

ONLINE APPENDIX

for

Modeling Behavioral Complexities of Warning Issuance for Domestic Security: A Simulation

Approach to Develop Public Management Theories

1. Model interface

Figure A1 shows a screen-shot of the model interface. There are five different parameters through which different scenarios can be examined. The interface includes the parameters that represent behavioral characteristics of a society, and behavioral characteristics of adversaries, as well as the frequency of planning attacks, and the probability that adversaries might cancel or post-pone their plans. The interface also includes decision model parameters through which decisions of the security management can be represented. They are private warning threshold, public warning threshold, and accuracy in the intelligence system report.

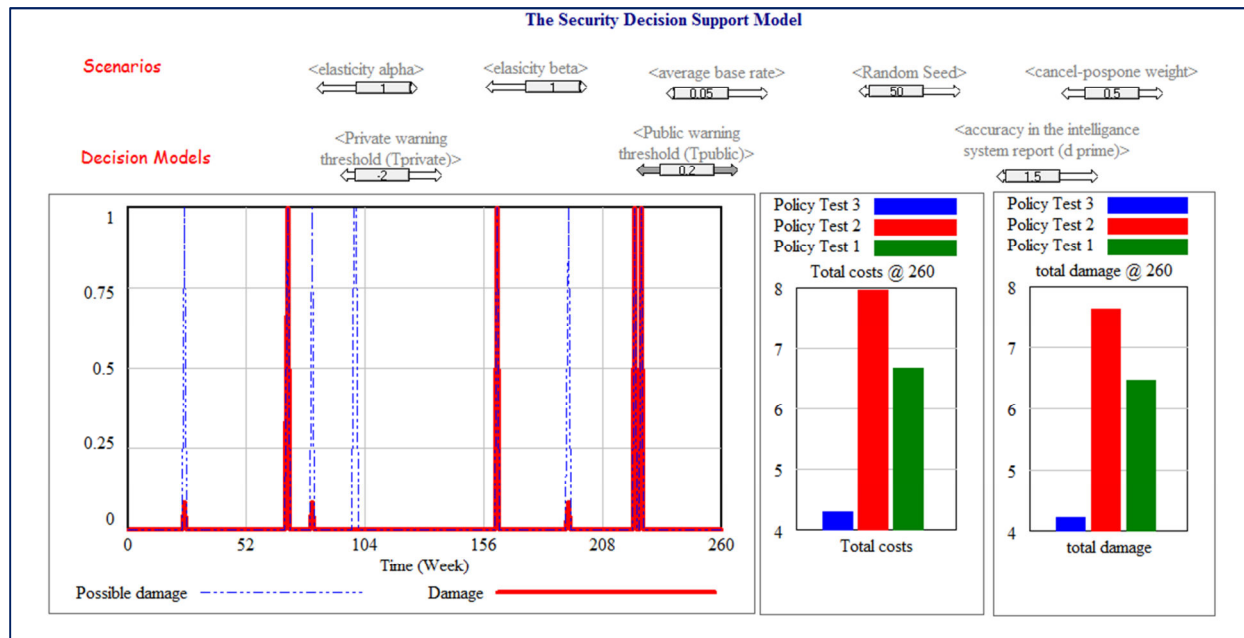
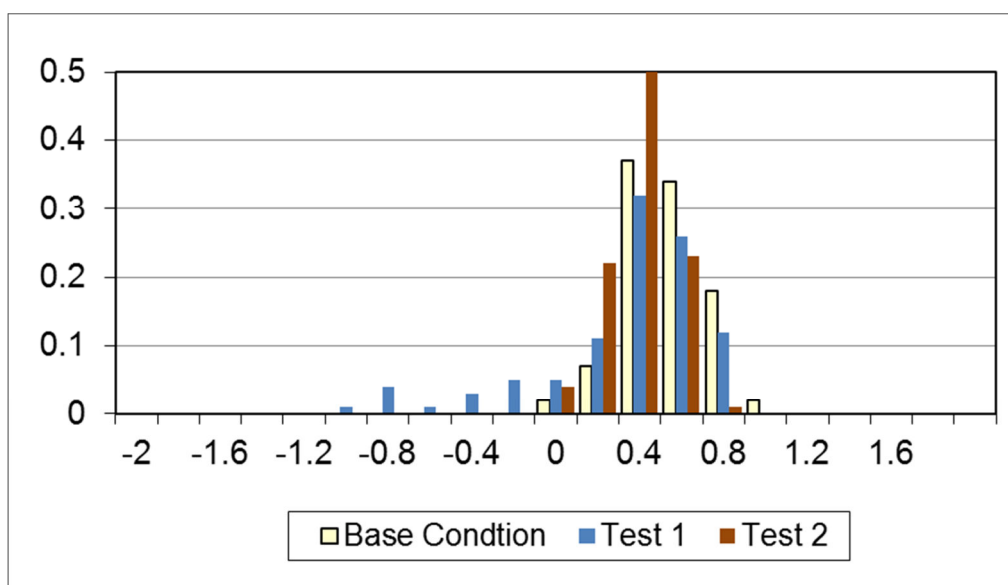


Figure A1: A Screen Shot of the Model Interface

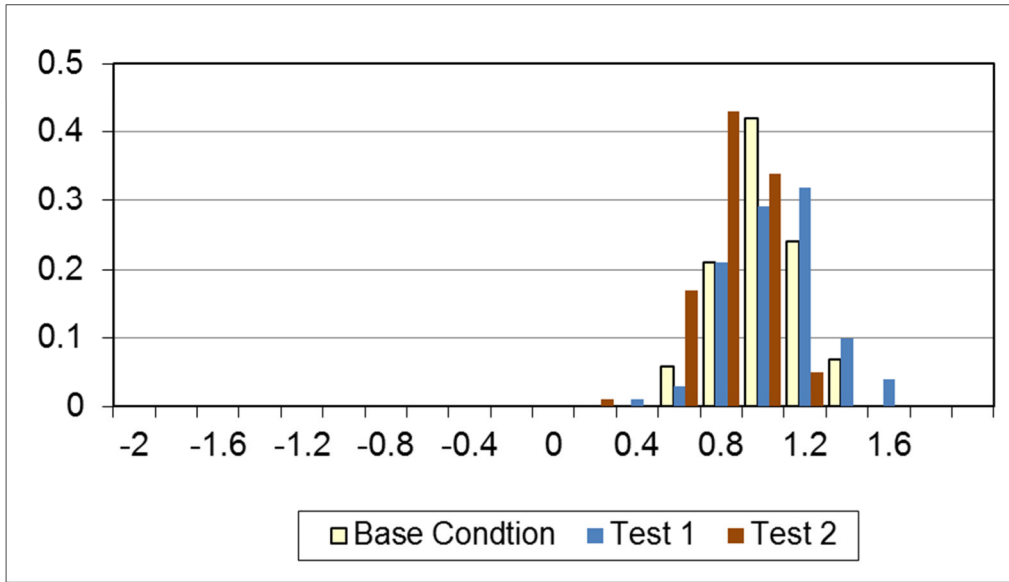
The graphs show the dynamic trend of attacks in a 260-week time period (easily adjustable to a longer time period) and the total damage and costs to the security system, which helps comparing different scenarios and policies.

2. Sensitivity Analysis

We conduct two sets of sensitivity analysis. In the first analysis, we examine the model's robustness for a change in d' , judgmental accuracy. In the second sensitivity analysis, we examine the model's robustness for a change in the base rate, probability of planning an attack. The model proves robustness in both tests, as the optimal results are not significantly different from the base run. Figures A2-a and A2-b demonstrate a sample of results from the sensitivity analysis. Test 1 is for a condition when d' is equal to one third of d' in the base run, and Test 2 is for a condition when the base rate is as twice as the base rate in the base run. The distributions of optimal solutions in these three conditions are very close and statistically not different.



(a) Threshold of private warning



(b) Threshold of public warning

Figure A2: Sensitivity analysis for optimal threshold.

Note: Test 1 examines sensitivity to a change in d' , judgmental accuracy. Test 2 examines sensitivity to a change in the base rate, probability of planning an attack. The tests demonstrate the robustness of the model, that is optimal solutions under these two tests are not statistically different from the optimal solution for the base case.

We would like to re-state that each of our analyses in this paper includes extensive simulation runs with different random seeds, a common practice in a Monte Carlo simulation analysis. In many places, the distribution of the results is demonstrated, showing that our results are not sensitive to random seeds.

3. How to optimize warnings for n possible different targets for adversaries

In the current paper, we model a security system in an aggregate level. For example, if the model is applied to a university context, we assume, when public warnings are issued it is sent to all students, and when private warnings are issued it is sent to all police officers in the university. The model does not include what would happen if criminals changed their target and attacked somebody in the street. The current model also does not disaggregate the target to different potential targets, such as different departments or dormitories. Under the current assumption, we have one target, one type of private warning, and one type of public warning.

Due to a reviewer's suggestion, in this section we lay out how the model can be expanded to account for different possible targets, and different warning audiences, and take the first step with simulating a small model with two targets. This can be considered as a first step toward spatial modeling of the problem, where different geographical characteristics of the environment can be included in the model (such as locations of different departments, offices, dormitories of a university in a security model of a university).

Figure A3 lays out a model structure for $n=6$ possible targets for adversaries. In the macro level, people in each of these targets interact with each other. For example, if this figure represents a university, students of different departments are in contact, and they can go to different areas during a day. The security management team (in this example, a team of the chief of the university police department, the university president, and other decision makers) might send warnings through emails, automatic phone calls, phone messages, or the university websites. These public warnings can be sent to all students, or a specific group of students or staff. Warnings can also be posted on the website, or can be broadcasted through the local media, where the target will be all students, professors, and staff (in the figure, warning to everybody).

Adversaries will choose between different targets, and they might switch to a new target if they become aware of a warning.

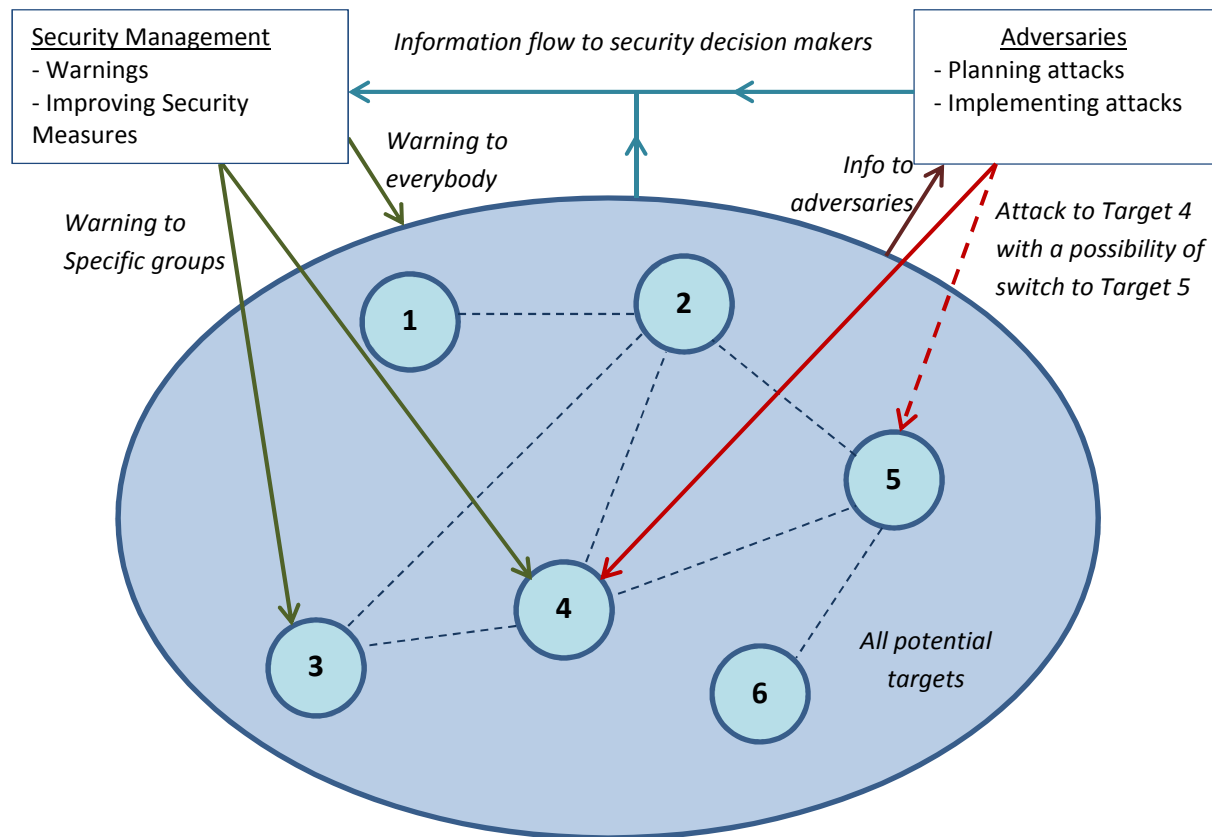


Figure A3: A diagram of a model of warning issuance in existence of multiple potential targets connected through social and infra-structural networks.

Note: In this conceptual model, the security management team issues public and private warnings to a specific audience or to everybody in the system. Adversaries, on the other side, target a specific place while being able to switch to another target. The connections between the targets represent the social and infra-structural networks (e.g., how people are in contact in different places, how the electricity supply system works).

Obviously, formulation of this model is very complicated. For example, it is important to model the social network of public (how people in different possible targets are in contact), the indirect effect of warning issuance to other potential audiences, accuracy of information that comes from different parts of the model, and how the adversaries select between the targets and switch between them.

As a first step, in following, we develop our model to apply it for a context that has two possible targets. An example can be security management in a university that has two campuses that are far enough to require warning strategies, while criminals might switch between the campuses. We call the potential targets campus A and B.

This extended model includes three types of warnings (public warning through email to everybody, private warning to security staff in campus A, and private warning to the security staff in campus B). The model considers the possibility that criminals with a plan to attack to campus A can switch to B, vice versa, due to a change in security measures in campus A, operationalized by guards' level of reactions.

Disaggregating the targets requires adding an additional variable to the system representing the security system's ability to predict which potential target is actually the adversaries' target. Here, we test two main scenarios representing two extremes for target recognition (T.R.): first, the security system in each time period can differentiate which place is more inclined to an attack (represented by Good T. R. in Figures A4 and A5), and the second scenario when the security system is incapable of such a prediction (represented by Poor T. R. in Figures A4 and A5). In the first condition, the security management system will issue warnings to security staff in the specific campus that is more prone to an attack, and in the second condition they issue private warnings to security staff in both campuses.

Similar to how we proceeded in the paper, we first optimize public warning threshold assuming private warning thresholds are equal to public warning threshold minus 1. Figure A4 shows the results from the optimization and compares it with the base run (the situation with one target).

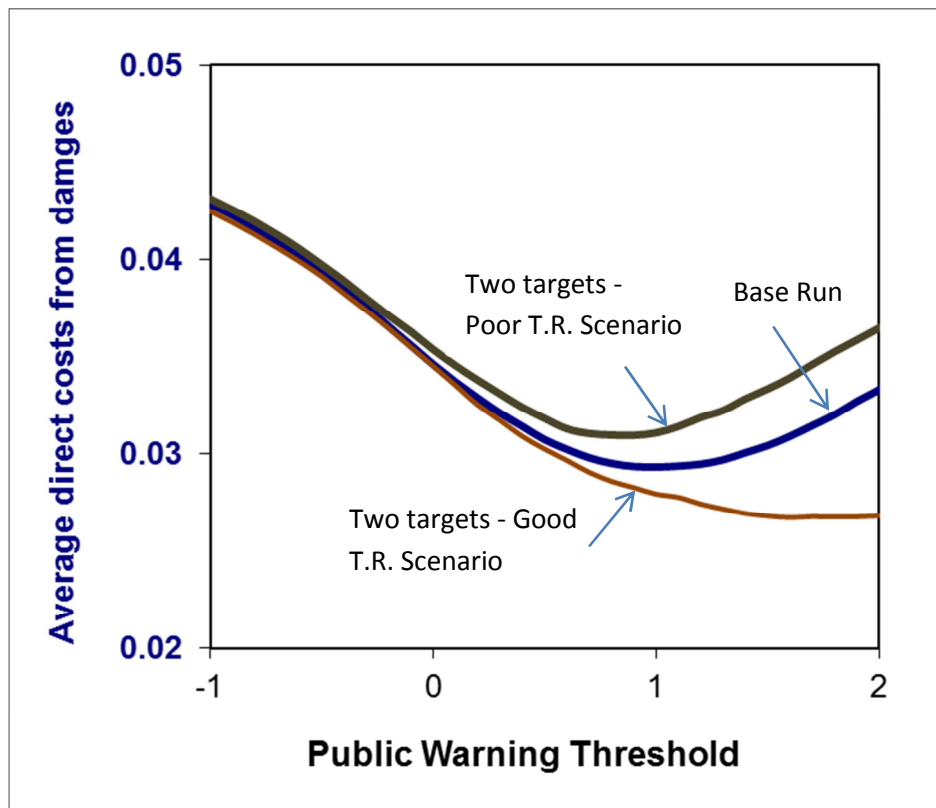


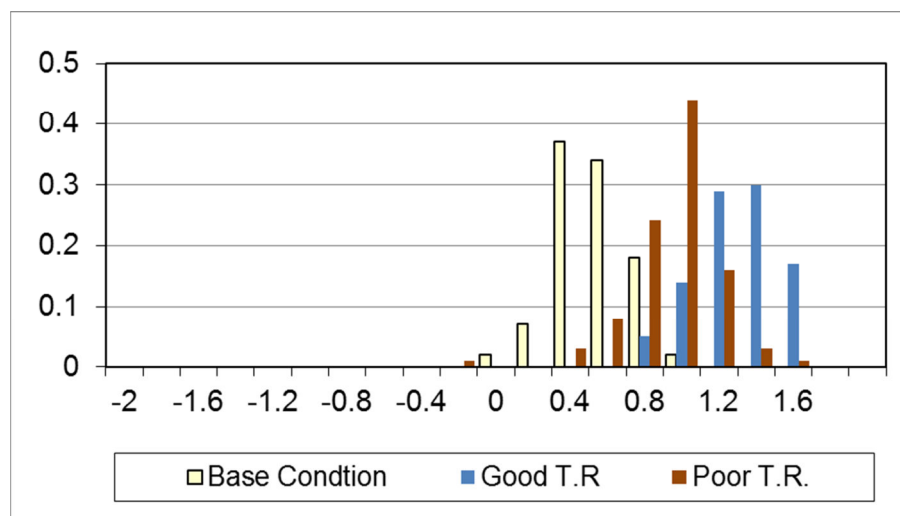
Figure A4: Effect of public warning threshold on damage costs in three different conditions.

Note: T.R. stands for Target Recognition. Good T.R. is the condition when security management is good at predicting which of the potential two targets are in adversaries plan. Poor T.R. is the scenario where the security system cannot easily recognize the target.

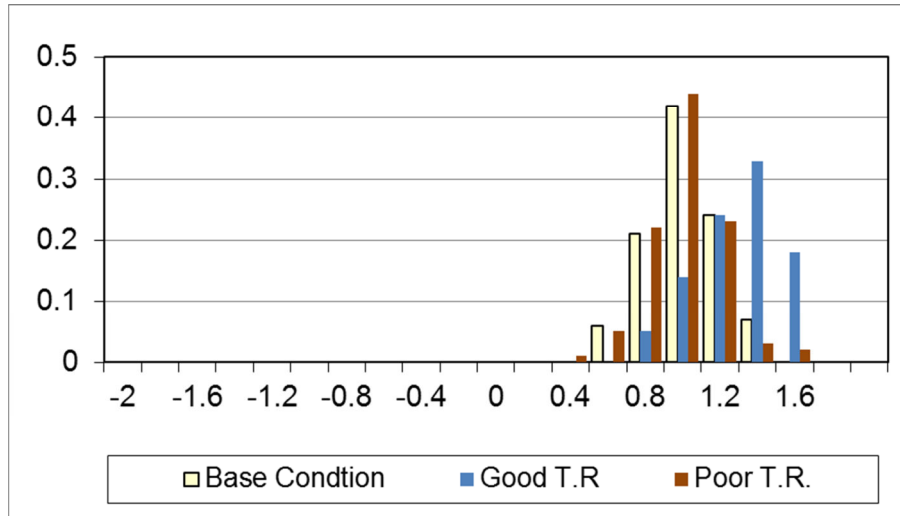
Figure A4 has two major points. First, as we expect, the level of damage depends on how good the security system is in recognizing the targets: under the poor T.R. scenario, we face higher costs than the good T.R. scenario. Second, the optimal solution is completely different

under these conditions (the point where cost is minimum). Under the good T.R. condition optimal threshold is high, i.e., the system suggests much lower frequency of warnings, unless there is a high chance of an attack.

We conduct an additional analysis, and optimize both public and private thresholds. Figure A5 shows the results under both scenarios and compare them with the base run (an aggregate target). The figure shows that the model suggests, overall, higher thresholds for private warning in both scenarios than in the base run. The reason is that in extended model, security staff works in different campuses (targets), and once one campus, say, campus A, is under an attack, staff in campus A get more cautious while staff in campus B thinks that campus A is more likely to be the target. This mechanism, in turn, increases the vulnerability of campus B in comparison to campus A in the next time period. Such a mechanism decreases the effectiveness of warning issuance. While, this lesson is insightful, it is preliminary, needs an empirical examination and consideration of other mechanisms through which loss of sensitivity might be balanced (such as learning the pattern of attacks).



(a) Threshold of private warning



(b) Threshold of public warning

Figure A5: Optimal threshold when there are two potentially different targets.

Note: T.R. stands for Target Recognition. Good T.R. is the condition when security management is good at predicting which one of the potential two targets are in adversaries' plan. Poor T.R. is the scenario where the security system cannot easily recognize the target. The scenario analysis shows how optimal threshold can change if the targets are geographically different enough, and how the optimal policy depends on the security systems' ability of target recognition.

Overall, the goal of this section was to lay out how the model can be further developed for considering multiple targets, and taking the first steps to model when having multiple, geographically and behaviorally different potential targets.