flexible model was solved using CVXPY with the MOSEK solver. Appendix I contains the Python code that was used to solve the flexible model for each generated situation.

The interpretation of the model evaluation results should be done carefully since they highly depend on the model's assumptions. This is particularly crucial for the flexible model compared to the semi-flexible model, as the former relies on more assumptions.

5.2.1 Model comparison

To objectively evaluate the flexible allocation model, it was compared to the base, optimized base, and semi-flexible models using the generated situations. For an explanation of the base and optimized base models, see Subsubsection 4.2.3. The flexible model includes all possible assignments available in the base, optimized base, and semi-flexible models, making it superior by definition.

The same method as in Subsubsection 4.2.3 was employed to compare the models for a scenario where the threatened individuals to be protected all have the same threat level and a scenario where they all have different threat levels.

Identical threat levels To compare the base, optimized base, semi-flexible, and flexible models when considering threatened individuals with the same threat level, it was assumed that individuals 1, 2, and 3 were assigned general threat levels, such that $\overline{d}_1 = \overline{d}_2 = \overline{d}_3 = 2$ and $f_1(d) = f_2(d) = f_3(d) = e^{-2.303d}$, as derived from Equation 8. Furthermore, it was assumed that $B_{early} = B_{late} = \overline{d}_1 + \overline{d}_2 + \overline{d}_3 = 6$.

Figure 9 displays four box plots of objective values: one for the base model, one for the optimized base model, one for the semi-flexible model, and one for the flexible model. Table 8 shows the improvements in the mean objective value and the number of improved instances for each combination of models that includes the flexible model.

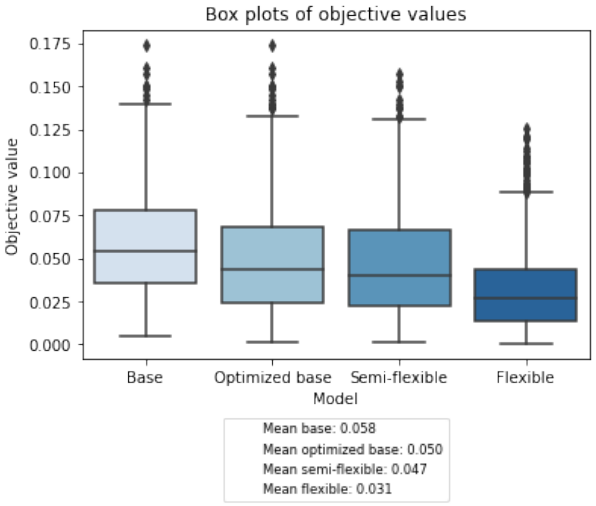


Figure 9: Box plot of objective values for the base, optimized base, semi-flexible, and flexible models, assuming identical threat levels.

Model		Improvement	
Reference	**Alternative**	**Mean objective value**	**Improved instances**
Base	Flexible	46.3%	99.2%
Optimized base	Flexible	37.7%	98.8%
Semi-flexible	Flexible	33.2%	97.9%

Table 8: Comparison of the base, optimized base, semi-flexible, and flexible models, assuming identical threat levels.

Different threat levels To compare the base, optimized base, semi-flexible, and flexible models when considering threatened individuals with different threat levels, it was assumed that individuals 1, 2, and 3 were assigned threat levels low, general, and high, respectively, such that $\overline{d}_1 = 1$, $\overline{d}_2 = 2$, $\overline{d}_3 = 3$, $f_1(d) = e^{-4.605d}$, $f_2(d) = e^{-2.303d}$, and $f_3(d) = e^{-1.535d}$, as derived from Equation 8. Furthermore, it was assumed that $B_{early} = B_{late} = \overline{d}_1 + \overline{d}_2 + \overline{d}_3 = 6$.

Figure 10 displays the box plots of objective values for each model, and Table 9 shows the improvements in the mean objective value and the number of improved instances for each combination of models that includes the flexible model.

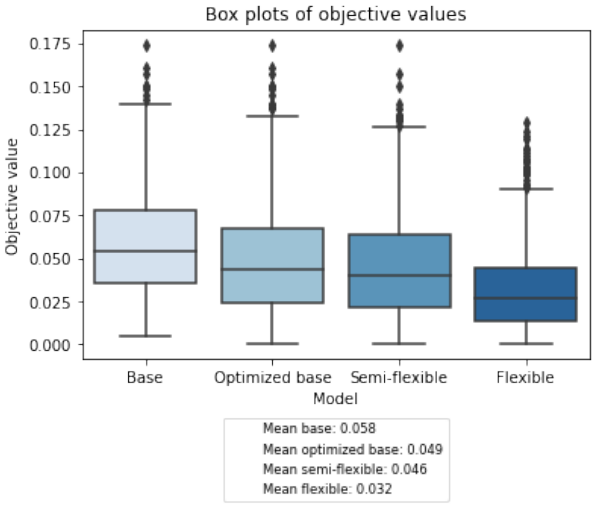


Figure 10: Box plot of objective values for the base, optimized base, semi-flexible, and flexible models, assuming different threat levels.

Model		Improvement	
Reference	**Alternative**	**Mean objective value**	**Improved instances**
Base	Flexible	45.5%	98.9%
Optimized base	Flexible	35.1%	98.4%
Semi-flexible	flexible	30.4%	97.6%

Table 9: Comparison of the base, optimized base, semi-flexible, and flexible models, assuming different threat levels.

Scenario comparison The figures and tables for both scenarios show that the flexible model outperforms the other models. The greatest difference is observed with the base model, followed by the optimized base and semi-flexible models. Although Figure 9 and Figure 10 are very similar, Table 8 and Table 9 reveal some interesting differences. The improvement in mean objective value and number of improved instances for each combination of models is slightly

greater when assuming identical threat levels than when assuming different threat levels. This means that using the flexible model can lead to slightly greater improvements when the threat levels of the individuals to be protected are similar than when the threat levels are different. This finding is surprising because the results from Subsubsection 4.2.3 revealed that the semi-flexible model showed slightly greater improvements for different threat levels than for identical threat levels.

5.2.2 Sensitivity analysis

Again, these two questions were to be answered for the sensitivity analysis:

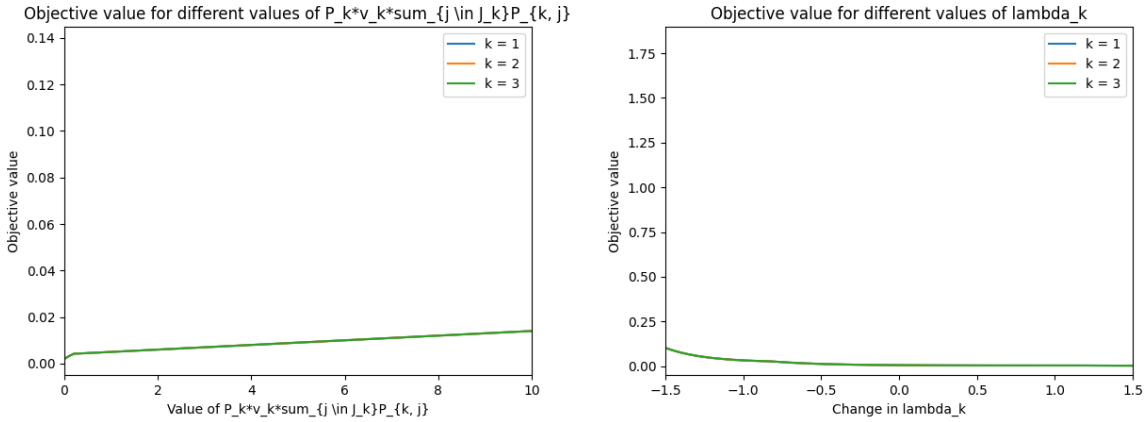
1. How does a change in a parameter affect the optimal solution?
2. How does a change in a parameter affect the objective value?

The method used to answer these questions was the same as the one employed for SQ2. However, in this case, the times of the activities were relevant as they affect the model. Therefore, it was assumed that $T_{1,0} = T_{3,0} = \{1\}$, $T_{1,1} = T_{2,0} = \{3\}$, $T_{2,1} = T_{3,1} = \{5\}$, $T_{1,2} = T_{3,2} = \{8\}$, $T_{1,3} = T_{2,3} = \{10\}$, and $T_{2,3} = T_{3,3} = \{12\}$. These times were arbitrarily chosen to ensure that guards could travel between threatened individuals during both shifts, but must be divided among the threatened individuals.

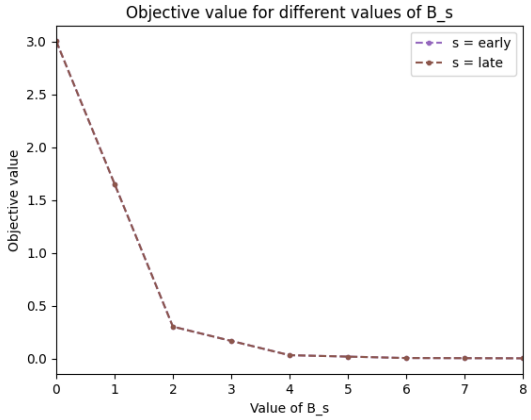
The parameters that were tested included the product of P_k, v_k, and $\sum_{j \in J_k} P_{k,j}$ for each $k \in K$, λ_k for each $k \in K$, and B_s for each $s \in S$. When testing the parameter $P_k v_k \sum_{j \in J_k} P_{k,j}$ for each $k \in K$, the value of each $P_k v_k P_{k,j}$ for each $j \in J_k$ was set to $\frac{P_k v_k \sum_{j \in J_k} P_{k,j}}{|J_k|}$. The parameter $T_{k,j}$ for each $j \in J_k$ and $k \in K$ was not tested because it was not possible to consistently change $T_{k,j}$ for each $j \in J_k$ and $k \in K$, making it impossible to fairly asses the sensitivity for this parameter.

Parameter	Possible values	Base value	Step size	Stability range
$P_1 v_1 \sum_{j \in J_1} P_{1,j}$	[0, 10)	2	0.001	(0.200, 10)
$P_2 v_2 \sum_{j \in J_2} P_{2,j}$	[0, 10)	2	0.001	(0.200, 10)
$P_3 v_3 \sum_{j \in J_3} P_{3,j}$	[0, 10)	2	0.001	(0.200, 10)
λ_1	$(0, \infty)$	2.303	0.001	(1.484, 3.492)
λ_2	$(0, \infty)$	2.303	0.001	(1.484, 3.492)
λ_3	$(0, \infty)$	2.303	0.001	(1.484, 3.492)
B_{early}	$\{1, 2, \ldots, \infty\}$	6	1	{6}
B_{late}	$\{1, 2, \ldots, \infty\}$	6	1	{6}

Table 10: Results of parameter effects on the optimal solution for the flexible model, assuming identical threat levels.



(a) $P_k v_k \sum_{j \in J_k} P_{k,j}$ **(b)** λ_k



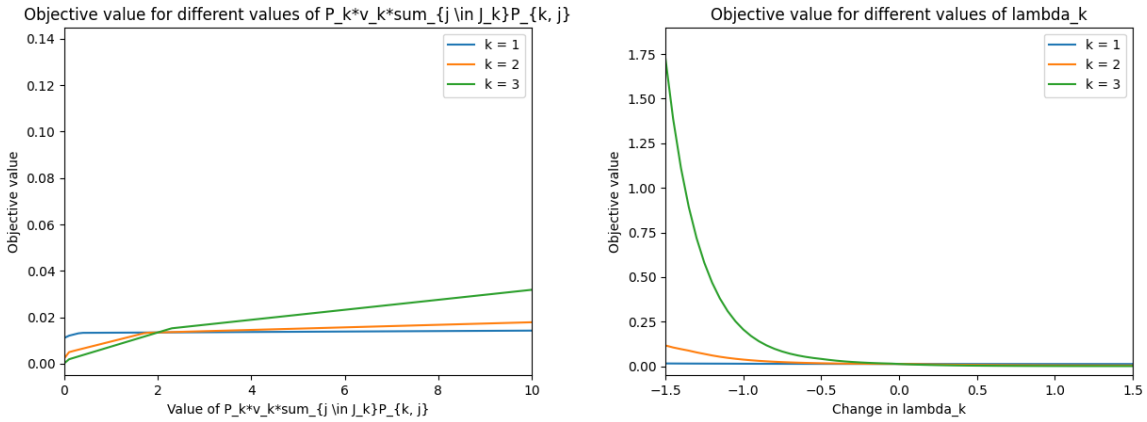
(c) B_s

Figure 11: Line plots of parameter effects on the objective value for the flexible model, assuming identical threat levels.

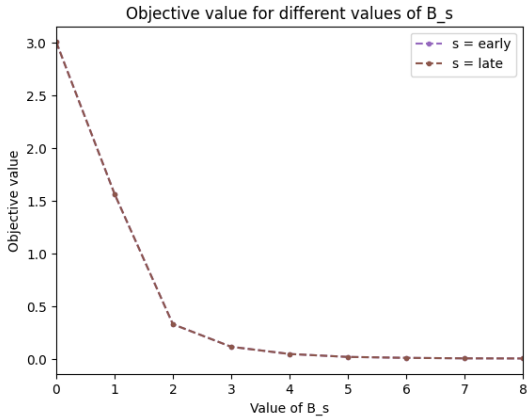
Identical threat levels Table 10 and Figure 11 contain the results for the sensitivity analysis for the flexible model, assuming identical threat levels. The stability range for the products of the parameters P_k, v_k, and $\sum_{j \in J_k} P_{k,j}$ for all $k \in K$ are identical and relatively wide, denoting low sensitivity. This is supported by the line plots, showing a small positive relation with the objective value. Additionally, although the stability range for λ_k for each $k \in K$ is narrow, indicating high sensitivity, λ_k shows a relatively weak relationship with the objective value. This implies that the optimal solution is less stable to changes in λ_k for each $k \in K$, but the effects of these changes on the objective value are smaller. Finally, the stability range of B_s for each $s \in S$ remains constant at its base value, signaling extreme sensitivity to alterations in B_s.

Parameter	Possible values	Base value	Step size	Stability range
$P_1 v_1 \sum_{j \in J_1} P_{1,j}$	[0, 10)	2	0.001	(0.341, 10)
$P_2 v_2 \sum_{j \in J_2} P_{2,j}$	[0, 10)	2	0.001	(1.743, 10)
$P_3 v_3 \sum_{j \in J_3} P_{3,j}$	[0, 10)	2	0.001	(0.049, 2.295)
λ_1	$(0, \infty)$	4.605	0.001	(3.491, 6.382)
λ_2	$(0, \infty)$	2.303	0.001	(1.045, 2.376)
λ_3	$(0, \infty)$	1.535	0.001	(1.484, 3.058)
B_{early}	$\{1, 2, \ldots, \infty\}$	6	1	{6}
B_{late}	$\{1, 2, \ldots, \infty\}$	6	1	{6}

Table 11: Results of parameter effects on the optimal solution for the flexible model, assuming different threat levels.



(a) $P_k v_k \sum_{j \in J_k} P_{k,j}$ **(b)** λ_k



(c) B_s

Figure 12: Line plots of parameter effects on the objective value for the flexible model, assuming different threat levels.

Different threat levels Table 11 and Figure 12 contain the results for the sensitivity analysis for the flexible model, assuming different threat levels. The stability range for the product of the parameters P_k, v_k, and $\sum_{j \in J_k} P_{k,j}$ differ for each $k \in K$. For $k = 1, 2$, the ranges are wide, indicating low sensitivity, but for $k = 3$, the range is narrow, showing high sensitivity. In contrast to the scenario where identical threat levels were assumed, differences in the slopes of the effect on the objective value seem to depend on the threat level of each threatened individual $k \in K$ and thus λ_k. Furthermore, the stability range for λ_1 is a bit wider, and for λ_2 and λ_3 a bit smaller. The impact on the objective value appears to vary based on the level of threat, with a lesser effect for lower threat levels. Finally, the stability range of B_s for each $s \in S$ remains constant at its base value, signaling extreme sensitivity to alterations in B_s.

5.3 Conclusion

Currently, a fixed number of guards is assigned to threatened individuals solely based on their threat level. However, this SQ has proposed an alternative allocation method that allows for variation in the number of guards assigned to each individual from activity to activity. To implement this method, an NLIP model was formulated to minimize the total expected damage on any given day. The travel time between all threatened individuals was assumed to be one hour. The total expected damage for a given day was calculated in a similar way as for the semi-flexible model in SQ2. The generated situations from SQ2 were solved to evaluate the flexible model objectively. The results showed that assuming identical threat levels for the individuals to be protected, the flexible model decreased the expected damage on a day by 46.3% on average, and reduced the expected damage for 99.2% of the generated situations. When assuming different threat levels, the expected damage on a day decreased by 45.5% on average, and the expected damage was reduced for 98.9% of the generated situations. The model evaluation revealed the need for caution when determining the success functions.

6 Attacker-defender game

SQ4. How can guard allocation to threatened individuals be improved by anticipating strategic behavior?

SQ4 aimed to propose a model to improve the allocation of guards to threatened individuals while anticipating strategic behavior. In contrast to SQ2 and SQ3, which only consider defensive strategic behavior, SQ4 includes both strategic offensive and defensive behavior. Similar to SQ2, however, the number of guards assigned to threatened individuals was only allowed to vary between shifts.

6.1 Modeling

To anticipate strategic behavior in guard allocation, game theory can be used. Game theory is a mathematical framework for analyzing interactions between decision-makers, often used to understand strategic behavior in competitive situations. This subsection proposes a game theoretic model to improve the allocation of guards to threatened individuals. First, the conceptual model is presented. Next, the conceptual model is translated into a mathematical model.

6.1.1 Conceptual model

A single-period, simultaneous-move, two-player, zero-sum attacker-defender game is proposed. For a general explanation of attacker-defender games, please refer to Subsection 2.3. The modeling approaches for the proposed game will be explained following the structure of Hunt and Zhuang (2023). These approaches include the time horizon, number of players, nature of decision variables, objective functions, and sequence of moves.

Time horizon A single-period game is considered, which involves analysis after a single round of play. In contrast, a multi-period game requires the defender to choose a defensive strategy for each period, with the attacker selecting an attacking strategy after observing the defensive strategy. Although a multi-period game might be more representative, a single-period game is considered to maintain simplicity. For this single-period game, one round of play (i.e., one period) includes a shift where the number of available guards is assumed to remain constant.

Number of players The model involves two players: a defender and an attacker. The attacker aims to attack one of the threatened individuals that the defender is trying to protect during the considered shift. In practice, these individuals may not be threatened by the same attacker, and a game with multiple attackers and one defender would be more appropriate. However, it is assumed that the probability of multiple attackers actively targeting different threatened individuals during a shift is so small that these attackers can be combined into a single intelligent attacker. This eliminates the need for an excessively large and complex model.

Players' decision variables The decision variables for the attacker and defender are related to target selection and resource allocation, respectively. These variables represent the strategic choices available to each player to achieve their objectives. During the considered shift, the attacker's objective is to select a threatened individual to attack, while the defender's objective is to allocate guards among the threatened individuals to protect them against the attacker's attack. Thus, the decision variables for the attacker and defender consist of the individual to be attacked during the considered shift and the number of guards to be assigned to each individual during that shift, respectively.

Players' objective functions In attacker-defender games, the strategic goals of the players are determined by their objective functions. For this study, a zero-sum game is assumed, i.e., the attacker aims to maximize the payoff and the defender aims to minimize it. A game is considered zero-sum when the objective functions of both players add up to zero, meaning that the payoffs of the attacker and defender are directly opposed. The attacker's payoff is calculated based on the strategies chosen by both players. The defender's payoff is calculated by subtracting the attacker's payoff from zero, ensuring that the total of both players' payoffs equals zero.

For this study, the attacker's payoff includes the expected damage for an attack during the considered shift and is calculated using the Fine-Kinney method (Fine, 1971; Kinney & Wiruth, 1976). This method presents a mathematical formula for calculating the risk score associated with a given threat. The risk score is calculated by multiplying the exposure, probability, and consequence. Exposure refers to the likelihood of the threat manifesting, while probability refers to the likelihood that a threat will result in an expected consequence. Consequence refers to the most likely effect of a potential threat. The Fine-Kinney risk score is a useful tool for estimating the conditional expected damage from an attack on a given threatened individual during a given shift, if one were to occur, by multiplying the following factors.

1. The risk exposure of the threatened individual.
2. The probability of success for an attack on the threatened individual given the assigned guards and the present threat level, given that the attacker attacks.
3. The value of the threatened individual.

Factor 1 involves the probability of the attacker attacking during one of the planned activities of the threatened individual during the shift, given that the attacker intends to attack the individual. This probability is calculated by summing the individual conditional probabilities for the planned activities during the shift, which can be calculated using the conditional logit model from SQ1. Factor 1 can be interpreted as the 'exposure' used in the Fine-Kinney method. Although this method assumes a range of 0 to 10, this study uses a range of 0 to 1 for factor 1.

Factor 2 is the conditional probability of a successful attack if the attacker attacks. The probability of success can be determined by using the success functions from the semi-flexible and flexible models in SQ2 and SQ3, which take into account the threat level of the threatened individual and the number of guards assigned to the individual during the shift. This factor can be interpreted as the 'probability' used in the Fine-Kinney method. Although this method assumes a range of 0 to 10, this study uses a range of 0 to 1 for factor 2.

Factor 3 relates to the attacker's valuation of the threatened individual. To maintain simplicity, it is assumed that values of the threatened individuals are the same to both the attacker and the defender. Therefore, factor 3 concerns the value of the threatened individual to the Dutch police, which may include the potential impact of an attack on the individual as well as organizational or political interests. The organization is responsible for determining an appropriate value for a threatened individual. This factor can be interpreted as the 'consequence' used in the Fine-Kinney method. Although this method assumes a range of 0 to 100, this study uses a range of 0 to 10 for factor 3.

Unlike the semi-flexible and flexible models in SQ2 and SQ3, the expected damage calculation in this SQ is conditional as it does not include the probability of the attacker intending to attack the threatened individual. This is because this probability reflects the strategy chosen by the attacker, which is therefore not exogenously but endogenously determined.

Sequence of moves Although less frequently used in the literature on attacker-defender games (Hunt & Zhuang, 2023), a simultaneous-move game is employed. Simultaneous-move games involve players selecting their strategies without knowledge of their opponent's choices. In contrast, sequential-move games involve the defender moving first, followed by the attacker after observing the defender's strategy. This study assumes that decisions are made simultaneously because the players may not be able to observe their opponent's decisions. The defender chooses the strategy for the allocation of guards before the shift, while the attacker's strategy can only be observed during or after the shift. Therefore, the defender can only observe the attacker's strategy after their own strategy has already been chosen. Additionally, the attacker may not be able to observe the defender's strategy during the shift when guards are not recognizable or visible. Therefore, it is assumed that players make simultaneous moves.

6.1.2 Mathematical model

The following notations are used to formulate the mathematical model.

$K \subseteq \mathbb{N}$	Set of threatened individuals
$x_k \in \mathbb{Z}_+$	Number of guards assigned to threatened individual $k \in K$
$f_k : \mathbb{Z}_+ \to [0,1]$	Success function for attacks on threatened individual $k \in K$ [9]
$R_k \in [0,1]$	Risk exposure of threatened individual $k \in K$
$v_k \in [0,10]$	Value of threatened individual $k \in K$
$B \in \mathbb{N}$	Number of available guards

The considered game is a single-period, simultaneous-move, two-player, zero-sum attacker-defender game. The strategic form of this game is a tuple $(S_A; S_D; \hat{f})$, with the following items:

- $S_A = K$ includes the attacker's strategy set, indicating the threatened individuals to potentially be attacked. A strategy from the attacker's strategy set is denoted by $k \in S_A$, where k is the threatened individual chosen to be attacked.
- S_D is the defender's strategy set, which consists of every possible allocation of the B available guards to the set of threatened individuals S_A, where all guards must be allocated. Strategies that do not assign all guards are not included in S_D because these are always dominated by strategies where all guards are assigned. The size of the defender's strategy set is given by $|S_D| = \binom{|S_A|+B-1}{B}$[10] $= \frac{(|S_A|+B-1)!}{B!\cdot(|S_A|-1)!}$. A strategy from the defender's strategy set is denoted by the vector $x = (x_k)_{k \in S_A} \in S_D$, where $x_k \in \mathbb{Z}_+$ is the number of guards assigned to threatened individual $k \in S_A$ under defender strategy $x \in S_D$. For each strategy $x \in S_D$ should always hold that $\sum_{k \in K} x_k = B$. For example, for a situation where $S_A = \{1,2\}$, each strategy $x \in S_D$ has the form $(x_1, x_2) \in \mathbb{N}^2$, where $x_1 + x_2 = B$. Furthermore, assume that the strategies in S_D are in reversed lexicographic order.[11]
- $\hat{f} : \prod_{n \in \{A,D\}} S_n \to \mathbb{R}$ is the payoff function of the game. This function describes the payoff to the attacker resulting from all possible choices of strategies by both players. Because the payoffs of the attacker and defender are directly opposing, it suffices to denote the payoffs for the attacker only. The payoff function is given by $\hat{f}(k,x) = R_k v_k f(x_k)$, where $k \in S_A$ and $x \in S_D$. For simplicity, $R_k v_k$ is replaced by the constant $r_k \in [0,10]$, such that $\hat{f}(k,x) = r_k f_k(x_k)$ for each $k \in S_A$.

[9] Similar to SQ2 and SQ3, it is assumed that $f_k(d) = e^{-4.605d}$, $f_k(d) = e^{-2.303d}$, and $f_k(d) = e^{-1.535d}$ if threatened individual $k \in K$ has threat level low, general, and high, respectively.

[10] This follows from the stars-and-bars theorem by Feller (1950), which states that the number of ways to place n indistinguishable objects into k distinguishable bins is $\binom{k+n-1}{n} = \binom{k+n-1}{k-1}$.

[11] Reversed lexicographic order in mathematics arranges elements based on their values in descending order, prioritizing higher-valued components first.

The tabular representation of the strategic-form game $(S_A; S_D; \hat{f})$ is as follows.

		Defender	
		$(x_1, \ldots, x_{\lvert S_A \rvert})$	$(x'_1, \ldots, x'_{\lvert S_A \rvert})$
Attacker	1	$r_1 f_1(x_1)$	$r_1 f_1(x'_1)$
	$\lvert S_A \rvert$	$r_{\lvert S_A \rvert} f_{\lvert S_A \rvert}(x_{\lvert S_A \rvert})$	$r_{\lvert S_A \rvert} f_{\lvert S_A \rvert}(x'_{\lvert S_A \rvert})$

Table 12: The game in strategic form.

The corresponding matrix game can be denoted by the triple (X_A, X_D, F). Here F is the payoff matrix, where $X_A = S_A$ denotes the rows of F that the attacker can choose, and $X_D = \{1, \ldots, |S_D|\}$ denotes the columns of F that the defender can choose. If the attacker and defender play row $i \in X_A$ and column $j \in X_D$, respectively, then the attacker receives and the defender pays $F(i, j) = \hat{f}(i, x^j) = r_i f_i(x_i^j)$, where $x^j = (x_i^j)_{i \in X_A}$ is the jth strategy in S_D. The matrix F can be represented as follows.

$$F = \begin{bmatrix} r_1 f_1(x_1^1) & \ldots & r_1 f_1(x_1^{|S_D|}) \\ \vdots & \ddots & \vdots \\ r_{|S_A|} f_{|S_A|}(x_{|S_A|}^1) & \ldots & r_{|S_A|} f_{|S_A|}(x_{|S_A|}^{|S_D|}) \end{bmatrix} \tag{11}$$

The simplest variant of the described game includes a situation where only one guard is available to protect two threatened individuals, such that $(\{1, 2\}; \{(1, 0), (0, 1)\}; \hat{f})$. Recall that $f_k(0) = 1$ for all $k \in S_A$.

		Defender	
		(1,0)	(0,1)
Attacker	1	$r_1 f_1(1)$	r_1
	2	r_2	$r_2 f_2(1)$

Table 13: Simplest variant of the game in strategic form.

The corresponding matrix game can be denoted by $(\{1, 2\}, \{1, 2\}, F)$.

$$F = \begin{bmatrix} r_1 f_1(1) & r_1 \\ r_2 & r_2 f_2(1) \end{bmatrix} \tag{12}$$

6.2 Model solving

Solving a game theoretic model refers to identifying one or more Nash equilibria. A Nash equilibrium (NE) (Nash, 1950) is a strategy profile that is stable in the sense that no player will find it profitable to unilaterally deviate from his strategy and choose another strategy.

Pure strategy NE An NE in pure strategies includes an NE where all players choose one of their available pure strategies (Slikker, 2022). To find a pure strategy NE in a matrix game, one must search for a saddle point. A game's saddle point is an NE in pure strategies, and vice versa. A game may, but need not, have one or more saddle points, and thus pure strategy NEs. The concept of a saddle point is based on the players' considerations. The minimum value in a row indicates the smallest gain for the row player if they select that particular strategy. The row player chooses the strategy that yields the highest gain among the minimum values in each row. The resulting (maximal minimum) gain is called the lower value of the matrix game. Similarly, the maximum value in a column represents the largest loss for the column player if they select that particular strategy. The column player chooses the strategy that yields the lowest loss among the maximum values in each column. The resulting (minimal maximum) loss is called the upper value of the matrix game. If the resulting row/column combination represents a strategy profile that corresponds to the lowest value in its row and the highest value in its column (meaning that the lower and upper values are equivalent), then the strategy profile is a saddle point. The value of a matrix game with a saddle point is equal to the lower and upper values.

In the matrix game (X_A, X_D, F), the attacker's maximal minimum gain includes the lower value $\underline{w}(F) = \max_{i \in X_A} \min_{j \in X_D} F(i,j)$ and the defender's minimal maximum loss includes the upper value $\overline{w}(F) = \min_{j \in X_D} \max_{i \in X_A} F(i,j)$. If $\underline{w}(F) = \overline{w}(F)$, then the matrix game (X_A, X_D, F) has one or more saddle points (i.e., strategy profiles that correspond to the lowest value in its row and the highest value in its column) and associated value $w(F)$, for which holds that $\underline{w}(F) = w(F) = \overline{w}(F)$.

Mixed strategy NE Note that not all games have an NE in pure strategies. However, incorpora- ting mixed strategies can effectively resolve this issue. A mixed strategy NE includes an NE where at least one player plays a randomized strategy (Slikker, 2022). A randomized strategy involves a probability distribution, or randomization, of the available set of pure strategies for a player in a game. In contrast to pure strategy NEs, a game always has one and only one pure strategy NE. To find the mixed strategy NE in a matrix game, one must search for probability distributions over the available sets of pure strategies such that no player has an incentive to use a different probability distribution.

The mixed extension of the matrix game (X_A, X_D, F) can be denoted by (Y_A, Y_D, π), with strategy sets $Y_A = \{p \in [0,1]^{|X_A|} | \sum_{i \in X_A} p_i = 1\}$ and $Y_D = \{q \in [0,1]^{|X_D|} | \sum_{j \in X_D} q_j = 1\}$, and payoff function $\pi(p,q) = \sum_{i \in X_A} \sum_{j \in X_D} p_i q_j F(i,j)$. The lower value and upper value are denoted by $\underline{v}(F) = \max_{p \in Y_A} \min_{j \in X_D} \pi(p, d_j)$ and $\overline{v}(F) = \min_{q \in Y_D} \max_{i \in X_A} \pi(e_i, q)$, respectively. Here $e_i = p$ with $p_i = 1$, and $d_j = q$ with $q_j = 1$. Finally, $v(F)$ denotes the value of the matrix game (X_A, X_D, F) in mixed strategies, for which always holds that $\underline{v}(F) = v(F) = \overline{v}(F)$..

6.2.1 Nash equilibria identification

Simple variant The simplest variant of the described game is $(\{1,2\}; \{(1,0),(0,1)\}; \hat{f})$. See Table 13, for a description of the situation in strategic form, and Equation 12 for the payoff matrix of the corresponding matrix game $(\{1,2\}, \{1,2\}, F)$. It can easily be observed that the following conditional pure NEs exist:

- $(1, (1,0))$ if $r_1 f_1(1) \geq r_2$
- $(2, (0,1))$ if $r_2 f_2(1) \geq r_1$

The attacker and defender choose to attack and defend threatened individual 1 during the shift that the matrix game represents, respectively, if the expected damage from an attack

on this individual, given the presence of the guard, is higher than the expected damage from an attack on threatened individual 2, given that no guards are present. The value of the matrix game is then $w(F) = r_1 f_1(1)$, which includes the expected damage from an attack on threatened individual 1, given one guard is present. Alternatively, they choose to attack and defend threatened individual 2, respectively, if the expected damage from attacking this individual, given the presence of the guard, is higher than the expected damage from attacking threatened individual 1, given the absence of guards. The value of the matrix game is then $w(F) = r_2 f_2(1)$, which includes the expected damage from an attack on threatened individual 2, given one guard is present. Note that these two conditional pure NEs can never co-exist.[12]

The mixed extension of $(\{1,2\},\{1,2\},F)$ can be denoted by (Y_A, Y_D, π). Here, $Y_A = \{p \in [0,1]^2 | \sum_{i\in\{1,2\}} p_i = 1\}$ and $Y_D = \{q \in [0,1]^2 | \sum_{j\in\{1,2\}} q_j = 1\}$. Furthermore, $e_i = p$ with $p_i = 1$, and $d_j = q$ with $q_j = 1$. To identify the NE in mixed strategies, it must be determined for which p the defender is indifferent between the strategies d_1 and d_2, and for which q the attacker is indifferent between the strategies e_1 and e_2. The defender is indifferent for p satisfying:

$$\begin{aligned} p_1 r_1 f_1(1) + p_2 r_2 &= p_1 r_1 + p_2 r_2 f_2(1) \\ p_1 r_1 f_1(1) + (1-p_1) r_2 &= p_1 r_1 + (1-p_1) r_2 f_2(1) \\ p_1 &= \frac{r_2(f_2(1)-1)}{r_1(f_1(1)-1) + r_2(f_2(1)-1)} \end{aligned} \tag{13}$$

$$\begin{aligned} p_2 &= 1 - p_1 \\ p_2 &= 1 - \frac{r_2(f_2(1)-1)}{r_1(f_1(1)-1) + r_2(f_2(1)-1)} \\ p_2 &= \frac{r_1(f_1(1)-1)}{r_1(f_1(1)-1) + r_2(f_2(1)-1)} \end{aligned} \tag{14}$$

Similarly, the attacker is indifferent for q satisfying:

$$\begin{aligned} q_1 r_1 f_1(1) + q_2 r_1 &= q_1 r_2 + q_2 r_2 f_2(1) \\ q_1 r_1 f_1(1) + (1-q_1) r_1 &= q_1 r_2 + (1-q_1) r_2 f_2(1) \\ q_1 &= \frac{-r_1 + r_2 f_2(1)}{r_1(f_1(1)-1) + r_2(f_2(1)-1)} \end{aligned} \tag{15}$$

$$\begin{aligned} q_2 &= 1 - q_1 \\ q_2 &= 1 - \frac{-r_1 + r_2 f_2(1)}{r_1(f_1(1)-1) + r_2(f_2(1)-1)} \\ q_2 &= \frac{r_1 f_1(1) - r_2}{r_1(f_1(1)-1) + r_2(f_2(1)-1)} \end{aligned} \tag{16}$$

Therefore, the following mixed NE exists:

- $(p_1 e_1 + p_2 e_2, q_1 d_1 + q_2 d_2)$

This means that the attacker will attack threatened individual 1 with $(p_1 \cdot 100)\%$ chance, and threatened individual 2 with $(p_2 \cdot 100)\%$ chance. p_i can be interpreted as the difference in expected damage when no guards instead of one guard are assigned to threatened individual

[12] Rewriting $r_2 f_2(1) \geq r_1$ to $r_2 \geq r_1 / f_2(1)$ lead to the inequality $r_1 f_1(1) \geq r_2 \geq r_1 / f_2(1)$, which never holds because $f_1(1), f_2(1) \in (0,1)$.

$i \in \{1,2\}$ proportionate to the sum of differences for both individuals. Similarly, the defender will allocate one guard to threatened individual 1 with $(q_1 \cdot 100)\%$ chance, and one to threatened individual 2 with $(p_2 \cdot 100)\%$ chance. q_i can be interpreted as the difference in expected damage when one guard is assigned to threatened individual $i' \in \{1,2\} \backslash \{i\}$ and when no guards are assigned to individual $i \in \{1,2\}$ proportionate to the sum of differences for both individuals. The value of the matrix game for this mixed NE is then $v(F) = p_1 q_1 r_1 f_1(1) + p_2 q_1 r_2 + p_1 q_2 r_1 + p_2 q_2 r_2 f_2(1)$, which includes the expected damage during the shift.

General variant Now, consider the matrix game $(S_A; S_D; \hat{f})$ from Table 12 and the corresponding matrix game (X_A, X_D, F) from Equation 11. Recall that the mixed extension is denoted by (Y_A, Y_D, π), where $Y_A = \{p \in [0,1]^{|X_A|} | \sum_{i \in X_A} p_i = 1\}$ and $Y_D = \{q \in [0,1]^{|X_D|} | \sum_{j \in X_D} q_j = 1\}$. Furthermore, $e_i = p$ with $p_i = 1$, and $d_j = q$ with $q_j = 1$. For the matrix game (X_A, X_D, F), the following conditional pure NEs exist:

- $(e_i, d^B_{i,j})$ if $r_i f_i(B) \geq r_{i'} \; \forall i \in X_A, i' \in X_A \backslash \{i\}$

For $x^j_i = B$, $d^B_{i,j} = q$ where $q_j = 1$. This means that $(e_i, d^B_{i,j})$ represents a strategy profile in which the attacker attacks threatened individual i, and the defender assigns all B guards to this individual. The attacker and defender choose to attack and defend threatened individual $i \in X_A$ during the shift that the matrix game represents, respectively, if the expected damage from an attack on this individual, given the presence of all B guards, is higher than the expected damage from an attack on threatened individual $i' \in X_A \backslash i$, given that no guards are present. The value of the matrix game for this conditional pure NE is $w(F) = r_i f_i(B)$, which includes the expected damage from an attack on threatened individual $i \in X_A$, given B guards are present. Note that there can never exist more than one of these conditional pure NEs.

To identify the NE in mixed strategies, it must be determined for which p the defender is indifferent between the strategies d_j for $j \in X_D$, and for which q the attacker is indifferent between the strategies e_i for $i \in X_A$. Recall that $\pi(p,q) = \sum_{i \in X_A} \sum_{j \in X_D} p_i q_j F(i,j)$ and that $F(i,j) = \hat{f}(i, x^j) = r_i f_i(x^j_i)$. The values of p and q can be determined by finding $\underline{v}(F) = \max_{p \in Y_A} \min_{j \in X_D} \pi(p, d_j)$ and $\overline{v}(F) = \min_{q \in Y_D} \max_{i \in X_A} \pi(e_i, q)$. This can be done by solving the following optimization problems.

$$
\begin{aligned}
\max_{p} \quad & \underline{v}(F) & & & \\
\text{s.t.} \quad & \textstyle\sum_{i \in X_A} p_i r_i f_i(x^j_i) & \geq \quad & \underline{v}(F) & \forall j \in X_D \\
& \textstyle\sum_{i \in X_A} p_i & = \quad & 1 & \\
& p_i & \geq \quad & 0 & \forall i \in X_A
\end{aligned}
\tag{17}
$$

$$
\begin{aligned}
\min_{q} \quad & \overline{v}(F) & & & \\
\text{s.t.} \quad & \textstyle\sum_{j \in X_D} q_j r_i f_i(x^j_i) & \leq \quad & \overline{v}(F) & \forall i \in X_A \\
& \textstyle\sum_{j \in X_D} q_j & = \quad & 1 & \\
& q_j & \geq \quad & 0 & \forall j \in X_D
\end{aligned}
\tag{18}
$$

These optimization problems include linear programming problems. Using the solutions for p and q, the following mixed NE can be formulated:

- $(\sum_{i \in X_A} p_i e_i, \sum_{j \in X_D} q_j d_j)$

This means that the attacker will attack threatened individual $i \in X_A$ with $(p_i \cdot 100)\%$ chance and that the defender will play allocation $j \in X_D$ with $(q_j \cdot 100)\%$ chance. The value of the matrix game for this mixed NE equals $\underline{v}(F) = v(F) = \overline{v}(F)$, which includes the expected damage during the shift, calculated from the optimization problems in Equation 17 and Equation 18.

6.3 Model evaluation

To evaluate the proposed matrix game, the solving method and generated situations from SQ2 and SQ3 were used. For each generated situation, the games for the early and late shifts were solved separately by finding the optimal values for p and q using CVXPY with the MOSEK solver. Appendix J includes the Python code utilized to compute the game value for each generated situation.

The interpretation of the model evaluation results should be done with care as they highly depend on the game's assumptions.

6.3.1 Model comparison

To objectively evaluate the proposed game, it was compared to a base and a relaxed base game using the generated situations from SQ2 and SQ3. In the base game, the defender's strategy set S_D included only the pure strategy where each threatened individual was assigned a predetermined number of guards (i.e., the strategy included $(\overline{d}_1, \overline{d}_2, \overline{d}_3)$, where $\overline{d}_1 + \overline{d}_2 + \overline{d}_3 = B$). The strategy set S_A of the attacker included the three threatened individuals. The attacker was allowed to play in mixed strategies. In the relaxed base game, S_D included all possible ways to allocate the available B guards among the three threatened individuals in S_A. The relaxed base game assumed the attacker could play in mixed strategies, while the defender was constrained to pure strategies. In the proposed game, S_D included all possible ways to allocate the available B guards among the three threatened individuals in S_A, and both players were allowed to play in mixed strategies. The proposed game includes all possible assignments in the base and relaxed base games, making it superior by definition. Similarly, by definition, the relaxed base game is superior to the base game.

The models were compared for two scenarios. These included a scenario where the threatened individuals to be protected all have the same threat level and a scenario where they all have different threat levels.

Identical threat levels To compare the base, relaxed base, and proposed games when considering threatened individuals with the same threat level, it was assumed that individuals 1, 2, and 3 were assigned general threat levels, such that $\overline{d}_1 = \overline{d}_2 = \overline{d}_3 = 2$ and $f_1(d) = f_2(d) = f_3(d) = e^{-2.303d}$, as derived from Equation 8. Furthermore, for all games, it was assumed that $B = \overline{d}_1 + \overline{d}_2 + \overline{d}_3 = 6$.

Figure 13 displays three box plots of values: one for the base game, one for the relaxed base game, and one for the proposed game. Table 14 shows the improvements in the mean value and the number of improved instances for each combination of games that includes the proposed game.

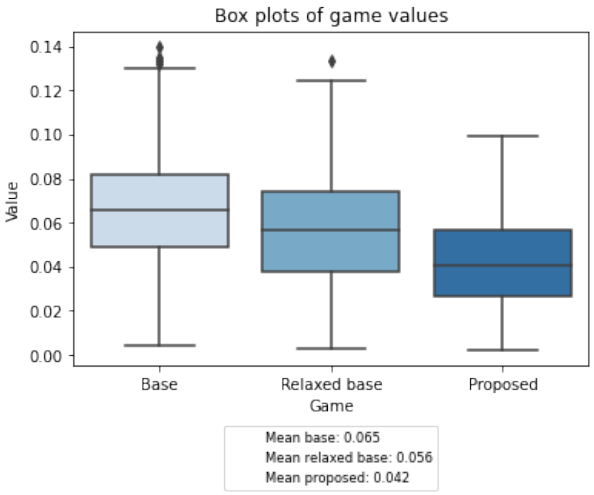


Figure 13: Box plot of values for the base, relaxed base, and proposed games, assuming identical threat levels.

Game		Improvement	
Reference	**Alternative**	**Mean value**	**Improved instances**
Base	Relaxed base	13.9%	42.8%
Base	Proposed	36.0%	100%
Relaxed base	Proposed	25.7%	100%

Table 14: Comparison of the base, relaxed base, and proposed game, assuming identical threat levels.

Different threat levels To compare the base, relaxed base, and proposed games when considering threatened individuals with different threat levels, it was assumed that individuals 1, 2, and 3 were assigned threat levels low, general, and high, respectively, such that $\overline{d}_1 = 1$, $\overline{d}_2 = 2$, $\overline{d}_3 = 3$, $f_1(d) = e^{-4.605d}$, $f_2(d) = e^{-2.303d}$, and $f_3(d) = e^{-1.535d}$, as derived from Equation 8. Furthermore, for all games, it was assumed that $B = \overline{d}_1 + \overline{d}_2 + \overline{d}_3 = 6$.

Figure 14 displays the box plots of values for each game, and Table 15 shows the improvements in the mean value and the number of improved instances for each combination of games that includes the proposed game.

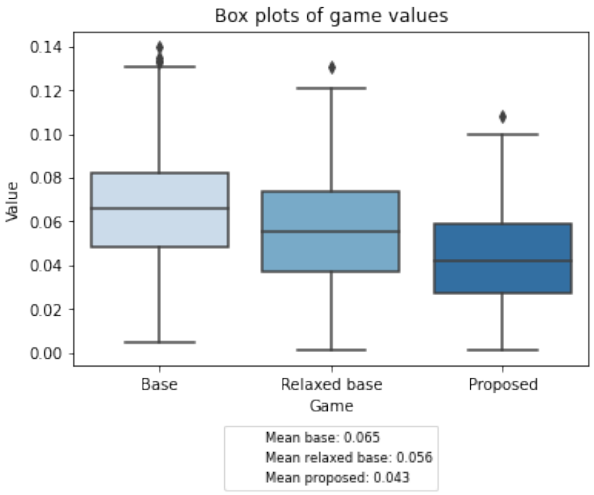


Figure 14: Box plot of values for the base, relaxed base, and proposed game, assuming different threat levels.

Game		Improvement	
Reference	**Alternative**	**Mean value**	**Improved instances**
Base	Relaxed base	14.7%	44.9%
Base	Proposed	34.1%	100%
Relaxed base	Proposed	22.7%	100%

Table 15: Comparison of the base, relaxed base, and proposed game, assuming different threat levels.

Scenario comparison The figures and tables for both scenarios show that the relaxed base game, and even more so the proposed game, outperforms the base game. Although Figure 13 and Figure 14 are very similar, Table 14 and Table 15 reveal some interesting differences. While all instances for the proposed game were improved relative to the base and relaxed base games for both scenarios, the number of improved instances for the relaxed base game relative to the base game was slightly greater when assuming different threat levels than when assuming identical threat levels. Furthermore, the mean value improvement in the relaxed base game was slightly greater when assuming different threat levels compared to identical threat levels. This means that using the relaxed base game can lead to slightly greater improvements when the threat levels of the individuals to be protected are different than when the threat levels are similar. For the remaining game combinations, the proposed game showed slightly greater improvements for identical threat levels compared to different threat levels. This means that using the proposed game can lead to slightly greater improvements when the threat levels of the individuals to be protected are similar than when the threat levels are different.

Relation to risk-based allocation models The results for the model evaluation of the proposed game and the semi-flexible and flexible allocation models from SQ2 and SQ3 can unfortunately not be compared because of a different determination for the probability of an attacker intending to attack a threatened individual. For the proposed game in this SQ, this

probability is endogenously determined for each threatened individual $k \in K$ and is denoted by $p_k \in [0, 1]$, where $\sum_{k \in K} p_k = 1$. For the risk-based allocation models from SQ2 and SQ3, the probability is exogenously determined for each threatened individual $k \in K$ and is denoted by $P_k \in [0, 1]$. However, $\sum_{k \in K} P_k$ does not necessarily equal 1. Comparing the model evaluation results of these models to those of the proposed game would therefore be unfair.

6.4 Conclusion

This SQ presented a method for assigning guards to threatened individuals while anticipating strategic behavior. To implement this method, a game-theoretic model was proposed that minimizes the total expected damage for a given shift. The model included a single-period, simultaneous-move, two-player, zero-sum attacker-defender game. The calculation of the payoff to the attacker included the conditional expected damage of an attack if one were to occur during a shift, which was computed using the Fine-Kinney method. This included multiplying the risk exposure of the threatened individual; the probability of success for an attack on the threatened individual given the assigned guards and the present threat level, given that the attacker attacks; and the value of the threatened individual. In this calculation, the same exponential success functions that depend on the threat level were used to calculate the success probability of an attack as for SQ2 and SQ3. It was shown under which conditions this game has an NE in pure strategies and that the game always has an NE in mixed strategies. Additionally, the methods for identifying these NEs were shown. To evaluate the proposed game objectively, it was solved for each of the generated situations from SQ2 and SQ3. The results showed that the proposed game reduced the expected damage on a day by 25.7% and 22.7% on average, assuming identical and different threat levels of the threatened individuals to be protected, respectively. Furthermore, the proposed game reduced the expected damage for all generated situations, independent of the threat levels.

7 Conclusion

This section presents the answers to the research questions and subsequent recommendations, followed by the scientific contributions of this study.

7.1 Conclusions, limitations, and recommendations

To answer the main research question of improving the B&B system's allocation of guards to threatened individuals, four subquestions were defined. The recommendations, limitations, and answers to these questions are presented first, followed by those for the main research question.

SQ1. How can the risk exposure of threatened individuals during the day be quantified?

To assess the risk exposure of threatened individuals throughout the day, their exposure was quantified as the probability of an attack. To achieve this, a mathematical model was developed to calculate the probability of an attack on each threatened individual for each planned activity on that day. Nine risk factors were identified for activities planned by threatened individuals. Of these, five are for visits to locations and four are for trips from one location to another. To analyze the effect of these risk factors on the probability of an attack, a discrete choice model, specifically a conditional logit model, was presented. It was demonstrated how to use a stated choice experiment to collect data from experts, after which it was shown how to estimate and evaluate the model using this data. Due to the variability in the collected data, some estimated coefficients did not show significance.

A limitation of this study is that the identified risk factors were based on an internal confidential study conducted by the Dutch police a few years ago, which may no longer be accurate. It is recommended to evaluate and possibly revise the list of risk factors as incorrect risk factors can affect the model's effectiveness. Additionally, attributes related to attackers and threatened individuals should also be considered. Furthermore, this study assumed that attackers have homogeneous preferences. In reality, however, preferences may vary. Therefore, it is recommended to examine how preference heterogeneity could be incorporated. To address this, a mixed or latent class logit model could be used instead of a conditional logit model. Furthermore, variability in the collected data caused some estimates to be insignificant. For future data collection, it is important to determine which employees of the Dutch police are experts on the topic. Additionally, collecting more responses will reduce variability in the data and help develop a more informative model.

SQ2. How can guard allocation be improved, assuming the number of guards assigned to each threatened individual does not vary during shifts?

SQ3. How can guard allocation be improved, assuming the number of guards assigned to each threatened individual can vary during shifts?

To improve the allocation of guards under the given assumptions, a semi-flexible and a flexible risk-based allocation method were proposed. The semi-flexible method allowed for variation in the number of guards assigned to each individual, but not during shifts. In contrast, the flexible method permitted variation in the number of guards assigned to each individual both between and during shifts. The flexible model included all possible assignments available in the base, optimized base, and semi-flexible models, making it superior by definition. To implement both methods, two nonlinear integer programming models were formulated, minimizing the total expected damage on any given day. The total expected damage for a given day was calculated as the sum of the expected damage for each activity planned by each threatened individual on that day. The expected damage for an activity was calculated using the Fine-Kinney method. This included multiplying the probability of an attacker intending to attack the threatened

individual; the probability of an attack on the threatened individual at the considered activity, given that an attacker intends to attack the individual; the probability of success of an attack on the threatened individual during the considered activity, given the assigned guards and the individual's threat level, and assuming an attacker attacks; and the value of the threatened individual. In this calculation, an exponential success function was used to calculate the success probability depending on the threat level of an individual. To evaluate the models objectively, 1000 situations were randomly generated and solved. The results showed that the semi-flexible model performed best when the threat levels of the individuals to be protected were different. The flexible model, however, performed best when the threat levels were identical.

The objective of the nonlinear integer programming models was to minimize the total expected damage on a given day. However, this study has a limitation in that it only considered one objective and its associated formulation. It is recommended that this objective and its formulation be reconsidered, as another objective may be preferable in practice. Furthermore, the probability of a successful attack for an activity was estimated using an exponential success function that depends on the individual's threat level. The evaluation of the model recommends caution when determining success functions. It is therefore recommended to reconsider which threat levels can be distinguished and which success functions correspond to these levels. Additionally, the flexible model assumes a travel time of one hour between all threatened individuals, which greatly influences the model's behavior. To make the model more dynamic, it is recommended to investigate how to determine and incorporate variable travel times between threatened individuals. Additionally, it is suggested to focus on further developing and implementing the flexible model, as it is superior to the semi-flexible model by definition. Although implementing the semi-flexible model may be easier, the benefits of the flexible model make it the better choice.

SQ4. How can guard allocation to threatened individuals be improved by anticipating strategic behavior?

To demonstrate how guard allocation can be improved while anticipating strategic behavior, a game-theoretic model was developed. More specifically, a single-period, simultaneous-move, two-player, zero-sum attacker-defender game was proposed. It was assumed that the probability of multiple attackers actively targeting different threatened individuals during a shift is so small that these attackers can be combined into a single intelligent attacker. Similar to SQ2, variation in the number of guards was allowed between but not during shifts. The decision of the defender was how many guards to assign to each threatened individual during a shift, and the attacker's decision concerned selecting one of those individuals to attack. The calculation of the attacker's payoff included the conditional expected damage of an attack if one were to occur during a shift, which was computed using the Fine-Kinney method. This included multiplying the risk exposure of the threatened individual; the probability of success for an attack on the threatened individual given the assigned guards and the present threat level, given that the attacker attacks; and the value of the threatened individual. In this calculation, the same exponential success functions that depend on the threat level were used to calculate the success probability of an attack as for SQ2 and SQ3. The conditions under which this game has a pure strategy Nash equilibrium and that the game always has a mixed strategy Nash equilibrium were shown. Furthermore, it was shown how these Nash equilibria could be identified. To objectively evaluate the game, it was solved for each generated situation from SQ2 and SQ3.

For illustrative purposes, a relatively simple attacker-defender game was used to model guard allocation, which limits this study. However, it is recommended to investigate a more complex type of attacker-defender game to better reflect reality. This could include a game with multiple attackers, a multi-period game, or a game with non-opposing payoffs for the players, depending on what best reflects reality. The payoff was calculated using the Fine-Kinney method and an

exponential success function. It is suggested to review the payoff and its calculation to ensure it accurately reflects reality. Furthermore, just as for SQ2 and SQ3, it is recommended to reconsider which threat levels can be distinguished and which success functions correspond to these levels. Finally, this study did not consider the effects of varying the number of guards on security quality. However, it is crucial to take this into account when deciding whether to use the semi-flexible or flexible model in practice. Based on the study results, however, it is recommended to focus on the development and implementation of the flexible model rather than the semi-flexible model, since the former is by definition superior to the latter.

RQ. How can the B&B system's guard allocation to threatened individuals inside the Netherlands be improved?

This study proposed and evaluated multiple quantitative methods to improve guard allocation to threatened individuals. The potential of these methods was demonstrated by comparing them to a base method representing the current allocation method. The results indicate that the use of these methods leads to large improvements.

The study recommends establishing a clear objective for guard allocation by seeking consensus. This can be achieved by organizing an interactive session, conducting interviews, or observing the current allocation method. Once the objective is established, appropriate mathematical models can be employed. It is also recommended to investigate whether the probability of an attacker intending to attack a threatened individual should be determined exogenously or endogenously. Exogenous determination involves estimating this probability based on external information. In that case, the risk-based allocation models would be the most appropriate to improve guard allocation. Endogenous determination involves estimating the probability based on strategic behavior. In that case, the attacker-defender game would be the most appropriate to improve guard allocation. A limitation of this study is that it used generated data. To gain more useful insights, future analyses should use actual data on situations instead of generating them. This includes data on threatened individuals and their agendas, potential attackers, guard availability, travel times, and decision-making. It is advisable to begin collecting this data in a structured manner as soon as possible.

7.2 Scientific contributions

In addition to the recommendations to the Dutch police, this study also contributes to scientific literature in three ways. First, it was shown how to use discrete choice modeling to quantify risk exposure. This was done by determining the probability of attack for the activities planned by threatened individuals. While discrete choice modeling has been commonly used in criminology, its application to attacks on threatened individuals has not been previously described in the literature. Second, an approach was proposed to improve the allocation of guards to threatened individuals while satisfying certain constraints. Two nonlinear integer programming models were formulated, and the Fine-Kinney method was used for the objective function. This approach for guard allocation to threatened individuals has not been previously reported in the literature. Third, a game-theoretic model was proposed to improve the allocation of guards to threatened individuals. The proposed model is an attacker-defender game, which is a type of game that has been widely studied in the literature. However, this study is unique in that it uses the Fine-Kinney method to determine the payoff.

References

Baudains, P. (2015). *Spatio-temporal modelling of civil violence: Four frameworks for obtaining policy-relevant insights* (Unpublished doctoral dissertation).

Baudains, P., Braithwaite, A., & Johnson, S. (2013). Target choice during extreme events: A discrete spatial choice model of the 2011 London riots. *Criminology*, *51*(2), 251–285. doi: https://doi.org/10.1111/1745-9125.12004

Baudains, P., Braithwaite, A., & Johnson, S. (2016). The London riots – 2: A discrete choice model. In *Approaches to geo-mathematical modelling* (pp. 170–191). doi: https://doi.org/10.1002/9781118937426.ch10

Bernasco, W. (2006). Co-offending and the choice of target areas in burglary. *Journal of Investigative Psychology and Offender Profiling*, *3*(3), 139–155. doi: https://doi.org/10.1002/jip.49

Bernasco, W. (2010a). Modeling micro-level crime location choice: Application of the discrete choice framework to crime at places. *Journal of Quantitative Criminology*, *26*(1), 113–138. doi: https://doi.org/10.1007/s10940-009-9086-6

Bernasco, W. (2010b). A sentimental journey to crime: Effects of residential history on crime location choice. *Criminology*, *48*(2), 389–416. doi: https://doi.org/10.1111/j.1745-9125.2010.00190.x

Bernasco, W., & Block, R. (2009). Where offenders choose to attack: A discrete choice model of robberies in Chicago. *Criminology: An Interdisciplinary Journal*, *47*(1), 93–130. doi: https://doi.org/10.1111/j.1745-9125.2009.00140.x

Bernasco, W., Johnson, S., & Ruiter, S. (2015). Learning where to offend: Effects of past on future burglary locations. *Applied Geography*, *60*, 120–129. doi: https://doi.org/10.1016/j.apgeog.2015.03.014

Bernasco, W., & Kooistra, T. (2010). Effects of residential history on commercial robbers' crime location choices. *European Journal of Criminology*, *7*(4), 251–265. doi: https://doi.org/10.1177/1477370810363372

Bernasco, W., & Nieuwbeerta, P. (2005). How do residential burglars select target areas? *The British Journal of Criminology*, *45*(3), 296–315. doi: https://doi.org/10.1093/bjc/azh070

Bernasco, W., & Ruiter, S. (2014). Crime location choice. In G. Bruinsma & D. Weisburd (Eds.), *Encyclopedia of criminology and criminal justice* (pp. 691–699). New York, NY: Springer New York. doi: https://doi.org/10.1007/978-1-4614-5690-2{_}440

Bernasco, W., Ruiter, S., & Block, R. (2017). Do street robbery location choices vary over time of day or day of week? A test in Chicago. *Journal of Research in Crime and Delinquency*, *54*(2), 244–275. doi: https://doi.org/10.1177/0022427816680681

Bier, V., Haphuriwat, N., Menoyo, J., Zimmerman, R., & Culpen, A. (2008). Optimal resource allocation for defense of targets based on differing measures of attractiveness. *Risk Analysis*, *28*(3), 763–770. doi: https://doi.org/10.1111/j.1539-6924.2008.01053.x

Bier, V., Oliveros, S., & Samuelson, L. (2007). Choosing what to protect: Strategic defensive allocation against an unknown attacker. *Journal of Public Economic Theory*, *9*(4), 563–587. doi: https://doi.org/10.1111/j.1467-9779.2007.00320.x

Bierlaire, M. (2023). A short introduction to Biogeme.

Bliemer, M., & Rose, J. (2014). Designing and conducting stated choice experiments . In *Handbook of choice modelling.* Cheltenham: Edward Elgar Publishing.

Casorrán, C., Fortz, B., Labbé, M., & Ordóñez, F. (2019). A study of general and security Stackelberg game formulations. *European Journal of Operational Research*, *278*(3), 855–868. doi: https://doi.org/110.1016/j.ejor.2019.05.012

ChoiceMetrics. (2021). Ngene 1.3 user manual & reference guide.

Clare, J., Fernandez, J., & Morgan, F. (2009). Formal evaluation of the impact of barriers

and connectors on residential burglars' macro-level offending location choices. *Australian & New Zealand Journal of Criminology*, *42*(2), 139–158. doi: https://doi.org/10.1375/acri.42.2.139

Cook, R., & Nachtsheim, C. (1980). A comparison of algorithms for constructing exact D-optimal designs. *Technometrics*, *22*(3), 315. doi: https://doi.org/10.2307/1268315

Curtis-Ham, S., Bernasco, W., Medvedev, O., & Polaschek, D. (2022). The importance of importance sampling: Exploring methods of sampling from alternatives in discrete choice models of crime location choice. *Journal of Quantitative Criminology*, *38*(4), 1003–1031. doi: https://doi.org/10.1007/s10940-021-09526-5

Dighe, N., Zhuang, J., & Bier, V. (2009). Secrecy in defensive allocations as a strategy for achieving more cost-effective attacker deterrence. *International Journal of Performability Engineering*, *5*(1), 31–43.

Feller, W. (1950). *An introduction to probability theory and its applications* (2nd ed., Vol. 1). Wiley.

Feng, Q., Cai, H., & Chen, Z. (2019). Using game theory to optimize the allocation of defensive resources on a city scale to protect chemical facilities against multiple types of attackers. *Reliability Engineering & System Safety*, *191*, 105900. doi: https://doi.org/10.1016/j.ress.2017.07.003

Fine, W. (1971). Mathematical evaluations for controlling hazards. *Journal of Safety Research*, *3*(4), 157–166.

Frith, M. (2019). *A discrete choice and configurational analysis of burglary offence location choices* (Unpublished doctoral dissertation). University College London.

Frith, M., Johnson, S., & Fry, H. (2017). Role of the street network in burglars' spatial decision-making. *Criminology*, *55*(2), 344–376. doi: https://doi.org/10.1111/1745-9125.12133

Golany, B., Goldberg, N., & Rothblum, U. (2017). A two-resource allocation algorithm with an application to large-scale zero-sum defensive games. *Computers & Operations Research*, *78*, 218–229. doi: https://doi.org/10.1016/j.cor.2016.08.013

Golany, B., Kaplan, E., Marmur, A., & Rothblum, U. (2009). Nature plays with dice – terrorists do not: Allocating resources to counter strategic versus probabilistic risks. *European Journal of Operational Research*, *192*(1), 198–208. doi: https://doi.org/10.1016/j.ejor.2007.09.001

Guan, P., He, M., Zhuang, J., & Hora, S. (2017). Modeling a multitarget attacker–defender game with budget constraints. *Decision Analysis*, *14*(2), 87–107. doi: https://doi.org/10.1287/deca.2017.0346

Guan, P., & Zhuang, J. (2016). Modeling resources allocation in attacker-defender dames with "warm up" CSF. *Risk Analysis*, *36*(4), 776–791. doi: https://doi.org/10.1111/risa.12502

Hao, M., Jin, S., & Zhuang, J. (2009). Robustness of optimal defensive resource allocations in the face of less fully rational attacker. *IIE Annual Conference. Proceedings*, 886–891.

Hausken, K. (2008). Strategic defense and attack for reliability systems. *Reliability Engineering & System Safety*, *93*(11), 1740–1750. doi: https://doi.org/10.1016/j.ress.2007.11.002

Hausken, K. (2011). Protecting complex infrastructures against multiple strategic attackers. *International Journal of Systems Science*, *42*(1), 11–29. doi: https://doi.org/10.1080/00207720903434789

Hausken, K. (2014). Choosing what to protect when attacker resources and asset valuations are uncertain. *Operations Research and Decisions*, *24*(3), 23–44. doi: https://doi.org/10.5277/ord140302

Hausken, K., & Bier, V. (2011). Defending against multiple different attackers. *European Journal of Operational Research*, *211*(2), 370–384. doi: https://doi.org/10.1016/j.ejor.2010.12.013

Hausken, K., & He, F. (2016). On the effectiveness of security countermeasures for critical

infrastructures. *Risk Analysis*, *36*(4), 711–726. doi: https://doi.org/10.1111/risa.12318
Hausken, K., & Zhuang, J. (2012). The timing and deterrence of terrorist attacks due to exogenous dynamics. *The Journal of the Operational Research Society*, *63*(6), 726–735.
Huber, J., & Zwerina, K. (1996). The importance of utility balance in efficient choice designs. *Journal of Marketing Research*, *33*(3), 307–317. doi: https://doi.org/10.1177/002224379603300305
Hunt, K., & Zhuang, J. (2023). A review of attacker-defender games: Current state and paths forward. *European Journal of Operational Research*. doi: https://doi.org/10.1016/j.ejor.2023.04.009
Jenelius, E., Westin, J., & Holmgren, (2010). Critical infrastructure protection under imperfect attacker perception. *International Journal of Critical Infrastructure Protection*, *3*(1), 16–26. doi: https://doi.org/10.1016/j.ijcip.2009.10.002
Johnson, S., & Summers, L. (2015). Testing ecological theories of offender spatial decision making using a discrete choice model. *Crime & Delinquency*, *61*(3), 454–480. doi: https://doi.org/10.1177/0011128714540276
Jonker, J. (2022). *Recordaantal meldingen bedreigde politici.* Retrieved from https://nos.nl/artikel/2456538-recordaantal-meldingen-bedreigde-politici
Keeney, R. (2007). Modeling values for anti-terrorism analysis. *Risk Analysis*, *27*(3), 585–596. doi: https://doi.org/10.1111/j.1539-6924.2007.00910.x
Kinney, G., & Wiruth, A. (1976). *Practical risk analysis for safety management* (Vol. 5865). China Lake, CA: Naval Weapons Center.
Kuralarasan, K., & Bernasco, W. (2022). Location choice of snatching offenders in Chennai city. *Journal of Quantitative Criminology*, *38*(3), 673–696. doi: https://doi.org/10.1007/s10940-021-09514-9
Levitin, G., & Hausken, K. (2009). Intelligence and impact contests in systems with redundancy, false targets, and partial protection. *Reliability Engineering & System Safety*, *94*(12), 1927–1941. doi: https://doi.org/10.1016/j.ress.2009.06.010
Marchment, Z., & Gill, P. (2019). Modelling the spatial decision making of terrorists: The discrete choice approach. *Applied Geography*, *104*, 21–31. doi: https://doi.org/10.1016/j.apgeog.2019.01.009
McFadden, D. (1979). Quantitative methods for analysing travel behaviour of individuals: Some recent developments. In D. Hensher & P. Stopher (Eds.), *Behavioural travel modelling* (pp. 279–318). London: Croom Helm.
McLay, L., Rothschild, C., & Guikema, S. (2012). Robust adversarial risk analysis: A level-k approach. *Decision Analysis*, *9*(1), 41–54. doi: https://doi.org/10.1287/deca.1110.0221
Menting, B., Lammers, M., Ruiter, S., & Bernasco, W. (2016). Family matters: Effects of family members' residential areas on crime location choice. *Criminology*, *54*(3), 413–433. doi: https://doi.org/10.1111/1745-9125.12109
Ministerie van Justitie en Veiligheid. (2023). Circulaire met betrekking tot de bewaking en beveiliging van personen, objecten en diensten 2023. *Staatscourant*, *39608*.
Musegaas, M., Schlicher, L., & Blok, H. (2022). Stackelberg production-protection games: Defending crop production against intentional attacks. *European Journal of Operational Research*, *297*(1), 102–119. doi: https://doi.org/10.1016/j.ejor.2021.04.012
Nash, J. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, *36*(1), 48–49.
Nikoofal, M., & Gümüs, M. (2015). On the value of terrorist's private information in a government's defensive resource allocation problem. *IIE Transactions*, *47*(6), 533–555. doi: https://doi.org/10.1080/0740817X.2014.938844
Nikoofal, M., & Zhuang, J. (2012). Robust allocation of a defensive budget considering an attacker's private information. *Risk Analysis*, *32*(5), 930–943. doi: https://doi.org/10.1111/j.1539-6924.2011.01702.x

Nikoofal, M., & Zhuang, J. (2015). On the value of exposure and secrecy of defense system: First-mover advantage vs. robustness. *European Journal of Operational Research*, *246*(1), 320–330. doi: https://doi.org/10.1016/j.ejor.2015.04.043

Orme, B. (1998). Sample size issues for conjoint analysis studies. In *Sawtooth software technical paper* (chap. 7). Sequim.

Oxford Learner's Dictionary of Academic English. (n.d.-a). *Risk.* Retrieved from https://www.oxfordlearnersdictionaries.com/definition/academic/risk1?q=risk

Oxford Learner's Dictionary of Academic English. (n.d.-b). *Threat.* Retrieved from https://www.oxfordlearnersdictionaries.com/definition/academic/threat?q=threat

Paulson, E., Linkov, I., & Keisler, J. (2016). A game theoretic model for resource allocation among countermeasures with multiple attributes. *European Journal of Operational Research*, *252*(2), 610–622. doi: http://dx.doi.org/10.1016/j.ejor.2016.01.026

Payyappalli, V., Zhuang, J., & Jose, V. (2017). Deterrence and risk preferences in sequential attacker-defender games with continuous efforts. *Risk Analysis*, *37*(11), 2229–2245. doi: https://doi.org/10.1111/risa.12768

PersVeilig. (2023). *Analyse meldingen.* Retrieved from https://www.persveilig.nl/over-persveilig/analyse-meldingen

Powell, R. (2007a). Allocating defensive resources with private information about vulnerability. *American Political Science Review*, *101*(4), 799–809. doi: https://doi.org/10.1017/S0003055407070530

Powell, R. (2007b). Defending against Terrorist Attacks with Limited Resources. *The American Political Science Review*, *101*(3), 527–541.

Powell, R. (2009). Sequential, nonzero-sum "Blotto": Allocating defensive resources prior to attack. *Games and Economic Behavior*, *67*(2), 611–615. doi: https://doi.org/10.1016/j.geb.2009.03.011

Rose, J., & Bliemer, M. (2013). Sample size requirements for stated choice experiments. *Transportation*, *40*(5), 1021–1041. Retrieved from https://doi.org/10.1007/s11116-013-9451-z doi: https://doi.org/10.1007/s11116-013-9451-z

Rothschild, C., McLay, L., & Guikema, S. (2012). Adversarial risk analysis with incomplete information: A level-kapproach. *Risk Analysis*, *32*(7), 1219–1231. doi: https://doi.org/10.1111/j.1539-6924.2011.01701.x

Sándor, Z., & Wedel, M. (2001). Designing conjoint choice experiments using managers' prior beliefs. *Journal of Marketing Research*, *38*(4), 430–444. doi: https://doi.org/10.1509/jmkr.38.4.430.18904

Schlicher, L., & Lurkin, V. (2024). Fighting pickpocketing using a choice-based resource allocation model. *European Journal of Operational Research*, *315*(2), 580–595. doi: https://doi.org/10.1016/j.ejor.2023.12.007

Shan, X., & Zhuang, J. (2013a). Cost of equity in homeland security resource allocation in the race of a strategic attacker. *Risk Analysis*, *33*(6), 1083–1099. doi: https://doi.org/10.1111/j.1539-6924.2012.01919.x

Shan, X., & Zhuang, J. (2013b). Hybrid defensive resource allocations in the face of partially strategic attackers in a sequential defender–attacker game. *European Journal of Operational Research*, *228*(1), 262–272. doi: http://dx.doi.org/10.1016/j.ejor.2013.01.029

Shan, X., & Zhuang, J. (2018). Modeling cumulative defensive resource allocation against a strategic attacker in a multi-period multi-target sequential game. *Reliability Engineering & System Safety*, *179*, 12–26. doi: http://dx.doi.org/10.1016/j.ress.2017.03.022

Slikker, M. (2022). *Game theory with applications to supply chain management* (Tech. Rep.). Eindhoven: Eindhoven University of Technology.

Smith, M., & Brown, D. (2007). Discrete choice analysis of spatial attack sites. *Information Systems and e-Business Management*, *5*(3), 255–274. doi: https://doi.org/10.1007/s10257

-007-0045-1

Song, G., Bernasco, W., Liu, L., Xiao, L., Zhou, S., & Liao, W. (2019). Crime feeds on legal activities: Daily mobility flows help to explain thieves' target location choices. *Journal of Quantitative Criminology*, *35*(4), 831–854. doi: https://doi.org/10.1007/s10940-019-09406-z

Townsley, M., Birks, D., Bernasco, W., Ruiter, S., Johnson, S., White, G., & Baum, S. (2015). Burglar target selection: A cross-national comparison. *Journal of Research in Crime and Delinquency*, *52*(1), 3–31. doi: https://doi.org/10.1177/0022427814541447

Townsley, M., Birks, D., Ruiter, S., Bernasco, W., & White, G. (2016). Target selection models with preference variation between offenders. *Journal of Quantitative Criminology*, *32*(2), 283–304. doi: https://doi.org/10.1007/s10940-015-9264-7

Train, K. (2009). *Discrete choice methods with simulation* (2nd ed.). Cambridge: Cambridge University Press. doi: https://doi.org/10.1017/CBO9780511805271

Vandeviver, C., Neutens, T., van Daele, S., Geurts, D., & Vander Beken, T. (2015). A discrete spatial choice model of burglary target selection at the house-level. *Applied Geography*, *64*, 24–34. doi: https://doi.org/10.1016/j.apgeog.2015.08.004

van Miltenburg, C., van Straaten, G., & Bouwmeester, J. (2022). *Agressie, bedreiging en intimidatie bij advocaten* (Tech. Rep.). Amsterdam: I&O Research.

Wang, C., & Bier, V. (2011). Target-hardening decisions based on uncertain multiattribute terrorist utility. *Decision Analysis*, *8*(4), 286–302. doi: https://doi.org/10.1287/deca.1110.0218

Wu, D., Yan, X., Peng, R., & Wu, S. (2020a). Optimal defence-attack strategies between one defender and two attackers. *Journal of the Operational Research Society*, *71*(11), 1830–1846. doi: https://doi.org/10.1080/01605682.2019.1630332

Wu, D., Yan, X., Peng, R., & Wu, S. (2020b). Risk-attitude-based defense strategy considering proactive strike, preventive strike and imperfect false targets. *Reliability Engineering & System Safety*, *196*, 106778. doi: https://doi.org/10.1016/j.ress.2019.106778

Xu, Z., & Zhuang, J. (2019). A study on a sequential one-defender-N-attacker game. *Risk Analysis*, *39*(6), 1414–1432. doi: https://doi.org/10.1111/risa.13257

Xue, Y., & Brown, D. (2006). Spatial analysis with preference specification of latent decision makers for criminal event prediction. *Decision Support Systems*, *41*(3), 560–573. doi: https://doi.org/10.1016/j.dss.2004.06.007

Yang, R., Kiekintveld, C., Ordóñez, F., Tambe, M., & John, R. (2013). Improving resource allocation strategies against human adversaries in security games: An extended study. *Artificial Intelligence*, *195*, 440–469. doi: http://dx.doi.org/10.1016/j.artint.2012.11.004

Zhai, Q., Peng, R., & Zhuang, J. (2020). Defender–attacker games with asymmetric player utilities. *Risk Analysis*, *40*(2), 408–420. doi: https://doi.org/10.1111/risa.13399

Zhang, J., Wang, Y., & Zhuang, J. (2021). Modeling multi-target defender-attacker games with quantal response attack strategies. *Reliability Engineering & System Safety*, *205*, 107165. doi: https://doi.org/10.1016/j.ress.2020.107165

Zhang, J., Zhuang, J., & Jose, V. (2018). The role of risk preferences in a multi-target defender-attacker resource allocation game. *Reliability Engineering & System Safety*, *169*, 95–104. doi: http://dx.doi.org/10.1016/j.ress.2017.08.002

Zhuang, J., & Bier, V. (2007). Balancing terrorism and natural disasters—Defensive strategy with endogenous attacker effort. *Operations Research*, *55*(5), 976–991. doi: https://doi.org/10.1287/opre.1070.0434

Zhuang, J., & Bier, V. (2011). Secrecy and deception at equilibrium, with applications to anti-terrorism resource allocation. *Defence and Peace Economics*, *22*(1), 43–61. doi: https://doi.org/10.1080/10242694.2010.491668

Zhuang, J., Bier, V., & Alagoz, O. (2010). Modeling secrecy and deception in a multiple-period attacker–defender signaling game. *European Journal of Operational Research*, *203*(2),

409–418. doi: https://doi.org/10.1016/j.ejor.2009.07.028

A Literature

A.1 Related work discrete choice models

Discrete choice modeling was first introduced to the field of criminology by Bernasco and Nieuwbeerta (2005). The researchers presented the discrete spatial choice approach to studying criminal target choice. Using neighborhood and offender characteristics, a conditional logit model was built to investigate how residential burglars select target areas. Parameter estimation of the model was performed with maximum likelihood methods. In a follow-up study, a similar model was adopted by Bernasco (2006) to investigate whether solitary burglars choose their target areas differently from burglar groups. Whereas these two studies used the distance between the homes of burglars and their potential target areas, Clare, Fernandez, and Morgan (2009) evaluated the impact of barriers and connectors on macro-level offending location choices of residential burglars. Furthermore, Bernasco and Block (2009) analyzed how street robbers choose target areas based on data related to distance and population, social barriers, attractions, and guardianship. A random sampling of the data was used to reduce the number of required computing resources.

Random sampling was also used by Bernasco (2010b), who studied the effects of residential history on crime location choices of offenders charged with robbery, burglary, theft from a vehicle, or assault. Demographic, social, and economic information about postal code areas was used to build the model. The author showed that offenders are not solely attracted to target areas close to their current homes but also to the environment of their past homes. Similar results were found in a study by Bernasco and Kooistra (2010), who focused on commercial robberies. While these studies showed that offenders often commit crimes within their current and former residential areas and in areas they previously targeted, Menting, Lammers, Ruiter, and Bernasco (2016) focused on the effects of residential areas of family members on crime location choice. The results showed that residential areas of family members are at increased risk of being targeted.

The previous studies analyzed criminal location decisions using a maximum likelihood estimated conditional logit model with relatively large spatial units of analysis. (Bernasco, 2010a), however, noted that the choice processes involved in crime target selection operate at considerably smaller geographic scales and therefore analyzed the location choices of offenders at more detailed spatial resolutions. Furthermore, they used importance sampling instead of random sampling to estimate the model. The results of this study justify the use of smaller geographic units and emphasize the importance of considering spatial interdependence in the conditional logit model. Bernasco, Johnson, and Ruiter (2015) consequently used smaller spatial units of analysis to build a maximum likelihood estimated conditional logit model to study the effects of past on future burglary locations. Similarly, Vandeviver, Neutens, van Daele, Geurts, and Vander Beken (2015) explored the burglary target selection process at the house level.

Xue and Brown (2006) compared two spatial choice models and a kernel density estimation (KDE) model to analyze and predict the spatial behavior of criminals and latent decision-makers. These spatial choice models included conditional logit models where the observable utility of each spatial alternative is weighted by its evaluation probability. The first model, the uniform spatial choice model, assumes criminals have the same choice sets and preferences. The second model, the distinct spatial choice model, relaxes this assumption. For both spatial choice models, attributes of spatial alternatives came from census data and calculated distance values, and the alternatives were sampled using an importance sampling technique. Results showed that the spatial choice models significantly increased the predictive accuracy of future criminal locations compared to the KDE model.

Smith and Brown (2007) compared a KDE model with two discrete choice models to predict

the spatial preferences of attackers. More specifically, they used residential breaking crime, geographic, and demographic data to compare the efficacy of these three models. The two discrete choice models included a conditional logit and a spatial hierarchy model. The latter model is a modification of the conditional logit in which the utility of each alternative is weighted by the probability that it was evaluated, similar to the spatial choice models of Xue and Brown (2006). For both the discrete choice models, parameters were estimated via maximum likelihood. Results demonstrated that the discrete choice models significantly outperform the KDE model.

Townsley et al. (2015) explored the generalizability of a theoretically derived offender target selection model in a cross-national comparison. This model included a maximum likelihood conditional logit model to estimate the impact of both space- and offender-level factors on residential burglary placement in three regions. They concluded that the effect of spatial factors is significant across all regions, while the effect of offender factors is not.

Bernasco, Ruiter, and Block (2017) examined whether, in street robbery location choices, the importance of location attributes is conditional on the time of the day and day of the week. A conditional logit model was used to estimate the effects of census block attributes on the probability of the census block being selected for robbery. Results indicated that the importance of the attributes hardly depends on the time of the day or day of the week.

While most studies employing the discrete choice approach have focused on burglary or robbery, other crime types have been explored too. Johnson and Summers (2015) used a maximum likelihood estimated conditional model to examine how neighborhood characteristics and proximity to offender home locations affect offender spatial decision-making for vehicle theft. Song et al. (2019) also focused on this crime type. They investigated whether the daily mobility flows of the urban population can explain the target location choices of thieves. Kuralarasan and Bernasco (2022) studied the location choices of snatching offenders. Furthermore, Marchment and Gill (2019) introduced discrete choice modeling to the study of spatial decision-making of terrorists. They built a maximum likelihood estimated conditional logit model using data on offenses and offenders. Results indicated that an increase in distance from the home of a terrorist to a potential attack area decreased the probability of this area being chosen, and areas containing a main road, police station, or military base were more likely to be selected. Schlicher and Lurkin (2024) used a conditional logit model to represent a pickpocket location choice process of a thief. Based on this conditional logit model, the researchers developed a choice-based resource allocation model, helping policymakers reduce the number of pickpocket attempts. Lastly, Baudains, Braithwaite, and Johnson (2013) used the same model to study the decision-making of rioters. They examined how factors of the selected destinations of rioters, the origins of their journeys, and the characteristics of rioters influence their spatial decision-making. Similar studies were conducted by Baudains (2015) and Baudains, Braithwaite, and Johnson (2016), the first suggesting that future research should consider a mixed logit model of discrete choice in which selections of model parameters are correlated over decision-makers.

Townsley, Birks, Ruiter, Bernasco, and White (2016) were the first to contrast a mixed logit model, until then unused in studies of offender location choice, with the widely used conditional logit model. Whereas the latter model assumes all offenders with the same observed characteristics share the same preferences, the former allows for preference variation between offenders having the same observed characteristics. The mixed logit model was estimated with hierarchical Bayes and the conditional logit model with maximum likelihood. Variables in the models included location choice characteristics from census data. The findings demonstrate considerable preference variation between offenders. Therefore, the researchers concluded that the mixed logit model could replace the conditional logit model in studying the target selection of offenders.

The hierarchical Bayes estimated mixed logit model was also adopted by Frith, Johnson,

and Fry (2017), who investigated the role of the street network in spatial decision-making of residential burglars. Graph theory metrics were used to estimate offender awareness, location accessibility, and ambient guardianship. They found that offender awareness and accessibility significantly predict residential burglary location choices. Global ambient guardianship was found to be associated with an increase in burglary risk and vice versa. The researchers also emphasized that the mixed logit model is promising as it accounts for individual taste preference variation across offenders in contrast to the conditional logit model.

Frith (2019) introduced another sophisticated discrete choice model to the study field; the latent class logit model. They compared the conditional logit, mixed logit, and latent class models to analyze the role of guardianship and familiarity in burglary location choices and offender preferences. These are estimated using maximum likelihood, hierarchical Bayes, and expectation maximization, respectively. Random sampling was used for all three models. The results showed that the mixed logit and latent class model significantly fit the data better than the conditional logit model and that the mixed logit and latent class models are (equally) suitable. This highlights the benefits of accounting for preference variation and relaxing other assumptions of the conditional logit model.

Curtis-Ham, Bernasco, Medvedev, and Polaschek (2022) explored methods to overcome computational challenges involved in using large datasets by sampling from alternatives. The researchers built a maximum likelihood estimated conditional logit model and examined non-residential burglary and sexual offenses. They compared importance sampling variants with simple random sampling and found that the first produced more consistent results than the latter and provided considerable computational savings.

A.2 Related work attacker-defender games

Attacker-defender games for resource allocation typically address the allocation of defensive resources among multiple targets by the defender, as well as the strategic attacker's decisions regarding target selection after observing the defensive allocation. Golany, Kaplan, Marmur, and Rothblum (2009) compared optimal resource allocation policies under probabilistic and strategic risk. The first refers to risk resulting from coincidence (or chance), and the latter to risk resulting from actions taken by interested parties. They concluded that for probabilistic risk, the optimal policy focuses on sites where resources yield the highest impact, and for strategic risk, the optimal policy focuses on the most vulnerable sites to decrease the total potential damage level. Furthermore, Bier et al. (2008) showed that the effectiveness of resources significantly influences optimal resource allocation. Others that studied sequential-move attacker-defender games to optimize defensive resource allocation are, for instance, Powell (2009), and Hausken and He (2016).

The most often relaxed assumption in attacker-defender game articles involving resource allocation is the complete information assumption. Bier, Oliveros, and Samuelson (2007); Powell (2007b); Wang and Bier (2011); Nikoofal and Zhuang (2012); Hausken (2014); and Nikoofal and Gümüs (2015), for example, studied a game where the preferences of the attacker are unknown to the defender. Bier et al. (2007) and Bier et al. (2008) concluded that the best option for the defender sometimes is to leave a target undefended and that different measures of target preferences yield other optimal resource allocations, respectively. Nikoofal and Zhuang (2012) modeled the preferences of the attacker as uncertain parameters on bounded intervals to propose a robust equilibrium for the defender. Incomplete information in attacker-defender games for resource allocation purposes can, however, also be of other types. In the study of Powell (2007a), for instance, the target vulnerabilities are only known to the defender. Zhuang and Bier (2011) and Zhuang, Bier, and Alagoz (2010) assumed that only the defender knows his type. Furthermore, Zhuang and Bier (2011) and Feng, Cai, and Chen (2019) considered a game

in which the type of the attacker is unknown to the defender. Hausken (2014) and Zhai, Peng, and Zhuang (2020) modeled a situation where the resources and objectives of the attacker are unknown to the defender, respectively. Jenelius, Westin, and Holmgren (2010), Rothschild, McLay, and Guikema (2012) and McLay, Rothschild, and Guikema (2012) assumed imperfect attacker perception of the strategy of the defender in their models.

Other less often relaxed assumptions are that of perfect rationality and risk neutrality. Yang, Kiekintveld, Ordóñez, Tambe, and John (2013) compared a model of bounded rationality to a model of perfect rationality and found that attackers follow the bounded rationality model more closely. Nikoofal and Gümüs (2015) explored the effect of the degree of rationality of the attacker on the defensive resource allocation decisions in a single-period game. Zhang, Wang, and Zhuang (2021) analyzed defensive resource allocation in a single-period attacker-defender game assuming bounded rationality for the attacker. Rothschild et al. (2012) and McLay et al. (2012) modeled bounded rationality of the attacker through level-k reasoning. Furthermore, Zhuang and Bier (2007) explored the effects of risk preferences on the decisions of the attacker and defender. Payyappalli, Zhuang, and Jose (2017) studied the effect of the risk preferences of the attacker and defender on resource allocation and deterrence levels. Zhang, Zhuang, and Jose (2018) analyzed the impact of the risk preferences and target valuation of the attacker on the defense allocation and concluded that when an attacker becomes more risk-seeking (or risk-averse), the most valuable targets to the attacker receive more (or less) resources. Other examples of authors that relaxed the risk neutrality assumption are Wu, Yan, Peng, and Wu (2020a) and Wu, Yan, Peng, and Wu (2020b).

Some articles specifically focus on secrecy and deception to improve resource allocation strategies by the defender. Secrecy refers to how information on resource allocation should be released to the public, and deception refers to the effort to mislead attackers. Bier et al. (2007) examined a game in which the defender is unaware of the attacker's target preferences. The study found that, at optimality, the defender may leave some targets unprotected. Alternatively, in a study on strategic interactions between an attacker and either centralized or decentralized defenders, Dighe, Zhuang, and Bier (2009) concluded that partial secrecy about the defensive allocation of discrete resources could lead to more cost-effective attack deterrence. This partial secrecy includes disclosing the total amount of defensive resources, but secrecy about which targets are defended. Zhuang and Bier (2011) found that defender secrecy and deception could be strictly preferred in a one-period game in which the defender has private information. Zhuang et al. (2010) uncovered similar findings in a multi-period game. They showed that secrecy could be an equilibrium strategy when the attacker is uncertain about the defensive resource effectiveness and the target valuation. Deception can be an equilibrium strategy when the attacker is unsure about the defensive resources. Nikoofal and Zhuang (2015) found that the defender cannot benefit from the first-mover advantage by exposing defense levels when the difference between the defender and attacker preferences towards targets is high.

Whereas in most attacker-defender game literature, two players are considered, other articles study games with one defender and multiple strategic attackers. Hausken (2011) developed a framework to analyze targets subject to defense by a strategic defender and attack by multiple strategic attackers. These targets can be parallel, in series, interlinked, interdependent, independent, or multi-use. Hausken and Bier (2011) found that a relatively strong attacker among a heterogeneous population of attackers may force the remaining attackers to withdraw from the contest. In such cases, the defender devotes resources toward the strong attacker only. Xu and Zhuang (2019) concluded that a defender tends to allocate too little or too many protection resources to a target when treating multiple heterogeneous attackers as one monolithic attacker. It depends on the conditions, whether the defender allocates too little or too many resources. Casorrán, Fortz, Labbé, and Ordóñez (2019) provided a detailed discussion on MILP formulations to solve attacker-defender games with multiple attackers. None of these

attacker-defender games has considered differences in target preferences among the attackers.

Most of the previous articles considered single-period games. However, a less common approach is to model multi-period games. Levitin and Hausken (2009) considered a three-period game between a defender and an attacker and determined the optimal strategies of the players. Wang and Bier (2011) modeled a two-period game with incomplete information. Wu et al. (2020a) also considered a two-period game. Zhuang et al. (2010) assumed a multi-period game in which the defender selects the defensive and information disclosure strategies in each period, and the attacker chooses his attack strategy after updating his beliefs about the defensive posture. Finally, Shan and Zhuang (2018) also studied a multi-period game, with different scenarios where the defender could be either myopic or long-sighted, and the defense could or could not be carried over to future periods. For each of the four scenarios, they identified the optimal defense allocations.

Many authors of attacker-defender resource allocation games developed custom solution techniques to solve their games, whereas others applied previously developed solution algorithms and heuristics. Levitin and Hausken (2009) used an enumerative algorithm in a multi-period game, and Jenelius et al. (2010) used a path-following algorithm in a single-period game, both for determining the players' optimal strategies. Wang and Bier (2011) and Yang et al. (2013) adopted a nested partition method and a local search method with random restarts, respectively. Shan and Zhuang (2013a), Shan and Zhuang (2013b), Nikoofal and Gümüs (2015), Paulson, Linkov, and Keisler (2016), and Golany, Goldberg, and Rothblum (2017) developed custom algorithms to search for equilibrium defense resource allocations, the algorithm of the latter being based on optimal solution properties. Zhuang et al. (2010) applied dynamic programming in a multi-period attacker-defender resource allocation game. Nikoofal and Zhuang (2012) and McLay et al. (2012) both applied robust linear programming to identify optimal defense resource allocation strategies, the latter being an extension of the study of Rothschild et al. (2012) that used a recursive algorithm to model bounded rationality. Casorrán et al. (2019) compared existing MILP formulations with the tightest known linear relaxations and proposed an improved (i.e., tighter) formulation to solve attacker-defender games with multiple attackers. Shan and Zhuang (2018) used an algorithm based on backward induction to solve a multi-period game. Payyappalli et al. (2017) and Wu et al. (2020a) also used backward induction to solve a single-period and two-period game, respectively, using numerical examples. Finally, Musegaas, Schlicher, and Blok (2022) used a custom algorithm that selects the worst-case optimal strategy from a set of potential optimal defense strategies.

B Stated choice experiments

The purpose of conducting experiments is to determine the independent influence of various variables (i.e., attributes or factors) on an observed outcome. In stated choice studies, this translates into a desire to determine the influence of the design attributes on the choices observed by the sampled respondents. Stated choice studies typically consist of a large number of respondents who are asked to complete a series of choice tasks in which they are asked to select one or more alternatives from a finite set of alternatives. In each task, the alternatives, whether labeled or unlabeled, are typically defined along several different attribute dimensions, each of which is further described by pre-specified levels derived from an underlying experimental design. Typically, each respondent is asked to complete (a subset of) all choice tasks created from the experimental design.

B.1 Creating stated choice experiments

Creating a stated choice experiment involves three primary steps: model specification, experimental design generation, and questionnaire construction. These steps are explained in this subsection, following the Ngene user manual & reference guide published by ChoiceMetrics (2021).

Model specification The first step in creating an experimental design is to determine the objective of the problem being studied. It is important to decide which alternatives and their respective attributes are to be included. Additionally, one must choose a model suitable for the problem. Furthermore, it is required to know the complete specification of the utility functions. Another important decision to be made is whether an attribute is generic across alternatives or alternative-specific. In addition, it is important to decide whether to include interaction effects in the model besides the main effects. Lastly, it is necessary to decide whether to incorporate nonlinear effects. Once the model is entirely specified, the experimental design can be generated.

Experimental design generation Once the model specification is known, the experimental design can be generated. An experimental design explains which choice situations will be presented to respondents in the questionnaire. Several experimental designs are feasible, but the aim is to identify the most optimal among them. To identify the best design, the following choices must be determined.

First, it must be determined whether the design will be labeled or unlabeled. Alternatives with alternative-specific coefficients must be labeled in the experiment if they are included in the model specification. Alternatives that have generic coefficients can go without labeling. Second, it must be decided whether the design should be balanced at the attribute level. An attribute level is a possible value of an attribute. Attribute-level balance requires that each attribute level appears an equal number of times for every attribute. The third step is to determine the number of attribute levels. Increasing the number of levels will increase the number of choice situations. Using different numbers of attribute levels for various attributes can also create a higher number of choice situations. This happens because of attribute level balance. For instance, if three attributes among which one has 2 levels, another 3, and a third 5 are used, the minimum number of choice situations will be 30. 30 is the lowest possible number, and it occurs because it is divisible by 2, 3, and 5. Alternatively, if one uses 2, 4, and 6 levels, then a minimum of only 12 choice situations will be necessary. Fourth, the attribute level ranges must be determined. It is suggested that using a wide range is statistically preferable to using a narrow range. More importantly, the attribute levels must make sense to the respondents. Fifth, the type of design must be selected. There are various types of designs to consider, and each type has its advantages and disadvantages. Sixth, the number of choice situations must be determined. The number of choice situations is bottom-constrained by the maximum number of coefficients (including constants) to be estimated and the number of choice situations needed to ensure balance at the attribute level (if attribute-level balance is maintained in the design). The design type may also constrain the number of choice situations.

Questionnaire construction Once the experimental design has been determined, the actual questionnaire can be constructed. The experimental design needs to be modified to become comprehensible to the respondent. Additionally, the order of choice situations should be randomized for each respondent to eliminate any potential effects of the order on the estimation.

B.2 Design types

When designing a stated choice experiment, several different types of designs can be considered. This subsection will elaborate on the most renowned design types as described in the Ngene user manual & reference guide published by ChoiceMetrics (2021).

B.2.1 Full factorial design

A full factorial design involves all possible distinct choice situations, and in so doing, facilitates the estimation of all possible effects, including main and interaction effects. In practice, the number of alternatives in a full factorial design is often too great.

B.2.2 Fractional factorial design

One way to overcome the problem of full factorial designs is to use fractional factorial designs, which use a subset of choice situations from the full factorial. Several designs fall into this category. In the standard fractional factorial design, choice situations are randomly selected from the full factorial. However, this is not the best way to do so. A better approach is to select choice situations in a structured way so that the best data from the given choice experiment are used to estimate the model.

Orthogonal design The most well-known type of fractional factorial design is the orthogonal design, which attempts to decrease the correlation among attribute levels in choice situations. However, orthogonal designs have limitations and cannot avoid choice situations in which one alternative is clearly preferred over the others (and thus does not provide much information).

An orthogonal design is considered orthogonal if it meets the criteria of attribute-level balance and independent coefficient estimation. This means that the attribute levels for each attribute in the design must be uncorrelated.

Ngene can generate a sequential orthogonal design (syntax: *seq* or *seq2*), where orthogonality applies only within each alternative, or a simultaneous orthogonal design (syntax: *sim*), where orthogonality also applies across alternatives (ChoiceMetrics, 2021). While attribute levels are not orthogonal across alternatives in a sequential orthogonal design, this method typically results in smaller designs in terms of the number of choice situations in the design. If all options have the same dimensions (i.e., the same attributes with the same levels), then *seq* is used. Alternatively, *seq2* necessitates that all alternatives have different dimensions. Lastly, if only some alternatives have the same dimensions, then *sim* is used.

If an orthogonal design has been identified, it may still be too large to present all choice situations to a single respondent. A popular technique, known as blocking, can break down the orthogonal design into smaller designs. Each block is not orthogonal by itself; only the combination of all blocks is orthogonal. The primary purpose of blocking is to ensure that attribute level balance is achieved within each block, thereby preventing respondents from facing only low or high attribute levels for a certain attribute. Blocks are typically determined by adding an uncorrelated column to the design with the number of levels equal to the number of blocks.

Efficient design Another type of fractional factorial design is the so-called efficient design. Instead of simply looking at the correlation between the levels of the attributes, the goal is to find designs that are statistically as efficient as possible in terms of the predicted standard errors of the coefficient estimates. Essentially, these designs attempt to maximize the information obtained from each choice situation. Efficient designs can outperform orthogonal designs, but prior coefficient estimates must be available. An orthogonal design may only be efficient when

prior knowledge of coefficients is lacking, as the design can be improved upon whenever such information is available.

Two main types of priors can be distinguished: informative priors and non-informative priors (Bliemer & Rose, 2014). Informative priors are established based on knowledge gained from a pilot study, literature, or expert judgment, whereas non-informative priors are based on knowledge of the sign of the coefficient (if available). Each of these two types of priors can be set either as a fixed value, referred to as a local prior, or as a probability distribution (e.g., normal, lognormal, etc.), known as a Bayesian prior. Local and Bayesian priors can be combined to create an effective design. Table 16 shows examples of the various types of priors.

	Local	Bayesian
Informative priors	$\beta_k = -0.5$	$\beta_k \sim Normal(-0.5, 0.2)$
	$\beta_k = 0.8$	$\beta_k \sim Normal(0.8, 0.5)$
	$\beta_k = 1.2$	$\beta_k \sim Lognormal(1.2, 0.9)$
Non-informative priors	$\beta_k = 0$	$\beta_k \sim Uniform(-1, 1)$
	$\beta_k = -0.00001$	$\beta_k \sim Uniform(-1, 0)$
	$\beta_k = 0.00001$	$\beta_k \sim Uniform(0, 2)$

Table 16: Types of priors and examples. From: Bliemer and Rose (2014).

An experimental design is considered efficient if it yields data that enables the estimation of coefficients with minimal standard errors. These errors can be predicted by determining the asymptotic variance-covariance (AVC) matrix based on the underlying experiment and some prior information about the coefficient estimates. There are two approaches to calculating the AVC matrix - either through Monte Carlo simulation or analytical methods. For a Monte Carlo simulation, a sample is generated and coefficients are estimated through simulated choices. This process involves computing observed utilities with prior coefficient estimates, adding random draws for the unobserved utilities, and then determining the chosen alternative by assuming that each respondent chooses the alternative with the highest utility. The estimation method produces results for the AVC matrix. This procedure is repeated numerous times, and the AVC matrix is obtained from the average of the results. The second derivative of the log-likelihood function is calculated and evaluated analytically when the AVC matrix is calculated analytically.

The design efficiency can be derived from the AVC matrix of a design. The commonly used measure is the D-error, which assumes a single respondent[13] and is derived from the determinant of the AVC matrix. Another commonly recognized efficiency error in matrix calculations is the A-error, which involves taking the trace of the AVC matrix rather than its determinant. Specifically, the A-error entails summing up all the diagonal elements of the matrix. Consequently, the A-error only considers the variances and not the covariances.

There are two types of algorithms for finding an efficient design: row-based and column-based. In a row-based algorithm, choice situations are selected from a predefined candidate set (either a full factorial or a fractional factorial) in each iteration. At the outset, row-based algorithms can readily eliminate unfavorable choice situations from the candidate set (e.g., by applying a utility balance criterion). However, achieving attribute-level balance is more

[13]The use of a single respondent is merely for convenience and comparison purposes and does not carry any further implications. While any sample size could have been used, it is common in the literature to rely on a single respondent.

challenging. The Modified Federov algorithm (Cook & Nachtsheim, 1980) is a row-based algorithm and includes an iterative process that begins by generating a set of candidate choice situations. This is done by using either a full factorial selection (for small problems) or a fractional factorial selection (for larger problems) from the complete range of possible choice situations for the given problem. The algorithm then selects choice situations from the candidature set, after which the D-error of the design is computed from the AVC matrix. The algorithm saves the design with the lowest D-error as the most efficient design up to that point. The iterative process continues until all possible combinations of choice situations have been evaluated. It is possible to stop the algorithm after a certain amount of time, a certain number of iterations, or a specified amount of time or number of iterations since the latest improvement was found. On the other hand, column-based algorithms create a design by selecting attribute levels over all choice situations for each attribute. Column-based algorithms can easily satisfy attribute-level balance, but finding good combinations of attribute levels in each choice situation is more difficult. In general, column-based algorithms offer more flexibility and can handle larger designs. However, for unlabeled designs and specific designs such as constrained designs, row-based algorithms may be more suitable. RSC (Relabeling, Swapping & Cycling) algorithms (Huber & Zwerina, 1996; Sándor & Wedel, 2001) are column-based. In each iteration, the algorithm creates different columns for each attribute, which together form a design. The design is evaluated, and if it has a lower efficiency error than the current best design, it is stored. The columns are not created randomly but are generated in a structured way using relabeling, swapping, and cycling techniques. Beginning with an initial design, each column could be modified by relabeling the attribute levels. Swapping involves switching some attribute levels. Cycling replaces all attribute levels in each choice situation by replacing the first attribute level with the second level, the second level with the third, and so on. Since this affects all columns, cycling can only be performed if all attributes have the same sets of feasible levels. In some cases, only swapping or relabeling and swapping are used as special cases of this algorithm type.

Although efficient designs typically require fewer choice situations than an orthogonal design, the number of choice situations may still be too large to give to a single respondent. Similar to creating blocks for orthogonal designs, choice situations in an efficient design can be divided based on the minimum correlation principle. In this case, the correlation between the blocking column and all other design columns is minimized on average.

C Ngene

C.1 Efficient design

C.1.1 Syntax

```
? Efficient design
Design
;alts = location1, location2, route1, route2, none
;rows = 10
;block = 2
;eff = (mnl, d)
;alg = mfederov(stop=noimprov(30 mins))
;reject:
location1.x1 >= location2.x1 and location1.x2 >= location2.x2 and
    location1.x3 >= location2.x3 and location1.x4 >= location2.x4 and
    location1.x5 >= location2.x5,
location1.x1 <= location2.x1 and location1.x2 <= location2.x2 and
    location1.x3 <= location2.x3 and location1.x4 <= location2.x4 and
    location1.x5 <= location2.x5,
route1.x6 >= route2.x6 and route1.x7 >= route2.x7 and
    route1.x8 >= route2.x8 and route1.x9 >= route2.x9,
route1.x6 <= route2.x6 and route1.x7 <= route2.x7 and
    route1.x8 <= route2.x8 and route1.x9 <= route2.x9
;model:
U(location1) = a_location + b1[0.001] * x1[0,1] + b2[0.001] * x2[0,1]
                          + b3[0.001] * x3[0,1] + b4[0.001] * x4[0,1]
                          + b5[0.001] * x5[0,1]                         /
U(location2) = a_location + b1        * x1      + b2        * x2
                          + b3        * x3      + b4        * x4
                          + b5        * x5                              /
U(route1)    = a_route    + b6[0.001] * x6[0,1] + b7[0.001] * x7[0,1]
                          + b8[0.001] * x8[0,1] + b9[0.001] * x9[0,1]   /
U(route2)    = a_route    + b6        * x6      + b7        * x7
                          + b8        * x8      + b9        * x9
$
```

C.1.2 Output

Design												
Choice situation	location1.x1	location1.x2	location1.x3	location1.x4	location1.x5	location2.x1	location2.x2	location2.x3	location2.x4	location2.x5		Block
1	1	0	0	0	1	1	1	1	1	0		2
2	1	1	0	1	1	0	1	1	0	0		1
3	0	1	0	0	1	1	0	1	1	1		1
4	1	0	0	0	0	0	1	1	0	1		2
5	1	1	0	0	1	0	0	1	1	0		2
6	1	0	1	0	0	0	1	0	1	1		1
7	0	0	0	0	1	1	1	1	1	0		2
8	0	0	1	0	1	0	1	0	1	0		2
9	1	1	0	1	1	0	0	1	0	0		2
10	1	0	0	1	0	0	1	1	0	1		1
11	1	1	0	0	0	0	0	1	1	1		1
12	1	0	0	0	1	0	0	1	1	0		1
13	1	1	1	1	0	0	0	1	0	1		2
14	1	1	1	0	0	0	0	0	1	1		2
15	0	1	1	0	1	0	0	0	1	0		1
16	1	0	1	1	1	1	1	0	0	0		1
17	0	1	0	1	1	1	0	1	0	0		1
18	0	1	1	1	0	1	0	0	0	0		2
19	1	1	1	0	1	0	0	0	1	0		1
20	1	0	1	1	1	0	1	0	0	0		2

Table 17: Efficient design (part 1).

Design										
Choice situation	route1.x6	route1.x7	route1.x8	route1.x9	route2.x6	route2.x7	route2.x8	route2.x9	Block
1	1	0	1	1	0	1	0	0	2
2	1	1	1	0	0	0	0	1	1
3	1	0	1	1	0	1	0	0	1
4	1	1	1	0	0	0	0	1	2
5	0	1	0	0	1	0	1	1	2
6	1	1	0	1	0	0	1	0	1
7	0	1	1	1	1	0	0	1	2
8	1	1	1	0	0	0	0	1	2
9	1	0	0	0	0	1	1	0	2
10	0	0	1	0	1	1	0	1	1
11	0	1	0	0	1	0	1	1	1
12	0	1	1	1	1	0	0	0	1
13	0	1	1	1	1	0	0	0	2
14	0	0	1	1	1	1	0	0	2
15	0	0	1	0	1	1	0	1	1
16	0	1	1	1	1	0	0	0	1
17	0	0	0	1	1	1	1	0	1
18	0	0	1	0	1	1	0	1	2
19	1	0	1	0	0	1	0	1	1
20	0	1	0	1	1	0	0	0	2

Table 18: Efficient design (part 2).

C.2 Orthogonal design

C.2.1 Syntax

```
? Orthogonal design
Design
;alts = location1, location2, route1, route2, none
;rows = 20
;block = 2
;orth = sim
;model:
U(location1) = a_location + b1 * x1[0,1] + b2 * x2[0,1] + b3 * x3[0,1]
                          + b4 * x4[0,1] + b5 * x5[0,1] /
U(location2) = a_location + b1 * x1      + b2 * x2      + b3 * x3
                          + b4 * x4      + b5 * x5      /
U(route1)    = a_route    + b6 * x6[0,1] + b7 * x7[0,1] + b8 * x8[0,1]
                          + b9 * x9[0,1]                /
U(route2)    = a_route    + b6 * x6      + b7 * x7      + b8 * x8
                          + b9 * x9
$
```

C.2.2 Output

Design												
Choice situation	location1.x1	location1.x2	location1.x3	location1.x4	location1.x5	location2.x1	location2.x2	location2.x3	location2.x4	location2.x5		Block
1	0	0	0	0	0	0	0	0	0	0		1
2	1	1	1	1	1	0	1	0	0	0		2
3	0	1	1	1	0	1	0	0	1	1		2
4	0	1	1	0	1	0	1	1	1	0		1
5	1	1	0	1	0	1	0	1	1	0		1
6	1	0	1	0	1	0	0	1	1	0		2
7	1	1	0	1	0	0	0	1	0	0		2
8	1	0	1	0	0	0	0	0	1	1		2
9	0	1	0	0	0	0	1	1	0	1		2
10	1	0	0	0	0	1	1	0	1	0		1
11	0	0	0	0	1	1	0	1	0	1		2
12	1	0	0	1	1	0	1	0	0	1		1
13	0	0	1	1	0	1	1	0	0	0		2
14	0	1	1	0	1	1	0	0	0	0		1
15	0	1	0	1	1	0	0	0	1	1		1
16	0	0	1	1	0	0	1	1	1	1		1
17	1	1	1	0	0	1	1	1	0	1		1
18	1	1	0	0	1	1	1	0	1	1		2
19	0	0	0	1	1	1	1	1	1	0		2
20	1	0	1	1	1	1	0	1	0	1		1

Table 19: Orthogonal design (part 1).

Design										
Choice situation	route1.x6	route1.x7	route1.x8	route1.x9	route2.x6	route2.x7	route2.x8	route2.x9	Block
1	0	0	0	0	0	0	0	0	1
2	0	0	0	1	1	0	1	0	2
3	0	0	1	1	0	1	0	0	2
4	0	1	1	0	1	1	0	0	1
5	1	1	0	1	1	0	0	0	1
6	1	0	1	1	0	0	0	1	2
7	0	1	1	0	0	1	1	1	2
8	1	1	0	0	1	1	1	0	2
9	1	0	0	1	1	1	0	1	2
10	0	0	1	1	1	1	1	1	1
11	0	1	1	1	1	0	1	0	2
12	1	1	1	1	0	1	0	0	1
13	1	1	1	0	1	0	0	1	2
14	1	1	0	1	0	1	1	1	1
15	1	0	1	0	1	0	1	1	1
16	0	1	0	1	0	0	1	1	1
17	1	0	1	0	0	0	1	0	1
18	0	1	0	0	0	0	0	1	2
19	1	0	0	0	0	1	1	0	2
20	0	0	0	0	1	1	0	1	1

Table 20: Orthogonal design (part 2).

D Questionnaires

D.1 Questions in Dutch

Deel A

1. *Bij welke van de onderstaande activiteiten is de kans het grootst dat er een aanslag op de TBP wordt gepleegd? Ga er vanuit dat er géén beveiligers aanwezig zijn.*
   (a) *Locatie 1*
   (b) *Locatie 2*
   (c) *Route 1*
   (d) *Route 2*
2. *Is de kans groter dat er wel of geen aanslag wordt gepleegd in de bovenstaande situatie?*
   (a) *De kans is groter dat er **wel** een aanslag wordt gepleegd.*
   (b) *De kans is groter dat er **geen** aanslag wordt gepleegd.*

Deel B

1. *Hoeveel beveiligers zijn er minimaal nodig om volledige weerstand te bieden tegen een daadwerkelijke aanslag op een TBP met **lage dreiging**?*
2. *Hoeveel beveiligers zijn er minimaal nodig om volledige weerstand te bieden tegen een daadwerkelijke aanslag op een TBP met a**lgemene dreiging**?*
3. *Hoeveel beveiligers zijn er minimaal nodig om volledige weerstand te bieden tegen een daadwerkelijke aanslag op een TBP met **hoge dreiging**?*

D.2 Responses

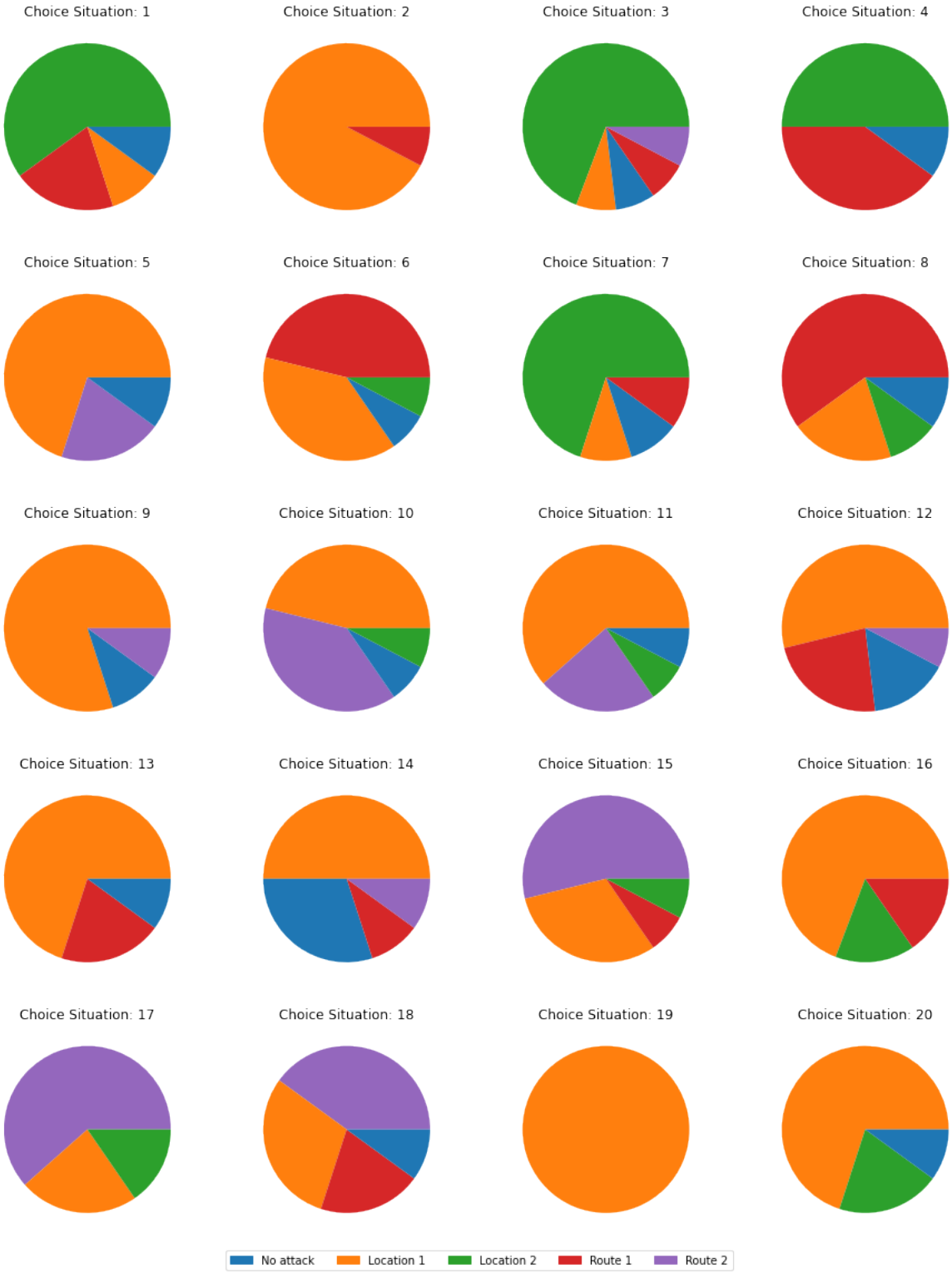


Figure 15: Pie charts of selected choices per choice situation.

E Biogeme

E.1 Python code

```
# Import libraries
import pandas as pd
import biogeme.database as db
import biogeme.biogeme as bio
import biogeme.models as models
from biogeme.expressions import Beta

# Load data
data = pd.read_csv("data.csv")

# Create database
database = db.Database("Conditional logit", data)

# Define variables
globals().update(database.variables)

# Define coefficients to be estimated
ASC_none = Beta('ASC_none', 0, None, None, 1)
ASC_location = Beta('ASC_location', 0, None, None, 0)
ASC_route = Beta('ASC_route', 0, None, None, 0)

B1 = Beta('B1', 0, 0, None, 0)
B2 = Beta('B2', 0, 0, None, 0)
B3 = Beta('B3', 0, 0, None, 0)
B4 = Beta('B4', 0, 0, None, 0)
B5 = Beta('B5', 0, 0, None, 0)
B6 = Beta('B6', 0, 0, None, 0)
B7 = Beta('B7', 0, 0, None, 0)
B8 = Beta('B8', 0, 0, None, 0)
B9 = Beta('B9', 0, 0, None, 0)

# Define utility functions
V_none = ASC_none
V_location1 = (
    ASC_location +
    B1 * location1_x1 +
    B2 * location1_x2 +
    B3 * location1_x3 +
    B4 * location1_x4 +
    B5 * location1_x5)
V_location2 = (
    ASC_location +
    B1 * location2_x1 +
    B2 * location2_x2 +
    B3 * location2_x3 +
    B4 * location2_x4 +
    B5 * location2_x5)
V_route1 = (
```

```
    ASC_route +
    B6 * route1_x6 +
    B7 * route1_x7 +
    B8 * route1_x8 +
    B9 * route1_x9)
V_route2 = (
    ASC_route +
    B6 * route2_x6 +
    B7 * route2_x7 +
    B8 * route2_x8 +
    B9 * route2_x9)

# Associate utility functions with identifiers
V = {0: V_none , 1: V_location1 , 2: V_location2 , 3: V_route1, 4: V_route2}

# Define the availability conditions
av = {0: 1, 1: 1, 2: 1, 3: 1, 4: 1}

# Create logit model
logprob = models.logit(V, av, Choice)

# Define Biogeme model
biogeme = bio.BIOGEME(database, logprob)
biogeme.modelName = "Conditional logit"

# Calculate null log-likelihood
biogeme.calculateNullLoglikelihood(av)

# Estimate coefficients
results = biogeme.estimate()

# Print results
print(results.getEstimatedParameters())
print(results.short_summary())
```

E.2 Output

Estimation report

Number of estimated parameters:	11
Number of free parameters:	9
Sample size:	230
Excluded observations:	0
Null log likelihood:	-370.1707
Init log likelihood:	46
Final log likelihood:	132
Likelihood ratio test for the null model:	1004.341
Rho-square for the null model:	1.36
Rho-square-bar for the null model:	1.33
Likelihood ratio test for the init. model:	172
Rho-square for the init. model:	-1.87
Rho-square-bar for the init. model:	-1.63
Akaike Information Criterion:	-242
Bayesian Information Criterion:	-204.1811
Final gradient norm:	1.3504E-05
Nbr of threads:	8
Relative gradient:	3.8387520779785375e-06
Cause of termination:	Relative gradient = 3.8e-06 <= 6.1e-06
Number of function evaluations:	59
Number of gradient evaluations:	27
Number of hessian evaluations:	26
Algorithm:	Newton with trust region for simple bound constraints
Number of iterations:	58
Proportion of Hessian calculation:	26/26 = 100.0%
Optimization time:	0:00:00.875104

Estimated parameters

Name	Value	Active bound	Rob. Std err	Rob. t-test	Rob. p-value
ASC_location	2.5	0	7.27e-09	3.44e+08	0
ASC_route	-2.5	0	1.8e+308	-1.39e-308	1
B1	81.6	0	2.26	36.1	0
B2	27.6	0	0.562	49.2	0
B3	43.3	0	1.86	23.3	0
B4	26.5	0	0.693	38.2	0
B5	24.2	0	1.19	20.3	0
B6	2.08	0	7.89e-23	2.64e+22	0
B7	2.5	0	7.89e-23	3.17e+22	0
B8	6.51e-17	1	7.89e-23	8.25e+05	0
B9	0	1	4.05e-25	0	1

Figure 16: Biogeme estimation report.

F Success functions

F.1 Python code

```
# Set mean values for different threat levels and a threshold value
#data = pd.read_csv("data.csv")
#mean_low = data.iloc[:, 0].mean()
#mean_general = data.iloc[:, 1].mean()
#mean_high = data.iloc[:, 2].mean()
mean_low, mean_general, mean_high = 1, 2, 3
threshold = 0.01

# Set the number of ticks and create a range for the number of guards
ticks = 6
d_range = np.linspace(0, ticks, 1000)

# Calculate lambda values
lambda_low = - np.log(threshold) / mean_low
lambda_general = - np.log(threshold) / mean_general
lambda_high = - np.log(threshold) / mean_high

# Calculate success probabilities for different threat levels
success_low = np.exp(-lambda_low * d_range)
success_general = np.exp(-lambda_general * d_range)
success_high = np.exp(-lambda_high * d_range)

# Plot the success functions for different threat levels
plt.plot(d_range, success_low,
            label=f'Low threat: e^({-lambda_low:.3f}d)')
plt.plot(d_range, success_general,
            label=f'General threat: e^({-lambda_general:.3f}d)')
plt.plot(d_range, success_high,
            label=f'Hight threat: e^({-lambda_high:.3f}d)')

# Customize and save the plot
plt.legend()
plt.xlim(0, ticks)
plt.xticks(np.arange(ticks+1))
plt.yticks(np.arange(0.0, 1.1, 0.1))
plt.xlabel('Number of guards assigned')
plt.ylabel('Success probability of attack')
plt.title('Success functions')
plt.savefig('success functions.png')
plt.show()

# Create and save DataFrame with the lambda values
lambda_list = {'low': -np.log(threshold) / mean_low,
               'general': -np.log(threshold) / mean_general,
               'high': -np.log(threshold) / mean_high}
df = pd.DataFrame([lambda_list])
df.to_csv('lambda.csv', index = False)
```

F.2 Plots

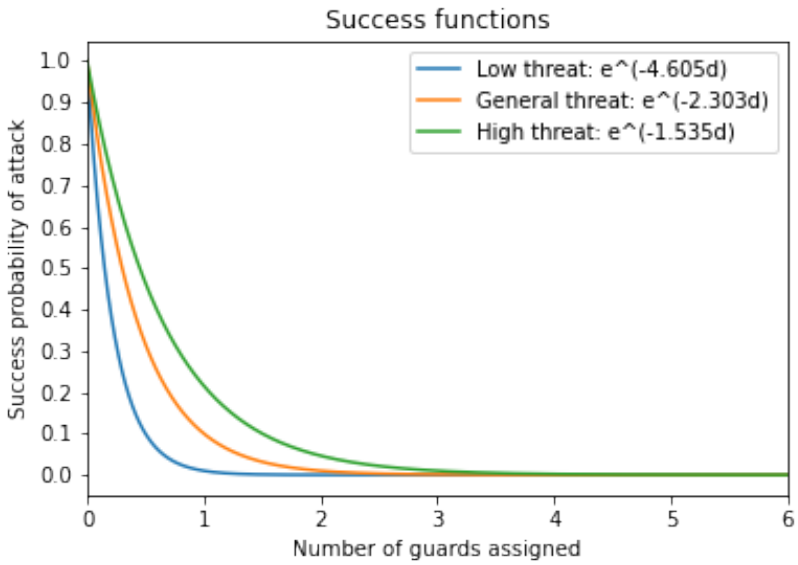


Figure 17: Line plots of success functions for each level of threat.

G Situations generation

G.1 Python code

```
# Import libraries
import pandas as pd
import numpy as np
import random

# Function to generate a random situation
def generate_situation(K):
    # Generate random values for P, v, J_early, and J_late
    P = {k: round(random.uniform(0, 1), 3) for k in K}
    v = {k: round(random.uniform(0, 10), 3) for k in K}
    J_early = {k: random.randint(1, 3) for k in K}
    J_late = {k: random.randint(1, 3) for k in K}

    # Initialize dictionaries for J, P_j, T_j
    J, P_j, T_j = {}, {}, {}

    for k in K:
        # Calculate values for J
        J[k] = J_early[k] + J_late[k]

        # Generate start_times for the activities
        start_early = sorted(np.random.choice(np.arange(7),
                                              J_early[k], replace=False))
        start_late = sorted(np.random.choice(np.arange(7, 14),
                                             J_late[k], replace=False))
        start_times = np.concatenate((start_early, start_late))

        # Generate random values for P_j
        rand_nums = np.random.rand(J[k]+1)
        probabilities = (rand_nums / np.sum(rand_nums)).round(3)

        # Populate dictionaries with generated values
        for j in range(J[k]):
            P_j[k,j] = probabilities[j]
            if j == J_early[k]-1:
                end_time = random.randint(start_times[j], 6)
            elif j == J[k]-1:
                end_time = random.randint(start_times[j], 13)
            else:
                end_time = random.randint(start_times[j], start_times[j+1]-1)
            T_j[k,j] = [i for i in list(range(start_times[j], end_time+1))]

    # Return a dictionary containing the generated values
    return {'P': P, 'v': v, 'T_j': T_j, 'P_j': P_j, 'J': J}

# Generate data for each situation using the generate_situation function
situations = 1000
K = [1, 2, 3]
```

```
data = [generate_situation(K) for _ in range(situations)]

# Create and save DataFrame from the generated data
df = pd.DataFrame(data)
df.to_csv('situations.csv', index=False)
```

G.2 Descriptives

Variable	Mean	Standard deviation	Median	Minimum	Maximum
P_1	0.492	0.285	0.488	0.001	0.999
P_2	0.506	0.288	0.517	0.001	1.000
P_3	0.496	0.291	0.506	0.000	0.998
v_1	4.957	2.861	4.987	0.010	9.998
v_2	5.019	2.853	4.908	0.000	9.980
v_3	5.037	2.891	4.972	0.007	9.980
$\lvert J_1 \rvert$	3.995	1.166	4	2	6
$\lvert J_2 \rvert$	4.028	1.171	4	2	6
$\lvert J_3 \rvert$	3.976	1.138	4	2	6
$P_{1,j}$	0.197	0.122	0.188	0.000	0.855
$P_{2,j}$	0.195	0.122	0.188	0.000	0.880
$P_{3,j}$	0.198	0.122	0.188	0.000	0.961
$\lvert T_{1,j} \rvert$	1.776	1.129	1	1	7
$\lvert T_{2,j} \rvert$	1.783	1.144	1	1	7
$\lvert T_{3,j} \rvert$	1.784	1.148	1	1	7

Table 21: Descriptive statistics of the generated situations.

H Semi-flexible risk-based model

H.1 Python code

```
# Import libraries
import pandas as pd
import ast
import cvxpy as cp
from tqdm import tqdm

# Function to solve the optimization problem
def solve_problem(df):
    objective_values = []
    for i in tqdm(range(len(df))):
        # Set the variables
        P, v, T_j = df['P'][i], df['v'][i], df['T_j'][i]
        P_j, J = df['P_j'][i], df['J'][i]

        # Define the decision variables
        d_j, d_t = {}, {}
        for k in K:
            d_j[k] = cp.Variable(J[k], nonpos=False, integer=True)
            d_t[k] = cp.Variable(sum(len(T[s]) for s in S),
                                 nonpos=False, integer=True)

        # Define the objective function
        objective = cp.Minimize(cp.sum([P[k] * v[k] * cp.sum([P_j[k,j] *
                                f[k](d_j[k][j]) for j in range(J[k])])
                                    for k in K]))

        # Define the constraints
        constraints = []
        for k in K:
            for j in range(J[k]):
                for t in T_j[k,j]:
                    constraints.append(d_j[k][j] == d_t[k][t])
            constraints.append(d_j[k] >= 0)
            constraints.append(d_t[k] >= 0)
            for s in S:
                for t in T[s][1:]:
                    constraints.append(d_t[k][t] == d_t[k][t-1])
        for s in S:
            for t in T[s]:
                constraints.append(sum(d_t[k][t] for k in K) <= B[s])

        # Define and solve the problem
        problem = cp.Problem(objective, constraints)
        problem.solve(solver='MOSEK', warm_start=True)

        # Save the objective value
        objective_values.append(problem.value)
```

```
    return objective_values

# Import lambda values
lambda_data = pd.read_csv('Success function/lambda.csv')
lambdas = {'low': lambda_data['low'][0],
           'general': lambda_data['general'][0],
           'high': lambda_data['high'][0]}

# Define the fixed sets
S = ['early', 'late']
T = {'early': [0, 1, 2, 3, 4, 5, 6], 'late': [7, 8, 9, 10, 11, 12, 13]}

# Set the base scenario
K = [1, 2, 3]
B = {'early': 6, 'late': 6}
f = {1: (lambda d: cp.exp(-lambdas['low']*d)),
     2: (lambda d: cp.exp(-lambdas['general']*d)),
     3: (lambda d: cp.exp(-lambdas['high']*d))}

# Import situations
df = pd.read_csv('Situations/situations.csv')
df = df.applymap(ast.literal_eval)

# Solve the model for each situation
objectives = pd.DataFrame(solve_problem(df), columns=['Objective value'])
objectives.to_csv('Situations/objectives.csv', index = False)
```

I Flexible risk-based model

I.1 Extension for longer travel times

Let τ be the required travel time in hours between all threatened individuals. Furthermore, let $y_t^i \in \mathbb{Z}_+$ be the number of guards that are unassigned for the i'th hour at time $t \in T$, where $i \in \{1, ..., \tau - 1\}$. Let $y_t^\tau \in \mathbb{Z}_+$ be the number of guards that are unassigned for the τ'th hour or longer at time $t \in T$. Then, for $\tau \geq 3$, the extension of the model in Equation 10 is as follows.

$$
\begin{aligned}
\min_{(d_{k,j})_{k\in K, j\in J_k}, (d'_{k,t})_{k\in K, t\in T}} \quad & \textstyle\sum_{k\in K} P_k v_k \sum_{j\in J_k} P_{k,j} f_k(d_{k,j}) & & \\
\text{s.t.} \quad & d_{k,j} = d'_{k,t} & & \forall k \in K, j \in J_k, \\
& & & t \in T_{k,j} \\
& \textstyle\sum_{i=1}^{\tau} y_t^i = B_s - \sum_{k\in K} d'_{k,t} & & \forall t \in T_s, s \in S \\
& y_t^1 = \textstyle\sum_{k\in K} \max(0, d'_{k,t-1} - d'_{k,t}) & & \forall t \in T'_s, s \in S \\
& y_t^i = y_{t-1}^{i-1} & & \forall i \in \{2, ..., \tau - 1\}, \\
& & & t \in T'_s, s \in S \\
& y_t^\tau = y_{t-1}^{\tau-1} + y_{t-1}^\tau & & \\
& \quad - \textstyle\sum_{k\in K} \max(0, d'_{k,t} - d'_{k,t-1}) & & \forall t \in T'_s, s \in S \\
& y_{t-1}^\tau \geq \textstyle\sum_{k\in K} \max(0, d'_{k,t} - d'_{k,t-1}) & & \forall t \in T'_s, s \in S \\
& d_{k,j} \in \mathbb{Z}_+ & & \forall k \in K, j \in J_k \\
& d'_{k,t} \in \mathbb{Z}_+ & & \forall k \in K, t \in T \\
& y_t^i \in \mathbb{Z}_+ & & \forall i \in \{1, ..., \tau\}, t \in T
\end{aligned}
\tag{19}
$$

The constraints in Equation 19 reflect the restrictions on the decision variables. The first constraint ensures that a constant number of guards are assigned to a threatened individual during an activity. The second constraint ensures that the total number of assigned plus unassigned guards equals the number of available guards at any given time. The third constraint ensures that guards are unavailable for the first hour when traveling from one threatened individual to another. The fourth constraint ensures that traveling guards remain unassigned for $\tau - 1$ hours. The fifth constraint ensures that guards who are traveling remain unassigned for at least τ hours, and the number of guards who have been unassigned for at least τ hours is adjusted when guards are newly assigned. The sixth constraint ensures that the total number of newly assigned guards does not exceed the number of guards who have been unassigned for at least τ from the previous period. The seventh, eighth, and ninth constraints ensure that the decision variables are non-negative integers.

If $\tau = 2$, the model in Equation 19 can be used without constraint 4. Moreover, if $\tau = 1$ and thus $y_t = y_t^1 = y_t^\tau$, the model in Equation 19 can be used without constraints 3 to 5 and with a new constraint that sums the right-hand sides of constraints 3 and 5. The resulting new constraint is then $y_t = \sum_{k\in K} \max(0, d'_{k,t-1} - d'_{k,t}) + y_{t-1} - \sum_{k\in K} \max(0, d'_{k,t} - d'_{k,t-1})$. If $(d'_{k,t-1} - d'_{k,t})$ is positive, the first element of this constraint is positive and the third element is 0. Similarly, if $(d'_{k,t-1} - d'_{k,t})$ is negative, the first element of this constraint is 0 and the third element is positive. Therefore, $\sum_{k\in K} \max(0, d'_{k,t-1} - d'_{k,t}) - \sum_{k\in K} \max(0, d'_{k,t} - d'_{k,t-1}) =$

$-\sum_{k\in K}(d'_{k,t}-d'_{k,t-1})$, and the constraint can be simplified to $y_t = y_{t-1} - \sum_{k\in K}(d'_{k,t}-d'_{k,t-1})$, which includes the third constraint in Equation 10.

If $\tau \geq 6$, there will be insufficient time for guards to travel between threatened individuals as the length of a shift includes 7 hours. This results in the model producing equivalent outcomes to the semi-flexible model in Equation 9, which assumes that guards are not permitted to travel during shifts.

I.2 Python code

```
# Import libraries
import pandas as pd
import ast
import cvxpy as cp
from tqdm import tqdm

# Function to solve the optimization problem
def solve_problem(df):
    objective_values = []
    for i in tqdm(range(len(df))):
        # Set the variables
        P, v, T_j = df['P'][i], df['v'][i], df['T_j'][i]
        P_j, J = df['P_j'][i], df['J'][i]

        # Define the decision variables
        d_j, d_t = {}, {}
        for k in K:
            d_j[k] = cp.Variable(J[k], nonpos=False, integer=True)
            d_t[k] = cp.Variable(sum(len(T[s]) for s in S),
                                 nonpos=False, integer=True)
        y = cp.Variable(sum(len(T[s]) for s in S), nonpos=False, integer=True)

        # Define the objective function
        objective = cp.Minimize(cp.sum([P[k] * v[k] * cp.sum([P_j[k,j] *
                                f[k](d_j[k][j]) for j in range(J[k])])
                                    for k in K]))

        # Define the constraints
        constraints = []
        for k in K:
            for j in range(J[k]):
                for t in T_j[k,j]:
                    constraints.append(d_j[k][j] == d_t[k][t])
            constraints.append(d_j[k] >= 0)
            constraints.append(d_t[k] >= 0)
        for s in S:
            for t in T[s]:
                constraints.append(sum(d_t[k][t] for k in K) + y[t] == B[s])
            for t in T[s][1:]:
                constraints.append(y[t] == y[t-1] -
                                   sum(d_t[k][t] - d_t[k][t-1] for k in K))
                constraints.append(sum(cp.maximum(0, d_t[k][t] - d_t[k][t-1])
```

```
                                        for k in K) <= y[t-1])
        constraints.append(y >= 0)

        # Define and solve the problem
        problem = cp.Problem(objective, constraints)
        problem.solve(solver='MOSEK', warm_start=True)

        # Save the objective value
        objective_values.append(problem.value)

    return objective_values

# Import lambda values
lambda_data = pd.read_csv('Success function/lambda.csv')
lambdas = {'low': lambda_data['low'][0],
           'general': lambda_data['general'][0],
           'high': lambda_data['high'][0]}

# Define the fixed sets
S = ['early', 'late']
T = {'early': [0, 1, 2, 3, 4, 5, 6], 'late': [7, 8, 9, 10, 11, 12, 13]}

# Set the base scenario
K = [1, 2, 3]
B = {'early': 6, 'late': 6}
f = {1: (lambda d: cp.exp(-lambdas['low']*d)),
     2: (lambda d: cp.exp(-lambdas['general']*d)),
     3: (lambda d: cp.exp(-lambdas['high']*d))}

# Import situations
df = pd.read_csv('Situations/situations.csv')
df = df.applymap(ast.literal_eval)

# Solve the model for each situation
objectives = pd.DataFrame(solve_problem(df), columns=['Objective value'])
objectives.to_csv('Situations/objectives.csv', index = False)
```

J Attacker-defender game

J.1 Python code

```
# Import libraries
import pandas as pd
import numpy as np
import ast
import cvxpy as cp

# Import situations
df = pd.read_csv(r'Situations\situations.csv')
df = df.applymap(ast.literal_eval)

# Add columns with r values
K = [1, 2, 3]
r_early, r_late = [], []
for i in range(len(df)):
    r_early.append({1:0, 2:0, 3:0})
    r_late.append({1:0, 2:0, 3:0})
    for k in K:
        early, late = 0, 0
        for j in range(df['J'][i][k]):
            if df['T_j'][i][k,j][1] <= 6:
                early += 1
            elif df['T_j'][i][k,j][1] >= 7:
                late += 1
        r_early[i][k] = df['v'][i][k] * sum(df['P_j'][i][k,j]
                                            for j in range(early))
        r_late[i][k] = df['v'][i][k] * sum(df['P_j'][i][k,j]
                                           for j in range(late))
df['r_early'], df['r_late'] = r_early, r_late

# Function to find all possible defender strategies
def find_strategies(total, num_elements):
    if num_elements == 1:
        return [[total]]
    strategies = []
    for i in range(total, -1, -1): # Start from total and go backwards
        for remainder in find_strategies(total - i, num_elements - 1):
            strategy = [i] + remainder
            strategies.append(strategy)
    return strategies

# Function to solve the game
def solve_game(S_A, S_D, F):
    # Define the decision variables
    p = cp.Variable(F.shape[0], nonneg=True)  # Probabilities for the attacker
    q = cp.Variable(F.shape[1], nonneg=True)  # Probabilities for the defender
    v = cp.Variable(nonneg=True)  # Value of the game

    # Define the constraints
```

```
    constraints = []
    for j in range(F.shape[1]):
        constraints.append(p @ F[:, j] >= v)
    constraints.append(cp.sum(p) == 1)
    for i in range(F.shape[0]):
        constraints.append(q @ F[i, :] <= v)
    constraints.append(cp.sum(q) == 1)

    # Define the objective
    objective = cp.Maximize(v)

    # Define the problem and solve it
    prob = cp.Problem(objective, constraints)
    prob.solve()
    return v.value

# Import lambda values
lambda_data = pd.read_csv('Success function/lambda.csv')
lambdas = {'low': lambda_data['low'][0],
           'general': lambda_data['general'][0],
           'high': lambda_data['high'][0]}

# Set the base values
f = {1: (lambda d: np.exp(-lambdas['low']*d)),
     2: (lambda d: np.exp(-lambdas['general']*d)),
     3: (lambda d: np.exp(-lambdas['high']*d))}
S_A = [1, 2, 3]
B = 6
S_D = np.array([[1,2,3]])
F = np.empty((len(S_A), len(S_D)))
v_F = []

# Solve the game for each situation
for x in range(len(df)):
    r = df['r_early'][x]
    for i in range(len(S_A)):
        for j in range(len(S_D)):
            F[i,j] = r[i+1]*f[i+1](S_D[j,i])
    v_F_1 = solve_game(S_A, S_D, F)
    r = df['r_late'][x]
    F = np.empty((len(S_A), len(S_D)))
    for i in range(len(S_A)):
        for j in range(len(S_D)):
            F[i,j] = r[i+1]*f[i+1](S_D[j,i])
    v_F_2 = solve_game(S_A, S_D, F)
    v_F.append(v_F_1 + v_F_2)
```

K Scientific poster

POLITIE TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

Risk-based guard allocation for threatened individuals

M.C.E. Nachtegael

Supervisors: L.P.J. Schlicher, M. Slikker, B. Smeulders, R. Zandbergen

Introduction

Crime and terrorism pose a constant threat to individuals, especially public figures. In the Netherlands, the Bewaken & Beveiligen system offers protection to those who are threatened. However, this system is under significant pressure due to the rising threat from organized crime. As a result, research efforts have emerged to strengthen the system structurally and ensure its future viability. This study aimed to enhance the allocation of guards to threatened individuals by utilizing quantitative methods.

Quantifying risk exposure

To measure the risk exposure of threatened individuals, attacker decision-making was analyzed using discrete choice modeling. A conditional logit model was formulated where the choice set for an attacker represented the activities planned by the threatened individual and the option to refrain from attacking. For each activity type (i.e., locations and routes), binary risk factors were identified and their effects were estimated. The model quantified the risk exposure of an activity by calculating the probability of it being selected by an attacker.

Semi-flexible risk-based allocation

A risk-based semi-flexible allocation model was developed, allowing the number of guards assigned to each threatened individual to vary between shifts. This model included a nonlinear integer program that minimized the total expected damage on a day subject to capacity constraints. The Fine-Kinney method was used to determine the total expected damage.

Flexible risk-based allocation

The allocation model was extended from semi-flexible to flexible, allowing for variation in the number of guards assigned to each threatened individual not only between shifts but also during shifts.

Attacker-defender game

In contrast to the allocation models that focused solely on defensive decision-making, an attacker-defender game was developed to anticipate both offensive and defensive strategic behavior. This involved a single-period, simultaneous-move, two-player, zero-sum attacker-defender game in which the defender assigned guards to threatened individuals and the attacker selected one of those individuals to attack.

Results

Model	Threat levels	Exp. damage decrease	Improved instances
Semi-flexible	Identical	19.6%	63.5%
	Different	21.6%	66.5%
Flexible	Identical	46.3%	99.2%
	Different	45.5%	98.9%
Game	Identical	36.0%	100%
	Different	34.1%	100%

Table 1: Performance of different models.

Conclusions

Risk-based allocation models were proposed based on the quantification of risk exposure. The flexible model outperformed the semi-flexible model when modeling attackers' decisions as exogenous probabilities. In contrast to the semi-flexible model, the flexible model performed slightly better when assuming identical threat levels compared to different threat levels. Alternatively, the game could be used to model attackers' decisions as endogenous probabilities. The game performed slightly better when identical threat levels were assumed, rather than different threat levels.

It is important to note that this study is limited by several assumptions. Therefore, the results should be interpreted with caution. Future research could potentially explore the relaxation of these assumptions.

Department of Industrial Engineering & Innovation Sciences

Figure 18: Scientific poster of thesis.