SECONOMICS

D6.4 – A Set of Policy Papers

Document author(s) and Company –

Mathew Collinson – UNIABDN Petra Guasti, Zdenka Mansfeldova – ISASCR Woohyun Shim, Fabio Massacci, Martina de Grammatica, Luca Allodi – UNITN Ugur Turhan – AU Raminder Ruprai – NGRID Julian Williams – UNIDUR Alessandra Tedeschi – DBL

Pending of approval from the Research Executive Agency - EC

Document Number	D6.4
Document Title	A Set of Policy Papers
Version	1.6
Status	Final
Work Package	WP 6
Deliverable Type	Report
Contractual Date of Delivery	31.01.2015
Actual Date of Delivery	31.01.2015
Responsible Unit	UNIDUR
Contributors	UNIABDN, ISASCR, UNITN, AU, NGRID, UNIDUR, DBL
Keyword List	Public Policy, Regulation, Legal Implementation, Legal Instruments, Game Theory, Security Insurance
Dissemination level	PU

SEVENTH FRAMEWORK PROGRAMME

SECONOMICS Consortium

SECONOMICS "Socio-Economics meets Security" (Contract No. 285223) is a Collaborative project) within the 7th Framework Programme, theme SEC-2011.6.4-1 SEC-2011.7.5-2 ICT. The consortium members are:

1	UNIVERSITÀ DEGLI STUDI DI TRENTO	Universite Degli Studi di Trento (UNITN) 38100 Trento, Italy http://www.unitn.it	Project Manager: Prof. Fabio Massacci fabio.massacci@unitn.it
2	DEEPBLUE	DEEP BLUE Srl (DBL) 00193 Roma, Italy http://www.dblue.it	Contact: Alessandra Tedeschi alessandra.tedeschi@dblue.it
3	Fraunhofer ISST	Fraunhofer-Gesellschaft zur Fiorderung der angewandten Forschung e.V., Hansastr. 27c, 80686 Munich, Germany http://www.isst.fraunhofer.de/en/	Contact: Prof. Jan Jurjens jan.juerjens@isst.fraunhofer.de
4	Universidad Rey Juan Carlos	UNIVERSIDAD REY JUAN CARLOS, Calle Tulipileon s/n, 28933, Miostoles (Madrid), Spain. http://www.urjc.es	Contact: Prof. David Rios Insua david.rios@urjc.es
5	1495 UNIVERSITY OF ABERDEEN	THE UNIVERSITY COURT OF THE UNIVERSITY OF ABERDEEN, a Scottish charity (No. SC013683). King's College Regent Walk, AB24 3FX, Aberdeen, United Kingdom http://www.abdn.ac.uk/	Contact: Dr Matthew Collinson matthew.collinson@abdn.ac.uk
6	TMB Transports Metropolitans de Barcelona	FERROCARRIL METROPOLITA DE BARCELONA SA, Carrer 60 Zona Franca, 21-23, 08040, Barcelona, Spain http://www.tmb.cat/ca/home	Contact: Michael Pellot mpellot@tmb.cat
7	AtoS	ATOS ORIGIN SOCIEDAD ANONIMA ESPANOLA, Calle Albarracin, 25, 28037, Madrid, Spain http://es.atos.net/es-es/	Contact: Alicia Garcia Medina alicia.garcia@atos.net
8	SECURENOK	SECURE-NOK AS, Professor Olav Hanssensvei, 7A, 4021, Stavanger , Norway Postadress: P.O. Box 8034, 4068, Stavanger, Norway http://www.securenok.com/	Contact: Siv Houmb sivhoumb@securenok.com
9	SOÚ Institute of Sociology AS CR	INSTITUTE OF SOCIOLOGY OF THE ACADEMY OF SCIENCES OF THE CZECH REPUBLIC PUBLIC RESEARCH INSTITUTION, Jilska 1, 11000, Praha 1, Czech Republic http://www.soc.cas.cz/	Contact: Dr. Zdenka Mansfeldova zdenka.mansfeldova@soc.cas.cz
10	nationalgrid THE POWER OF ACTION	NATIONAL GRID ELECTRICITY TRANSMISSION PLC, The Strand, 1-3, WC2N 5EH, London, United Kingdom http://www.nationalgrid.com/uk/	Contact: Dr. Ruprai Raminder raminder.ruprai@uk.ngrid.com
11	ANADOLU ÜNİVERSİTESİ	ANADOLU UNIVERSITY, SCHOOL OF CIVIL AVIATION Iki Eylul Kampusu, 26470, Eskisehir, Turkey http://www.anadolu.edu.tr/akademik/yo_svlhvc/	Contact: Nalan Ergun nergun@anadolu.edu.tr
12	Durham University	The Palatine Centre, Stockton Road, Durham, DH1 3LE, UK https://www.dur.ac.uk/	Contact: Prof. Julian Williams julian.williams@abdn.ac.uk

Document change record

Version	Date	Status	Author (Unit)	Description
0.1	25/09/2014	Draft	F. Massacci (UNITN)	Draft of Paper 2 completed and submitted.
0.2	1/10/2013	Draft	J. Williams (UDUR)	Draft table of contents and content for Annex 1.
0.3	13/10/2014	Draft	M. Collinson (UNIABDN)	Content of Paper 4 drafted.
0.4	1/11/2014	Draft	J. Williams (UDUR)	Draft of paper 1 submitted to ISASCR for review.
0.5	28/11/2014	Draft	J. Williams, M. Collinson (UDUR)	Draft of Deliverable and Final ToC completed.
0.6	11/12/2014	Draft	Z. Masfeldova, P. Guasti (ISASCR), J. Williams (UDUR)	Completion of paper 1.
0.7	28/12/2014	Draft	J. Williams (UDUR), W. Shim (UNITN)	Paper 5 completed in addition to summaries of papers 3 and 5.
0.9	28/12/2014	Draft	J. Williams (UDUR), M. Collinson (UNIADBN)	Main draft finalized for distribution.
0.95	28/12/2014	Draft	W. Shim (UNITN), M. Collinson (UNIADBN)	Main draft revised for distribution.
1.1	13/01/2015	Draft	A. Caranti, A. Massa (UNITN)	Scientific review, comments and revision on compendium.
1.2	14/01/2015	Draft	E. Lopez (URJC)	Scientific review.
1.3	20/01/2015	Draft	E. Chiarani (UNITN)	Quality Check.
1.4	24/01/2015	Draft	L. Allodi (UNITN)	Contribution to Section 5.
1.5	26/01/2015	Draft	F. Massacci (UNITN)	Final scientific review.
1.6	27/01/2015	Draft	J. Williams (UDUR), W. Shim (UNITN)	Revision reflecting comments from the final scientific review.

Index

Executive Summary

This deliverable presents a comprehensive set of policy papers. They outline both our specific policy recommendations and scientific underpinnings founded in our scientific work and the policy tool. The culmination of the SECONOMICS project is a series of policy recommendations and this is in the form of the policy oriented scientific papers found in this volume.

The SECONOMICS project is founded in the concept of the production of evidence led policy, based on structural models that carefully address the economic incentives of agents and the physical and technological environment that is subject to attack. These models are programmed in the Matlab programming language and translated using various other languages (such as Java) to integrate them into the Toolkit. The purpose of this compendium is threefold:

1. To outline the underlying theories used to build the SECONOMICS Toolkit.

2. To illustrate how these models have been used to build policy in the SECONOMICS context.

3. To demonstrate, in detail, how we have combined qualitative and quantitative evidence into our policy framework.

Short Summary

Each paper provides a summary of work accomplished in this project. The following document provides a detailed overview of each paper. However, we will now provide a very short roadmap to the architecture of the document and a brief executive summary on each paper to place them in context of the overall project. The first two papers provide qualitative insight from salience and case study analysis and motivate the specific quantitative choices for later models. Papers 3, 4 and 5 provide specific quantitative solutions that underpin the mathematical framework behind the user interface of the tool.

- Policy Paper 1: The Political Economy of Security Risk Management provides an overview of how we integrated the salience methodology outlined in Work Package 4, to influence the foundational choices of economic models used in the Toolkit. The paper provides a summary on the public policy implications of security provision and motivates the concept of security as a public good. The key policy insight is in providing the intellectual case for public policy coordination to reduce inequities in cost sharing in security provision. The paper is authored by members of WP4 and WP6.

- Policy Paper 2: Public Policy and Cyber Insurance. This paper provides the most refined version of our general public policy and security model with reactive threats. The major extension of this paper is the quantification of security risks by adding an insurance component to externality model (a marked improvement over the version of the model presented in Work Package 6 Deliverable D6.2). The mathematical architecture of this paper effectively underpins all of the WP6 models and provides guidance on feasible functional forms.

- Policy Paper 3: Agency Problems and Airport Security. This paper provides us with a summary of the qualitative evidence collected from interviews with various airport partners and authorities and motivates the investment functions used in the SECONOMICS Toolkit. The paper also provides basic game theoretic results that show that the interviews are consistent with many of the behavioural concepts outlined in Policy Paper 1. This paper provides policy insight on how contracts for the employment of security need to be structured and the most appropriate usage of training to incentivise the desired security outcomes.

- Policy Paper 4: Public Policy And The Security of Critical Infrastructure: Discretionary or Audit Based Regulation? This paper provides an effective instantiation of the foundational modelling concepts outlined in Policy Paper 3 for the area of Critical national infrastructure. The paper focused on the key policy issues for CNI, rules (with audit) or risk based (with tort driven penalties) regulation, and then quantitatively identifies the concepts of policy assurance audit in a mathematically rigorous setting. The model provides all of the solutions outlined in the CNI section of the SECONOMICS Toolkit and provides some insight into the optimal usage of this policy tool.

- Policy Paper 5: Fairness in Airport Security Expenditures: Equilibrium and Optimum. The final paper in the policy compendium provides the precise instantiation of the mathematical framework underpinning the Airport security model that allows us to identify between different security taxation regimes for various scales of airports across Europe. The paper illustrates how the foundational mathematical models from Policy Paper 2 can be instantiated into the quantitative outcomes presented in the tool and provides insight in how to utilize the tool for the design of such cost sharing systems.

1. The Political Economy of Security Risk Management

The primary focus of the results of the SECONOMICS project is on providing insight and guidance for public policy in respect to security. Public policy is implanted by legal instruments that are designed to produce social welfare maximizing outcomes for all citizens. This can be achieved by information sharing and advice as well as legally binding social and business contracts.

Designing legal structures is however an extremely complex task. In the lay-parlance, the law of unintended consequences may result in well-intentioned legislative instruments creating more issues that they solve. The economic approach to this issue is one of mechanism design or how to build constraints on behaviour that achieve socially optimal outcomes. Game theoretic approaches are generally used to tackle this problem and this deliverable will review our approaches to modelling and optimizing social policy in respect to our security scenarios and finally applying them in more general contexts.

The modern approaches in public economics to the design of legal frameworks dates back to Cournot in 1820s through to Stackelberg in the 1930s. First, set out the basic economic problem in absence of a legal framework and derive some form of equilibrium behaviour. In our context this is usually considered to be individual investments in a vector of security attributes. Next, formulate a measure of global social welfare, for instance you can aggregate the welfare of each individual participant as measured by some form of equilibrium pay-off or utility. Now postulate a series of legal instruments with their various parameters, such as clauses in contracts and varying levels of fines for non-compliance. Derive the optimal legal structure and trade off the improvement in global social welfare with the cost of implementing the optimal instrument. If the instrument is able to improve upon the base case, then this is the best approach.

However, this approach only works if the major drivers of behaviour are carefully modelled in the correct manner. Over the past 50 years game theory has developed from the raw rational expectations approaches that dominated the work of Cournot and Stackelberg. Bayes-Nash equilibria provide a framework for modelling the equilibrium when participants learn iteratively. We can design sequential models that allow agents to react strategically to changes in the legislative environment and we can more accurately model the directions of externalities transferring costs across groups of agents in economic environments.

At present, the use of game theory in security has focused on strategic games of attack and defence usually by two players (an attacker and a defender). However, the consequences of choices in these simple games do appear to have significant wider implications for other agents (attacker or defender). As such public policy games need to nest the individual interactions and statistically control for the wider cross interactions between games.

The deliverable is formed from both non-technical summary papers and technical papers in the public economics mould. In this introduction we provide a brief summary of the policy papers and how they have been influential in the development of the SECONOMICS Toolkit (in WP8).

1.1 Outline of the contribution to policy

The security of citizens has been one of, if not the most important, facets of public policy provision since the dawn of modern civilization. The need for society as a whole to engage in coordinated action to protect individuals from threats, both physical and economic; to their welfare has been a generally accepted principle throughout this time. However, the delegation and coordination of security related activities to a limited number of authorities is, of course, not without cost. The very nature of security enhancing activities lends themselves to the collection and storage of information by a centralized party. However, the political economy of the drivers to the trade-offs discussed previously are often presented most succinctly in the language of economics and game theory. It is to this aspect of the discourse that this chapter is primarily directed.

In contrast, the economics of privacy is a relatively understudied area in general (cf. for an early research on the topic Posner 1980). However, in certain areas, such as medical record disclosure and in certain countries such as the USA, the high level damages awarded for privacy breaches have driven a broad academic and industrial research agenda (see Acquisti 2004, Acquisti et al 2013 for a modern perspective). From a public policy view point, most advanced economies have extensive privacy legislation designed to protect individual citizens from both deliberate and accidental disclosure of confidential information. In the European framework, litigation is increasingly broad to the European Court of Human Rights, under articles 5 and 8 (Guasti, Stockemer, Siroky forthcoming).

The paper first briefly reviews the core economic concepts of utility theory, the adaptability of this concept to non-monetary consumption and the concept of consumption of security in §(2). The paper then moves on the concept of security as a pubic good, the 'production of security' and the non-excludable and non-rivalrous nature of certain aspects of security, §(3). From here having now outlined the basic treatments of security as a public good, we can then address the problems of production of a secure environment that achieves the outcome of society via a utilitarian social planner. This is covered in §(4) of the paper.

We now move from the first and second welfare theorems to cases when information clearing is inefficient and the preceding arguments are fundamentally based on fully informed citizens and social planners and provide the first best outcomes for the security of society as a whole. However, relaxing these assumptions results in social welfare outcomes in the absence of a planner (at a Nash or 'rational-expectations' equilibrium) substantially away from the first best outcomes. Indeed social planners are usually not fully informed and sometimes not benevolent and we will look at a series of examples and outcomes for these cases in §(5).

We will then introduce a further actor in the economic system, firms with capital owners. We will show that the presence of private markets in the security domain can have counter-intuitive effects. In§(6) we illustrate the effect of 'risk neutral' firms on aggregate cyber criminal behaviour. We then look at risk averse corporate officers in §(7) and introduce the notion of an insurance contract. We then speculate on the impact that delegation of security coordination by the social planner will have on risk-averse firms and whether this will lead us towards our first best outcomes. Our final discussion section §(8) will look at the economics of citizens' choices and collective structures in security and the reasoning behind their inception and we will discuss the area of 'institutional analysis and design' in the last part of the commentary and concludes.

1.2 Contribution to the tool

Policy paper 1 does not have a specific mathematical setting, however the ideas that are outlined here permeate the full gambit of every policy case covered in the Toolkit, inculcating the model from the perspectives of the unifying ideas in WP4 and 6. Furthermore, the concept of security provision as being influenced by economic agents (attackers, targets and policy-makers) is additionally a foundational idea behind the work accomplished in WP5.

Specifically, in this paper we outline how security is fundamentally a public good that needs to be produced by the collective of potential targets. Lack of coordination in the production of a public good, creates externalities that lead to inefficient cost allocations. This concept underpins the mathematical framework for all WP6 models. It is also how we translate the conceptual ideas founded in WP4 into the mathematical framework. Salience analysis does not directly lend itself to simulation based forecasting as it is inherently a hind-casting procedure, that very usefully sets current cultural frames. However, we can use Salience analysis to understand very specific features of pubic preferences and risk perceptions (this will be more carefully analysed in policy paper 3).

For instance, we know from WP4 deliverables D4.3 and D4.4 that two key themes permeate public perceptions and risk tolerance: plurality of media and experience within the country of significant security incidents. We can then convey these concepts into reasonable choices of utility functions for social planners and targets. In the SECONOMICS tool we take considerable care in modelling time and risk preferences to ensure that the public policy outcomes documented in the salience analysis correspond (in part) to those found in the WP6 analysis. We know that negative security events motivate governments to invest and this can be described as a mechanism of Bayesian updating. This is carefully analysed in Policy Paper 2 and again in Policy Paper 4 that look at training and security issues in airports and the design of policy assurance and subsidy mechanisms in CNI.

1.3 Policy Summary

- Security appears to have many properties that partially admit it to being a public good.
- Production of this good needs to be coordinated to ensure that appropriate and fair cost sharing amongst agents occurs.
- Coordination directives such as the EU NIS Directive that is currently under discussion address many of the information sharing and coordination requirements that would naturally reduce information asymmetries, however costly production of the public good aspect remains.
- If we look at organized and mandated security regulation, such as the NERC framework, the requirement to commit costly and audited investment in security is seen as a mechanism to reduce total costs to society as a whole.
- The basic theory on externalities that lead to inefficient cost sharing suggests that in almost all conceivable circumstances, the presence of strategic attackers means that firms can also strategically shift defensive costs, there appears to be no way of avoiding regulatory intervention in most security contexts.

2. Public Policy and Cyber Insurance

From the previous paper we see that the role of public policy, in a security context as well as all others, is in providing a mechanisms for overcoming externalities created by adverse selection and moral hazard issues. Both of these economic phenomena contribute to unequal cost sharing in the provision of security. In part the public policy maker has to choose the most appropriate regulatory mechanism to instantiate the public policy requirement from the underlying externalities. This is a very difficult step, in most economic settings the need for policy is complicated by inherent information asymmetries and the collection of data to overcome these information asymmetries can create costs that are greater than those overcome by the regulatory mechanism.

One of the core issues is assurance and its financial analogue insurance. We utilize Cyber Insurance as the object of study in this paper as it adds several contemporary aspects to the modelling problem. First, reactive attackers, strategic behaviour by attackers is foundational to the SECONOMICS modelling framework across all of the scientific Work Packages. Second, the concept of assurance and contingent valuation (fundamental to insurance) also permeates through all of the policy papers herein. Finally, the concept of risk aversion and risk neutral valuations are mathematical tools used in each of our models.

2.1 Outline of the contribution to policy

The impact of moral hazard and adverse selection in the presence of insurance has a long history of investigation in economics, see for instance [1, 2, 3, 4, 5] and [6] for an eclectic set of examples that directly relate to the notions of public policy, liability sharing and insurance considered herein. In contrast, the impact of the presence of insurance and how the collective behaviour of victims of crimes can influence the aggregate behaviour of criminals generating the risks that are being insured, has far lower profile in the literature, although [4] provide a materially similar treatment to our cyberinsurance case. We can think of these effects as moral hazard and adverse selection effects that are once removed from the actions of the insuree, i.e. not the direct influence of target behaviour on target risk, with and without insurance, ceteris paribus, but the impact that changes in the aggregate behaviour of the pool of externally-insured or self-insured have on the risk vectors generating the distribution of losses. Furthermore, adjustments in external environmental conditions have been shown to affect the distribution of insurance claims and their legitimacy, see for instance [7] for auto insurance fraud. Our attacker externality shares several similarities with the incentive to defraud suggested in [7], in that there is a systematic factor that varies across all targets. However, in our case this is fully endogenous, rather than driven by an external macroeconomic effect.

The background risk of being a victim of burglary will, in-the-main, be a function of security choices of the target and background exogenous factors such as the location of the property and regional and national crime trends. The latter components of the risk model are materially unaffected by the influence of insurer and insuree actions in a way that affects the individual and aggregate behavior of the criminals generating the risks, for the current period of coverage for a standard insurance contract, usually one year.

The role of adjusting aggregate behavior by mandating behavior on the insuree is mostly the remit of the public policy maker. For instance, the policy maker can collect higher taxes to invest in higher levels of physical security and increase the costs for criminals so as to dissuade them from choosing to engage in criminal behavior that results in insurance losses.

From the perspective of fraud and theft activities against firms, the focus of this paper, cybercrime has a number of different characteristics to more traditional forms of crime. For instance, the choice of a software engineer to work on either malicious software or software with a more legitimate business purpose is simply a matter of re-tasking code. Most cyber criminals are anonymous (indeed this is the name of a cyber criminal group), therefore the decision to work as a either a 'white-hat' or a 'black-hat' is simply a cost benefit analysis that assesses the opportunity costs, risk of detection and time investment between these two roles. There is recent evidence to suggest that the decision of a hacker to enter into criminal activity is fragile. [8] illustrates that much of the online crime is based on spatial opportunity. Spatial in the cyber crime sense is in terms of ease of access to particular systems and opportunity to illicitly monetize that access for personal benefit.

2.2 Critical Policy Comments

The paper outlines a general modelling treatment of Cyber Security with strategic attackers and illustrates many of the pitfalls in implementing security policy using insurance companies as proxies. Critically, we show that any form of insurance intermediary does not have the incentive to reduce overall attacking effort as this reduces the economic activity in the area of security for which the insurance company or provider of assurance such as a security vendor extract rents.

We illustrate carefully, that our results in the following models are *not* simply artefacts of the functional forms, but are intrinsically related to the preferences of the individual agents and that the need for proper coordination (for instance the NIS directive) inescapable.

2.3 Contribution to the tool

- The mathematical framework outlined in Propositions 1 to 4 of the paper provide all of the WP6 models with their basic motivation.

- The use of risk aversion and utility to model the certainty equivalence valuation of security risks is implemented using specific functional forms in all of the deployed models in the tool (WP5 and WP6 models inclusive).

- The motivating factor of the externalities created by underinvestment cross sectionally is foundational to each of the WP6 models for the Airport and CNI case studies.

- The results Theorem 2 provide us with a robustness check on the model assumptions deployed in the WP6 tools. We generally model preferences with hyperbolic utility and discounting, the results in this paper provide us with assurance that changes in the functional form do not severely impact the magnitude of the results (e.g. switching to polynomial preferences and moment based risks).

3. Agency Problems and Airport Security: Quantitative and Qualitative Evidence on the Impact of Security Training

Policy paper 3 is concerned with agency problems and the role of security training in civil aviation. In civil aviation, airport security and the potential for the systemic failure of airport security has been a central policy question. Moreover, the appropriate training of security personnel is commonly seen as an important policy instrument. Our models try to explore the issues related to agency problems and security training.

The quantitative economic model used in the Principal-Agent (P-A) model found in Policy paper 3 that illustrates a situation where one party (i.e., agent) makes a decision on behalf of another party (i.e., principal). A stated goal of airport security provision usually outlined by the government agencies entrusted with this task is managing risk. The implementation of policies on the ground, however, is performed by staff that is usually paid at or below the national average for their respective countries.

3.1 Outline of the contribution to policy

In the model, the focus is given to a P-A game where the players are both on the security provision side. We consider the principal as a government agency and the agent as a worker conducting security on the principal's behalf. As for the agent, we consider both the police who are hired by the government, and security staff (e.g., security guard and X-Ray screener) who are hired by an airport to meet the goals of the government (hereinafter, referred to as "the employee").

In order to model the interaction between the government and the employee, we consider that the employee needs to comply with various security rules to avoid any penalty, but his action to comply with these rules is costly to him: he is adverse to taking action. However, since the state of the world is uncertain, the employee's effort is not perfectly correlated with the outcome of his effort, e.g., sometimes, even if the employee makes high effort on his security work, there can be a security breach. On the other hand, the contract should not be based on the outcome, since an attack on an airport is very rare. If the contract is based on the outcome (e.g., no security breach), security staff might not need to exert any effort. As shown in Figure 1, however, if the employee's action can be fully observable costlessly, the government only needs to pay the employee for his action that can guarantee his participation. As a result, the employee gets a fixed wage based on his action.

In reality, however, the employee's action is commonly unverifiable and unobservable, and hence his action is not contractable. Therefore, while the government wants to maintain more than a certain level of security, the employee will shirk his responsibilities if he can do this without being discovered and if the expected net gains from shirking are higher than those from exerting due care. In this case, as shown in Figure 2, the government need to provide them with incentives to make them exert due effort.

The general result from the model is that, when the employee's action is unverifiable, the employee will not carry out any action if an incentive wage is not provided. Unlike the case with full information, unverifiability of the employee's action makes the government provide incentive wages that reduce the total surplus of the participating parties, or develop a mechanism that can motivate the employee to exert his due effort.

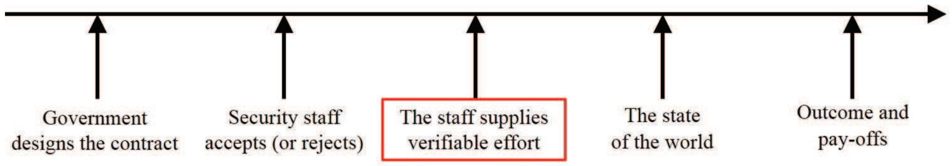

Government designs the contract
Security staff accepts (or rejects)
The staff supplies verifiable effort
The state of the world
Outcome and pay-offs

Figure 1: P-A relationship when the agent's effort can be verified costlessly.

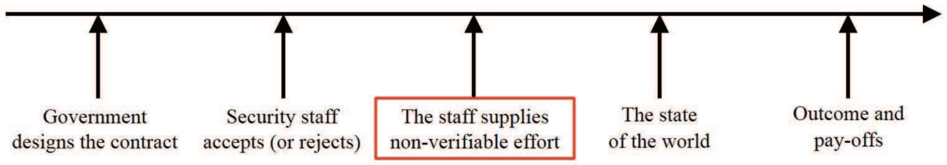

Government designs the contract
Security staff accepts (or rejects)
The staff supplies non-verifiable effort
The state of the world
Outcome and pay-offs

Figure 2: P-A relationship when the agent's effort is not verifiable.

In order to overcome this problem, the model in Paper 3 is extended to show how intrinsic factors motivating the employee affect his action when the action cannot be observable and verifiable, particularly focusing on security training. According to various scholars [9, 10, 11, 12], the employee's rewards might be influenced not only by direct monetary payoffs but also by intrinsic preferences such as job satisfaction and peer recognition. For example, as shown in Figure 3 Huselid et al. [13] argue that employee education and training might be able to increase the employee's intrinsic motivation, thereby raise his effort level and reduce a moral hazard issue. From the model, it is identified that, if security training has a transferable value for the employee, by raising future employability or providing certification and evidence of effort to become a qualified trainee, it is more effective in making the employee exert his due effort.

In order to validate the results of the model, a series of interviews has been conducted with key aviation stakeholders. The individual interviews are designed to explore a potential contradiction between what the employees need to mitigate risk and the risk management mechanisms that have been implemented. The results of our study specifically indicate that shared values exist in the management chain. However, the effectiveness of risk mitigation measures (e.g., managerial tools and extra training) might be undermined by poor incentive structures and asymmetric pay-offs and liabilities.

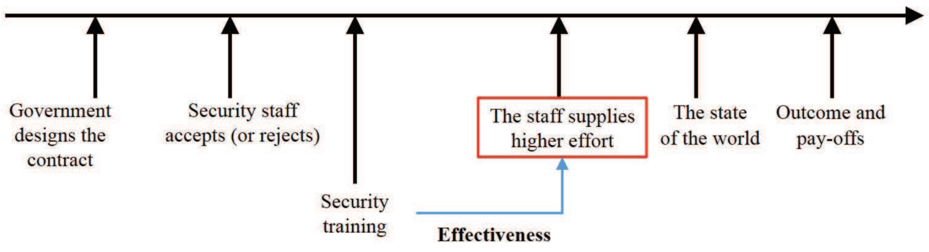

Government designs the contract
Security staff accepts (or rejects)
Security training
The staff supplies higher effort
Effectiveness
The state of the world
Outcome and pay-offs

Figure 3: P-A relationship when training is introduced.

In a situation where an employee's action is not observable, a typical P-A explanation says that a principal will try to design a contract that can make an employee bear costs caused by shirking. In airport security, however, this might not be feasible. A risk related to a terrorist event has extremely high impact on the society with very low probability of occurrence [14]. Once a terrorist event occurs, the participants might not be able to pay the damage, and hence the security risk cannot be transferred to the participant. Together with imperfect observation, therefore, this will make a contract produce a suboptimal outcome.

Some previous studies have indicated that motivating employees by increasing their intrinsic preferences [9, 10, 12] can improve the gap in optimal effort perceived between the principal and the agent. For security staff, motivation is essentially conducted through security training [13]. The argument being, that intrinsic motivation can be a mechanism for reducing the P–A effort gap.

In the interview process, training programs which use two different types of approaches are identified: 'strategic' and 'technical'. Training using a strategic approach aims at providing efficiency that ensure the achievement of a firm's general business objectives (henceforth, referred to as 'general training'), while training with a technical approach focuses on shaping a wide range of technical and professional practices (hereinafter, referred to as 'technical training').

The interview results show that general training programs might only incur a burden on the employees and will not provide the employees with the recognition of their role in ensuring airport security. Since general training does not provide a specific certification to a qualified trainee, it does not provide any information on the employee's repute and not increase his level of employability. Consequently, it is identified that general security training might not be helpful to increase employees' motivation and make them exert due effort.

On the other hand, the interview results indicate that specific technical training is deemed to be very effective in motivating trainees and in attaining skills while these cause higher burden on the trainees than a general training program. The mandatory renewal of employee's certification from technical training and the possible loss of the job due to the failure during this renewal process provide a degree of transferable value which entails a higher level of an employee's effort. A core conclusion appears to be that specific technical training can develop the employees' motivation and understanding of the rationale behind their tasks, hence mitigate a moral hazard problem. However transferability of value from effort appears to be a important factor in the employees pay-off function.

There is a significant body of literature on *ex-post* failings due to an agency problem in complex socio technical systems (in relation to both security events and accidents) and financial services. A study, such as this one, seeks to identify P–A issues *a-priori* to help reduce the likelihood of catastrophic security failures by illustrating to the policy maker the type of risk structure that they are faced with.

3.2 Policy Summary

- Principal-agent problems can exaggerate security risks when liability is unequally shared. The principal is a security manager or agency operating on behalf of society. The agent is the security practitioner directly engaged in identifying and mitigating security threats.

- Agents, although risk averse, exhibit heterogeneous beliefs in the individual outcomes in terms of maximising expected utility.
- Traditionally, training is theorised to provide two sources of benefits for principals: it enables the agents to be more skilled and hence more effective, and it provides agents with a greater understanding of the "importance" of their work, we call this buy-in additional "intrinsic motivation".
- However, we have identified another area, bargaining power, that provides incentives for agents to engage more fully in the activities for which the principals have contracted these agents. We show that if the training appears to lead to ex-post benefits (such as changing job or promotion) then the engagement with the programme appears to be more fully realized.
- The theoretical prediction appears to ties-up very closely with the observations made in the interviews.

3.3 Contribution to the tool

- The following models, need to be founded in real economic relationships between those agents engaged in security, attackers and policy makers.
- The basic theoretical framework is designed to map closely to things we can measure, qualitatively, though interviews with the subject matter experts.
- The key driver here is motivating and validating carefully the underlying assumptions of diminishing marginal returns to security investment and the economic motivations behind actions within a security setting.
- Whilst not directly modelled in the tool, this procedure motivates the functional forms used in Paper 5 that drives how we model the whole arline sector in Europe and identify optimal choices in investments.

4. Public Policy And The Security of Critical Infrastructure: Discretionary or Audit Based Regulation?

This is the primary paper used to build the WP6 CNI section of the SECONOMICS Tool. The paper contains all of the mathematical derivations needed to build the CNI model.

4.1 Policy Insights

The protection of *Critical Infrastructure* (CI) and the continuity and reliability of the *Critical Services* it provides is an essential task for any modern national government or supranational governmental body. Electricity transmission is an excellent example of a critical service. It has been the subject of study for both Work Packages 2 and 6 of the SECONOMICS project.

In many countries, bulk electricity transmission is now provided by private sector Transmission System Operator (TSO) firms. As private sector entities, their primary duty is to their owners. The interests of the owners may not always naturally and automatically align perfectly with the interests of society and citizens. Governments have therefore sought to regulate firms' behaviours in order to protect the interests of society, specifically to ensure adequate provision of the critical service and stewardship of the critical infrastructure. This regulation is typically mandated as part of the contract structure under which the TSO operates the infrastructure. The cost efficient provision of public services and electricity transmission services have been a subject of intense study for economists in Europe and the United States of America for the last thirty years.

A key aspect of electricity transmission is that the service must be reliable. Governments typically place specific contractual constraints on this reliability. Often the top-level reliability concerns are framed in terms of engineering measures of transmission output across the transmission grid. The economics of effective regulation for reliability of transmission has been the subject of a good deal of study by prominent economists over the last fifteen years.

Electricity transmission is increasingly dependent upon computer technology, computer networks and cyber-infrastructure. The changing landscape of technology provides both opportunities and challenges for TSOs and governments. One trend is toward control systems with hardware and software components based on modifications of general components (for example hardened variants of common operating systems), i.e., not completely bespoke technology. Another trend is toward connection of systems, including control systems, to the internet. On the one-hand, this offers the possibility of cost efficiency savings that can be passed on to consumers. However, it potentially introduces additional risks, not only to cyber-assets and operations that form part of the critical infrastructure, but to the reliable provision of service. Some of these are driven by natural processes, but deliberate, remote exploitation by a variety of threat actors is possible, as are accidental or incidental risks caused by non-specific cyber-threats present in the general environment. Thus there are adversarial cyber-security engineering challenges as well as traditional engineering challenges and simpler adversarial physical protection challenges.

Although TSOs naturally seek to avoid any direct damage to themselves it is not clear that, without behavioural constraints, they will take appropriate security measures to mitigate risks faced by society from disruption to service. A laissez-faire approach by policy-maker's is not appropriate in the present context. Policy-makers in many countries have thus sought to regulate to ensure that appropriate security measure are taken by their resident TSOs. Sometimes this has been done by amending the contractual requirements for reliability to include requirements for cyber-security protection.

There are two dominant paradigms in terms of the regulatory approach for cyber-security in this area. There is considerable disagreement in policy and industrial circles as to which is better for society.

The first regulatory approach is commonly referred to as *risk-based*. It is a decentralised approach in which the TSO is given the responsibility and the flexibility to judge, mitigate and report risk to which it is exposed via cyber-assets. Negligence can be punished via civil legal suits (torts) or possibly because of knock-on effects to breaches of higher-level reliability requirements. The security investment made by the firm is purely at its own discretion.

The second regulatory approach is commonly referred to as *rules-based*. In this approach, the regulator provides a mandatory schedule of security controls, measures and rules with which the TSO must comply. Verification of compliance is done via an external audit. A TSO can be punished for non-compliance in line with a pre-arranged schedule of punishments.

On the one-hand, the risk-based approach is seen to have the advantages of decentralization: the TSO can most efficiently invest to mitigate risk, and in theory it should understand its on costs better; moreover, this approach provides for flexibility and agility in the rapidly-evolving cyber-threat landscape. On the other-hand, policy-maker's have been concerned that firms lack maturity, awareness and understanding of cyber-threats and may not be taking appropriate security steps, and a rules-based system gives a definite form assurance to the policy-maker that at least some basic protections are in place. Of course, this is not the only way to provide assurance and risk-based approaches may integrate processes for this (as in the UK). A purely rules-based approach may also have the unintended counter-productive effect of producing TSOs that merely blindly follow rules to the letter and do not take adequate steps to mitigate broad cyber-risk.

SECONOMICS has given access to a unique opportunity in the form of National Grid, a private sector TSO that operates in countries with both of the above types of regulatory regime. In the United Kingdom, it operates the bulk electricity network (apart from parts of Scotland and Northern Ireland) where a risk-based system is in operation. National Grid also operates in the north east of the United States of America, which has probably the most sophisticated and technically detailed rules-based system in the world.

A key issue that policy-makers and TSOs would like to understand is, in which context rules-based systems are better than risk-based systems, and vice-versa. Moreover, they would like to know whether systems that combine the benefits of both approaches are possible and effective. Finally, an issue that is always present in the regulation of private provision of public services is to ensure that an appropriate subsidy (also known as a transfer) is allocated to the private sector entities by society. In electricity transmission this is often done via setting the rates which TSOs can charge. The subsidy must be large enough to allow the TSO to operate securely and in accordance with the regulatory mechanism, but beyond that it should be minimised to provide the most cost efficient service to society.

There has so far been relatively little academic work on understanding the public economics of security for critical infrastructure in the presence of adversarial threats, and very little indeed on the cyber-security of electricity transmission systems. The present work makes contributions to the academic literature and formulates models that could be used to inform practical discussions of future policy in the area. These ideas and models have been extensively discussed and validated with our industry colleagues from National Grid.

4.2 Contribution to the Tool: A Model of Subsidies and Security Incentives in Electricity Transmission and Critical Infrastructure

This paper provides all of the core mathematical modelling for the second CNI model in the tool, this includes derivations of the individual reactions functions and the solution space.

Specifically, we have constructed a model that captures subsidies and also risk and rules-based regulatory regimes. The approach that we take is economic. The regulatory regimes can be viewed as incentives for security behaviour carried out by TSOs. Regulatory mechanisms set incentives for CI firms: non-compliance with these carries the possibility of (direct or indirect) financial consequences for private CI firms. The intention of the regulatory mechanism is that anticipation of such consequences should drive ‘good’ behaviour of the firm, from the point of society. Specifically, a good policy should lead the firm to providing good security to the critical infrastructure.

The adversarial nature of security and the rapidly-changing nature of technology means that it is simply not possible, at least at the present time, to collect threat data and build statistical models that would give a definite scientific evidence base for answering the essential policy questions above. In this work, we have built structural models of the situation in the tradition of mathematical economics. The methodology is game-theoretic and takes into account the preferences of the key actors (policy-maker, TSO firms and attackers) over the outcomes resulting from the combination of choices of regulatory mechanism, TSO security investments and attacking effort, respectively.

In our model we represent a policy abstractly as a strength of incentive for rules compliance, a strength of incentive for risk-based security behaviour and a level of subsidy. This leads, via the game-theoretic analysis to a ‘phase diagram’ that shows how the firm and attacker’s investment levels respond to policy. A stylized version of this diagram is shown in Figure 4. This shows that there are policies under which there is no security investment, only rules-based investment, only risk-based investment, mixed rules and risk investment, and also a combination of pure rules investment and no attacking effort.

The reactions of the TSO firm and attacker can be used to figure out the resulting payoff for the policy-maker. One way to visualize this is as in Figure 5, where a surface is plotted over the phase diagram. A policy consists of a point in the phase diagram and a transfer. The subsidy determines a set of feasible policies in the phase diagram. Over each point in the phase diagram we can plot the resulting payoff. The policy-maker prefers policies with large payoffs, that is, high points on the surface. However, only those points in the feasible set are considered. Here, the payoffs of the feasible policies are separated from the payoffs of the infeasible policies by a dashed line projected onto the payoff surface.

4.3 Policy Insights

Our models do not provide a simple answer to which is ‘better’ between risk-based or rules-based regulation. Rather, they suggest that a careful framing of the policy-maker’s prefer-

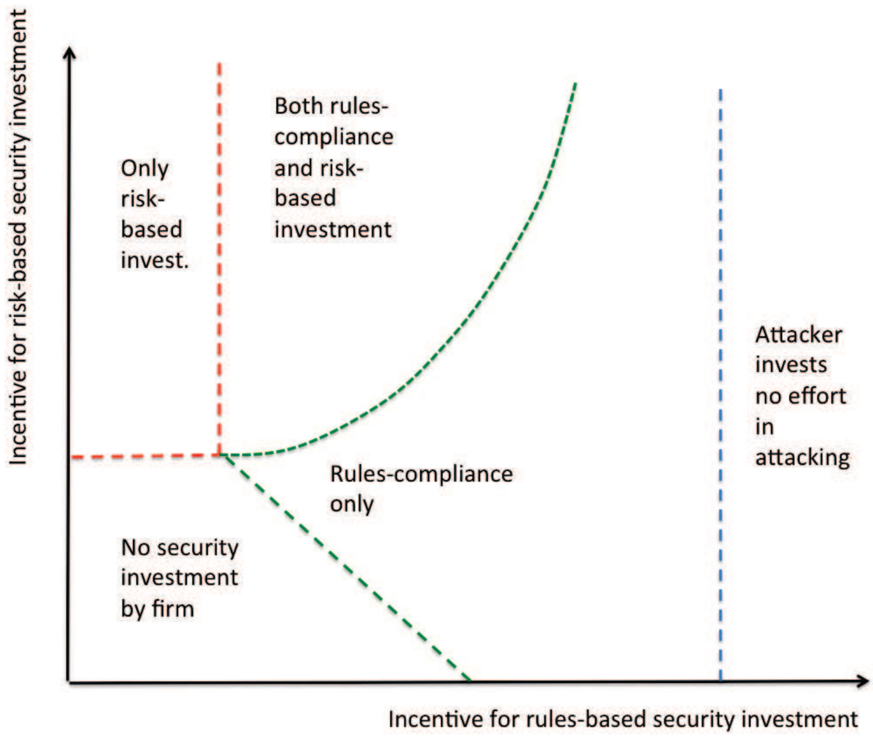

Incentive for risk-based security investment
Only risk-based invest.
Both rules-compliance and risk-based investment
Attacker invests no effort in attacking
Rules-compliance only
No security investment by firm
Incentive for rules-based security investment

Figure 4: The Basic Phase Diagram

ences is required and that the answer is sensitive to the problem context, for example the efficiencies of mitigation from rules and risk investment are key. Some of these parameters are hard to estimate, for example, parameters regarding the attributes of attackers. Nevertheless, our models provide qualitative insight into the key questions.

Our models indicate that treating regulation of reliability including security as a two actor interaction between policy-maker and TSO will lead to different policy recommendations than treating it as a three actor interaction that additionally includes rational adversaries with the ability to vary their attacks in response to policy and in anticipation of the behaviour of the TSO. For example, in the three actor treatment a strong system of rules can drive the TSOs risk-based investment to zero, where this does not happen in the two-actor treatment. It is therefore important that policy discussions do not forget the adversarial nature of security engineering problems, even where much of traditional reliability engineering and regulatory economics would tend to emphasize the two actor case.

From our model we see that unindemnified damage to the TSO arising from adverse security events has an important role. For example, it is instrumental in controlling whether or not a mixed rules and risk-ased investment pattern by the TSO will emerge from a purely rules-based regulatory regime. This reinforces and illustrates the view that a mature understanding and awareness of cyber-threat on behalf of TSOs is highly desirable.

The model suggests that asymmetries of information play a significant role in this problem. For example, the firm will often know better than the policy-maker the costs and efficiencies of security controls. Moreover, the attacker will know its costs and rewards better than the TSO or policy-maker. Such asymmetries lead to uncertainty about the results of

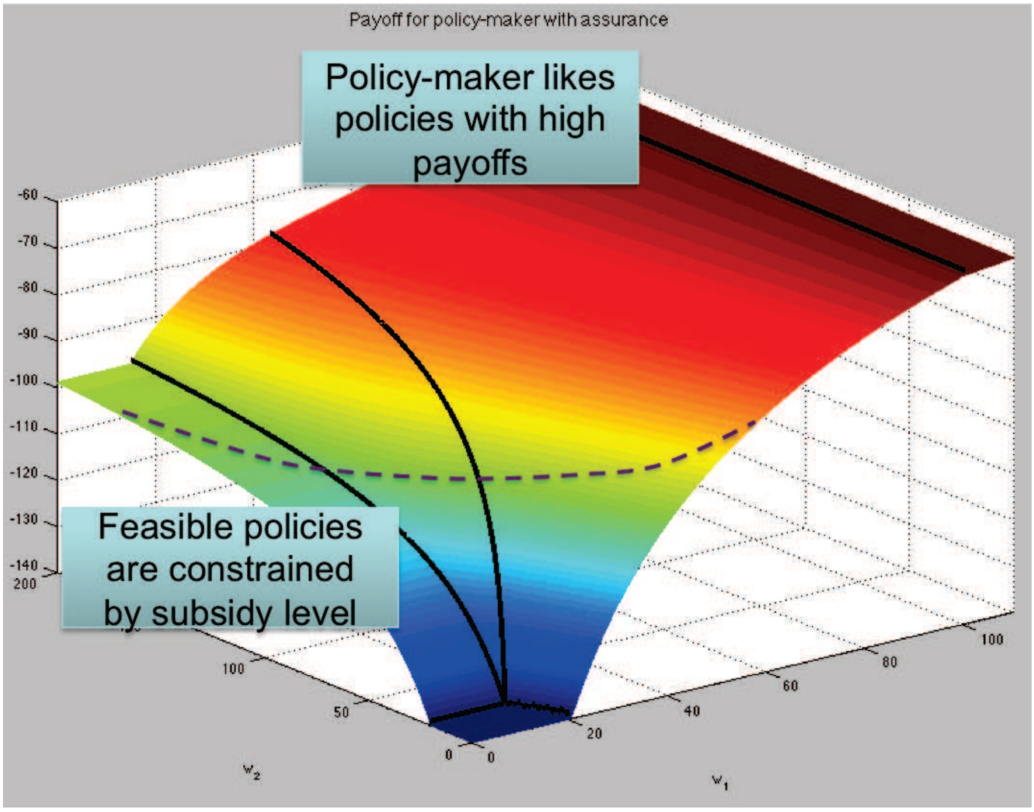


Figure 5: The Policy-maker's Payoff

policy, since exact values of the required parameters cannot be found. However, further to that, information asymmetry creates incentives for the actors to not necessarily represent themselves accurately to others. For example, under a purely risk-based regulatory regime it can be advantageous for the TSO to under-represent its own security efficiency in order to extract a larger subsidy. This might be expected to lead the policy-maker to mistrust claims for subsidy made by the TSO and therefore to allocate less. Such information effects have been widely studied in economics, and deserve to be further studied in such settings with the presence of rational adversaries.

5. Fairness in Airport Security Expenditures: Equilibrium and Optimum

This paper contains all of the required mathematical derivations for the WP6 Airport tool.

5.1 Outline of the contribution to policy

Following the September 11 attacks, various security strategies and regulations have been enacted to increase the airport security level. While these activities might increase overall security level, we presently do not have an in-depth discussion on whether they are effective and efficient. In WP6, we conduct various studies in order to directly address this issue. In our studies, we mainly employ a game-theoretical model to investigate strategic interactions among the players in the security ecosystem. The following sections offer the summaries of the two studies for airport security.

Policy Paper 5 focuses on the issues related to fair allocation of security expenditures in civil aviation of Europe. After September 11, 2001, security costs represent up to 35% of overall airport operating costs [15], and airport operators need to decide the best mechanism for the resource allocation in compliance with regulatory standards. As for a policy-maker, determining the optimal level of security expenditures has become a major task.

As airports' security expenditures are directed by the regulators, however, various questions regarding the fairness of these mandatory expenditures have arisen. For example, some authors have recently pointed out that the optimal security expenditures are likely to vary across airports (e.g., [16]), and each airport might have different security preference. In a series of interviews, we also found that the regulators' passions for making a sound security environment by mandatory expenditures do not align well with the interest of airport operators. The airport operators seem to think that mandatory security expenditures that are set uniformly might not align well with the airports' incentives since security activities required by heterogeneous airports are different.

This issue might be amplified when we consider emerging cyber-threats. As the successful deployment of SESAR and NextGen operational concepts will result in major integration of IT systems and major use of ATM services where the IT part becomes safety critical (e.g., the SWIM system and the Remote and Virtual Tower operational concept in SESAR), major IT interdependencies may lead to major cyber-security risks. At the same time, as indicated in the study of the cybercrime markets done by [17], cybercrime markets have evolved from unruly and disorganized market mechanisms to mature and regulated mechanisms, and have become more efficient in transactions.

In order to tackle this issue, the model from the paper and deployed in the tool employs a game-theoretic approach and investigate whether airport security investments mandated by the regulators are determined fairly for airports with different characteristics. In detail, the models demonstrate how a regulatory rule on security investments might undermine fairness in the context of airport security and how the rule might cause a divergence between unregulated private actions and those that would maximize the overall social surplus against terrorism.

The model captures the strategic interaction between airports, attackers and a regulator. Particularly, the model assumes that airports are heterogeneous in size, and are at risk of a potential terrorist attack. Attackers are assumed to be identical and uniformly distribute their attacks over the population of target airports. From these assumptions, the probability that one or more attacks mounted against specific types of airports are successful is defined. This probability is conditional on the vector of security expenditures of the airports and the attackers' strategic decisions. The probability of a successful attack is also assumed to be affected by the action of other airports.

The model first considers a game where there is no social coordination: Nash equilibrium. A strategy of an attacker is a choice whether or not to launch an attack on the target population and a strategy of an airport is a choice of security expenditure. In this game, there is a strategic interaction between the choices of attackers and airports. The expected payoff for an attacker is affected, in part, by the choices of airports' strategies on security expenditures. Similarly, the expected loss for an airport is determined partly by the choices of attack participation of attackers. As a result, in a Nash equilibrium, the strategies of both parties should be optimal given the expectations about the strategies chosen by other parties, and these expectations have to coincide with the results when all of them behave optimally.

The model then descries a game where there is a policy-maker who desires to minimize a weighted average of the expected losses suffered by the population of airports. Since it is important for him to consider socially desirable ecosystem conditions, the policy-maker is assumed to take into account an externality effect of security expenditures between airports. Particularly, the model assumes that an increase in the security expenditure of an airport has a positive effect on other airports' security levels. From this perspective, a policy-maker designs a regulatory rule which can minimize the overall social expected loss.

5.2 Contribution to the tool

The simulation analysis in the paper is a 'best-case' used to quantitatively investigate how the policy-maker's decision on the levels of security expenditures affects the aviation security ecosystem. For calibrating parameter values for a simulation analysis, we use the data gained from the literature review and a series of interviews. When the data for a parameter value is not available, the value is derived from reasonable guesses based on other research fields.

Table 2 shows the summary simulation results found in Paper 5 and those that provide the base case for the tool. When the interdependence between airports is small, the government's regulation for security expenditures is relatively close to Nash equilibrium security expenditures. However, as interdependence increases, the regulation makes medium and large airports underinvest in security, and small airports overinvest in security compared to Nash equilibrium security expenditures. This implies that medium and large airports can get benefits from the rule whereas small airports take greater costs. As a result, security interdependence makes small airports carry a security burden of medium and large airports.

Further investigation also identifies how the changes in the degree of interdependence between specific types of airports affect airports' security expenditure (e.g., an one-stop security check solution whereby passengers and their baggage does not need to be re-screened at a connecting airport if they had gone through the security check adequately at the airport of origin). The results indicate that, while a security regulation that increases security interdependence between large and medium airports, and large and small airports

Table 2: Nash equilibrium and socially optimal security expenditures with different levels of security interdependence.

	Security Expenditure per Passenger	
Type	Nash Equilibrium	Social Optimum
Low interdependence		
Large	**€5.1**	€4.8
Medium	€6.3	**€7.4**
Small	€7.6	**€9.9**
Medium interdependence		
Large	**€5.1**	€2.2
Medium	**€6.3**	€5.5
Small	€7.6	**€9.2**
High interdependence		
Large	**€5.1**	€0.8
Medium	**€6.3**	€3.0
Small	€7.6	**€8.1**

raises the problem of the unfairness in security expenditures, the regulation that increases security interdependence between medium and small airports might achieve socially fair cost allocation among different types of airports.

This study offers a contribution to the ongoing discussion on the fairness of public policy on security expenditures of airports with different nature. It is illustrated that a current regulatory rule on security expenditure might ask small airports spend more on security than medium and large airports. As a result, while the divergence between private and social incentives for security expenditures suggests the rationale for regulatory rules for security expenditures, it does not guarantee the fairness of such rules.

6. Concluding Remarks

This deliverable summarises the technical contributions from Work Package 6 combined with the results and intuition gained from working with Work Packages 1, 2 and 4. The deliverable presents a compendium of policy papers that outline the theoretical underpinnings of the SECONOMICS Toolkit and the graphical interface tool presented in Work Package 8. The references below represent some of the important external articles referred to in the content of this document. For a complete set of references please see the policy compendium.

BIBLIOGRAPHY

[1] Mark Pauly. Overinsurance and public provision of insurance: The roles of moral hazard and adverse selection. *Quarterly Journal of Economics*, 88(1):44–62, 1974.

[2] Steven Shavell. On moral hazard and insurance. *Quarterly Journal of Economics*, 93 (2):541–562, 1979.

[3] Steven Shavell. *Economic Analysis of Accident Law*. Harvard University Press, 1987.

[4] Richard Cornes and Todd Sandler. *The Theory of Externalities, Public Goods, and Club Goods*. Cambridge University Press, 1996.

[5] Paul K Freeman and Howard Kunreuther. *Managing Environmental Risk Through Insurance*. Kluwer Academic Publishing, 1997.

[6] Ken Binmore. *Natural Justice*. Oxford University Press, 2005.

[7] Georges Dionne and Kili C Wang. Does insurance fraud in automobile theft insurance fluctuate with the business cycle? *Journal of Risk and Uncertainty*, 47(1):67–92, 2013.

[8] Shane D. Johnson. How do offenders choose where to offend? perspectives from animal foraging. *Legal and Criminological Psychology*, 19(2):193–210, 2014. ISSN 2044-8333. doi: 10.1111/lcrp.12061. URL http://dx.doi.org/10.1111/lcrp.12061.

[9] Kevin Murdock. Intrinsic motivation and optimal incentive contracts. *The RAND Journal of Economics*, 33(4):pp. 650–671, 2002. ISSN 07416261. URL http://www.jstor.org/stable/3087479.

[10] Roland Bernabou and Jean Tirole. Intrinsic and extrinsic motivation. *The Review of Economic Studies*, 70(3):489–520, 2003. doi: 10.1111/1467-937X.00253. URL http://restud.oxfordjournals.org/content/70/3/489.abstract.

[11] Ramon Casadesus-Masanell. Trust in agency. *Journal of Economics & Management Strategy*, 13(3):375–404, 2004. ISSN 1530-9134. doi: 10.1111/j.1430-9134.2004.00016.x. URL http://dx.doi.org/10.1111/j.1430-9134.2004.00016.x.

[12] Erik Canton. Power of incentives in public organizations when employees are intrinsically motivated. *Journal of Institutional and Theoretical Economics JITE*, 161(4):664–680, 2005. doi: doi:10.1628/093245605775075942. URL http://www.ingentaconnect.com/content/mohr/jite/2005/00000161/00000004/art00006.

[13] Mark A. Huselid, Susan E. Jackson, and Randall S. Schuler. Technical and strategic human resources management effectiveness as determinants of firm performance. *Academy of Management Journal*, 40(1):171–188, 1997. doi: 10.2307/257025. URL http://amj.aom.org/content/40/1/171.abstract.

[14] Michael H. Belzer and Peter F. Swan. Supply chain security: Agency theory and port drayage drivers. *The Economic and Labour Relations Review*, 22(1):41–63, 2011. doi: 10.1177/103530461102200103. URL http://elr.sagepub.com/content/22/1/41.abstract.

[15] European Commission. Report from the commission on financing aviation security. European Commission, 2009.

[16] Vicki M Bier, Naraphorn Haphuriwat, Jaime Menoyo, Rae Zimmerman, and Alison M Culpen. Optimal resource allocation for defense of targets based on differing measures of attractiveness. *Risk Analysis*, 28(3):763–770, 2008.

[17] Luca Allodi, Marco Corradin, and Fabio Massacci. Then and now: On the maturity of the cybercrime markets: The lesson that black-hat marketeers learned. *To appear in IEEE Transactions on Emerging Topics in Computing*, 2015.

SECONOMICS POLICY PAPERS

Matthew Collinson, Martina de Gramatica, Petra Guasti, Zdenka Mansfeldova, Fabio Massacci, Raminder Ruprai, Woohyun Shim, Joe Swierzbinski, Alessandra Tedeschi, Uğur Turhan and Julian Williams

This deliverable presents a mix of policy papers. They outline both our specific policy recommendations and scientific underpinnings founded in our scientific work and the policy tool. The culmination of the SECONOMICS project is a series of policy recommendations and this is in the form of the policy orientated scientific papers found in this volume.

TABLE OF CONTENTS

[1]University of Aberdeen, UK
[2]University of Trento, Italy
[3]Institute Of Sociology Of The Academy Of Sciences Of The Czech Republic Public Research Institution, Czech Republic
[4]Institute Of Sociology Of The Academy Of Sciences Of The Czech Republic Public Research Institution, Czech Republic
[5]University of Trento, Italy
[6]National Grid plc, UK
[7]University of Trento, Italy
[8]University of Aberdeen, Italy
[9]Deep Blue, Italy
[10]Anadolu University, Turkey
[11]University of Durham, UK

1. EDITORS INTRODUCTION

The SECONOMICS project is founded in the concept of the production of evidence led policy, based on structural models that carefully address the economic incentives of agents and the physical and technological environment that is subject to attack. This work is encapsulated in the SECONOMICS Toolkit, that, in part, is represented by the front end product in the SECONOMICS graphical Tool (see Workpackage 8 deliverables). This tool has two parts, first the graphical visualisation that has been carefully tested by the various partners in WP1, 2 & 3 and second the underpinning mathematical models. These models are programmed in the Matlab programming language and translated using various other languages (such as Java) to integrate them into the Toolkit. The purpose of this compendium is threefold:

(1) To outline the underlying theories used to build the SECONOMICS Toolkit.
(2) To illustrate how these models have been used to build policy in the SECONOMICS context.
(3) To demonstrate, in detail, how we have combined qualitative and quantitative evidence into our policy framework.

SHORT SUMMARY OF THE CONTRIBUTIONS

- Policy Paper 1: The Political Economy of Security Risk Management provides an overview of how we integrated the salience methodology outlines in Workpackage 4, to influence the foundational choices of economic model used in the Toolkit. The paper provides a summary on the public policy implications of security provision and motivates the concept of security as a public good. The key policy insight is in providing the intellectual case for public policy coordination to reduce inequities in cost sharing in security provision. The paper is authored by members of WP4 and WP6.
- Policy Paper 2: Public Policy and Cyber Insurance. This paper provides the most refined version of our general public policy and security with reactive threats model. The major extension of this paper is the quantification of security risks by adding an insurance component to externality model (a marked improvement over the version of the model presented in Workpackage 6 Deliverable D6.2). The mathematical architecture of this paper effectively underpins all of the WP6 models and provides guidance on feasible functional forms.
- Policy Paper 3: Agency Problems and Airport Security. This paper provides us with a summary of the qualitative evidence collected from interviews with various airport partners and authorities and motivates the investment functions used in the SECONOMICS Toolkit. The paper also provides basic game theoretic results that show that the interviews are consistent with many of the behavioural concepts outlined in Policy Paper 1. This paper provides policy insight on how contracts for the employment of security need to structured and the most appropriate usage of training to incentivise good security outcomes.
- Policy Paper 4: Public Policy And The Security of Critical Infrastructure: Discretionary or Audit Based Regulation? This paper provides an effective instantiation of the foundational modelling concepts outlined in Policy Paper 3 for the area of Critical national infrastructure. The paper focused on the key policy issues for CNI, rules (with audit) or risk based (with tort driven penalties) regulation, and then quantitatively identifies the concepts of policy assurance audit in a mathematically rigorous setting. The model provides all of the solutions outlined in the CNI section of the SECONOMICS toolkit and provides some insight into the optimal usage of this policy tool.
- Policy Paper 5: Fairness in Airport Security Expenditures: Equilibrium and Optimum. The final paper in the policy compendium provides the precise instantiation of the mathematical framework underpinning the Airport security model that allows us to identify between different security taxation regimes for various scales of airport across Europe. The paper illustrates how the foundational mathematical models from Policy Paper 2 can be instantiated into the quantitative outcomes presented in the tool and provides insight in how to utilize the tool for the design of such cost sharing systems.

For more an extended summary please see the SECONOMICS Workpackage 6, Deliverable 6.4.

The Political Economy of Security Risk Management

Petra Guasti, Zdenka Mansfeldova and Julian Williams

Security is an inherently economic concept. This review provides a foundational insight into the relevant aspects of economics for the analysis of security problems. We first set up the discussion by providing some historical context on how economics has influenced security and then provide a context for the socio-political aspects of the SECONOMICS project. This chapter unifies the themes from Work Packages 4 and 6 and illustrates how these ideas have been integrated into the SECONOMICS toolkit.

KEY WORDS: Security Economics, The Political Economy of Security

1. THE POLITICAL ECONOMY OF SECURITY RISK MANAGEMENT

The security of citizens has been one of, if not the most important, facets of public policy provision since the dawn of modern civilization. The need for society as a whole to engage in coordinated action to protect individuals from threats, both physical and economic; to their welfare has been a generally accepted principle throughout this time. However, the delegation and coordination of security related activities to a limited number of authorities is, of course, not without cost. The very nature of security enhancing activities lends themselves to the collection and storage of information by a centralized party. However, the political economy of the drivers to the trade-offs discussed previously are often presented most succinctly in the language of economics and game theory. It is to this aspect of the discourse that this chapter is primarily directed.

In contrast, the economics of privacy is a relatively understudied area in general (cf. for an early research on the topic Posner 1980). However, in certain areas, such as medical record disclosure and in certain countries such as the USA, the high level damages awarded for privacy breaches have driven a broad academic and industrial research agenda (see Acquisti 2004, Acquisti et al 2013 for a modern perspective). From a public policy view point, most advanced economies have extensive privacy legislation designed to protect individual citizens from both deliberate and accidental disclosure of confidential information. In the European framework, litigation is increasingly broad to the European Court of Human Rights, under articles 5 and 8 (Guasti, Stockemer, Siroky forthcoming).

We will first briefly review the core economic concepts of utility theory, the adaptability of this concept to non-monetary consumption and the concept of consumption of security, §(2). We will then move on the concept of security as a public good, the ‘production of security and the non-excludable and non-rivalrous nature of certain aspects of security, §(3). Once we have outlined the basic treatments of security as a public good, we can then address the problems of production of a secure environment that achieves the outcome of society via a utilitarian social planner, §(4). The preceding arguments will be based on fully informed citizens and social planners and provide the first best outcomes for the security of society as a whole. However, relaxing these assumptions results in social welfare outcomes in the absence of a planner (at a Nash or ‘rational-

[1] Institute Of Sociology Of The Academy Of Sciences Of The Czech Republic Public Research Institution
[2] Institute Of Sociology Of The Academy Of Sciences Of The Czech Republic Public Research Institution
[3] University of Durham

expectations equilibrium) substantially away from the first best outcomes. Indeed social planners are usually not fully informed and sometimes not benevolent and we will look at a series of examples and outcomes for these cases in §(5).

We will then introduce a further actor in the economic system, firms with capital owners. We will show that the presence of private markets in the security domain can have counter-intuitive effects. In§(6) we illustrate the effect of 'risk neutral firms on aggregate cyber criminal behaviour. We then look at risk averse corporate officers in §(7) and introduce the notion of an insurance contract. We then speculate on the impact that delegation of security coordination by the social planner will have on the risk-averse firms and whether this will lead us towards our first best outcomes. Our final discussion section §(8) will look at the economics of citizens choices and collective structures in security and the reasoning behind their inception and we will discuss the area of 'institutional analysis and design in the last part of the commentary and concludes.

2. A UTILITY THEORY OF SECURITY

Utility theory in its most basic form acts as a form of book keeping for preferences between various different actions. Whilst this may seem a very abstract and foundational as we are looking at atomic preferences, actions and outcomes; understanding how individuals make decisions between different choices of actions is fundamental in establishing the globally optimal actions and outcomes that benefit society as a whole. Indeed, a 'utilitarian social planner is one that seeks to maximize a weighted sum of the utility of all individuals within their purview (indeed the most basic form of utilitarian social planner uses a Bentham-ite welfare function that weights are individuals equally.).

For an individual, every atomic action results in the consumption of some form of 'good. We often measure goods in monetary equivalents as the range of items that classify as a consumption good include abstract notions such as happiness, leisure, safety and wellbeing, in addition to physical items such as food. We can consider that a utility function acts as an accounting mechanism that combines the consumption of a list of goods into a measure of overall welfare or aggregate utility.

Once we have a utility function that maps multiple measures of consumption to a single measure of utility and we can identify the probability for each outcome state we can begin to make specific predictions regarding the choices of individuals. The two most basic assumptions are first, that when the quantity of any given consumption good increases the associated degree of utility increases; second, as the amount of consumption of a particular good increases the corresponding increase in utility gets smaller and smaller. This means that whilst the rate of increase in utility with consumption is always positive, the rate of change in the rate is negative.

We can view security in two ways. First, as an indirect feature of a utility function, for instance as an input into probability of a particular outcome for a consumption good. A simple example is the consumption of an income generating good, a net reduction in the security of supply an asset generating a financial return will lead to an increase in the volatility of future wealth outcomes. Therefore, the impact of security is via the variation in another good.

Second, we can also view the consumption of security as a direct contributor to the utility function. This maybe measured in terms of the direct sense of well-being that security conveys to the individual. The fact that the degree of security may not directly affect an individuals consumption of other goods, but simply engenders a sense of safety for which the individual may wish to sacrifice consumption in other areas to specifically maximize their welfare. It is relatively simple to specify a utility function using either approach, however determining which is the most relevant is very difficult to ascertain in practice.

Whilst the previous exposition on individual citizens preferences at first appears straightforward, determining the exact structure of preferences in practice is extremely difficult, and has been the subject of extended experimental research. One approach is to use various different combinations of lotteries which are given to subjects and their choices should elucidate the properties of the utility function. This approach has, in general, not provided results that are consistent between the experimental outcomes and the preferences implied by individuals regular economic activities. Indeed, this has been one of the major puzzles in economics over the last hundred years or more.

3. IS SECURITY A PUBLIC GOOD?

The first and second fundamental theorems of welfare economics in their current form most notably attributable to Arrow (1951) Debreu (1959) form a mathematical foundation to the ideas put forward by Adam Smith in the wealth of nations. Whilst the core principles of the first and second theorem are most applicable to the non-security economic debate, for instance for fiscal policy, the conditions under which the first and second fundamental theorems of welfare economics do not function are of considerable interest to the political economy of security provision.

The first fundamental theorem of welfare economics states that a competitive economy is always Pareto efficient; that externalities, the underproduction or over-consumption of public and common property goods are incorporated in each individuals welfare function and hence are not an issue for the sustainability of the economy. The second theorem suggests that every Pareto efficient allocation can be attained through a price system, such as a competitive market with a clearing price.

Why can't we provide security by simply letting each individual decide on their own security and allow them to provide joint surpluses for the production of a secure society? In essence the fact that something is a jointly produced good means that consumption of that good, in this case a secure society, requires joint production. Unlike many other types of goods that can have a market price attached to them from a clearing market, it is difficult to envision a market-based solution for security provision. Once we violate the second theorem, we can find natural objections to the first theorem. Therefore if we cannot create an efficient market to clear the value of security, we need to have a public policy-maker impose some mandatory liability sharing, such that all citizens contribute to the overall security of society. It is worth restating that is not a circular argument, if we cannot create an efficient market for the production of security, then we cannot price security, as such we cannot achieve a Pareto efficient outcome to provide security for citizens overall.

One of the major issues with security is that effective provision tends to rely on the social coordination and preferences of more than a single individual. As such the preferences of the individual (in choosing to allocate resources to security) have a net effect greater than that of a single individual. If we think of security as a good that needs to be produced, then individuals need to choose to contribute to its production. For instance, cost of production of security may not be solely monetary, privacy costs may need to be incurred.

Goods in most forms need to be produced via some form of costly allocation of labour and capital. In this instance, capital includes endowments that are both financial and non-financial, such as an individual perceived endowment of privacy. A public good has a set of particular properties; first, it is assumed to be non-rival that is one individuals consumption of the good does not reduce its availability to others; second that the good is non-excludable that is one or more individuals cannot prevent others from using it. As discussed in previous chapter, the choices of one individual to invest in security can effect the aggregate level of security overall. In this sense security has some of the attributes of a standard public good, such as public parks.

Production of a public good is the subject of extended debate in the economics literature. Underproduction or excessive use of a public good can lead to its degradation and eventual disappearance. When a good is non-excludable everyone can use it, without having to allocate resources to its production. Consumption of the good can then outstrip supply and eventually the good is exhausted. National security is often cited as an archetypal public good, the production of which needs to be regulated centrally by a social coordinator who ensures that all those, who consume the public good make a fair contribution to its production and prevent excess consumption. In contrast, public goods can also suffer from over production, if individual agents within an economy can convince the social coordinator that a public good deficiency exists.

The public good aspects of security are somewhat different to other typical public goods is that the consumption of security is not specifically decided by an individual. For instance, there is a growing critique in Western media that after 9-11 national security is being exploited by private companies for economic gain. Should this be the case, the overproduction of security may be considered as part of the public good problem, but in an unusual way. The mechanism is effectively driven by information asymmetries between those contracted to supply security, the social coordinator and citizens as a whole. This is in effect a sought rent by the contracted supplier of security.

3.1 Public policy and social planning in security

In the previous section we established the basic socio-economic structures needed to provide insights into production of security within an economy. Up to this point, we have used abstract notions of who the individual agents and social planners are. For the remainder of this chapter, we will now concretize the structure using the typical entities present in an advanced economy. We will also introduce the influence of firms, alluded to previously, providing security goods and risk management.

In Figure 1.5. we present an overview of our tour of the various entities and their interactions within the socio-political frame of security. There are other ways of interpreting these mechanisms, but each is more or less a derivative of the standard political-economic institutional setup (see for instance Laffont 2008). Furthermore, to substantiate our discussion we will focus on a single facet of security, one that is usually termed cyber security, and covers the disclosure and integrity of information assets held by the citizens, government and firms within the global economy.

In the reminder of this chapter we will cover the following topics policy coordination at the supranational and national government level; security strategy and planning for firms; the role of insurance in managing cyber risk; and a summary of the approaches to designing optimal institutions that ensure security and privacy of citizens at an appropriate cost.

4. NATIONAL AND SUPRANATIONAL PUBLIC POLICY

There is currently an on-going policy debate concerning the appropriate nature and extent of regulation to maintain the security of information assets. On one side of the debate, some policy-makers argue that the provision of advisory information by governments and the use of voluntary standards would be the best choice - see for instance the on-going consultation on the EU Network Information Security Directive. Others maintain that some form of compulsory regulation is required. For example, the UK Parliamentary Office of Science and Technology made the following observation concerning the issue of cyber security. Opinion is divided as to whether cyber security regulation by government would be the best way forwards. Regulation could increase the level of adherence to best practice; however it will always lag behind developments in technology and would be difficult to monitor. (see POSTNOTE 389, Sept. 2011).

In the European Union, ENISA (the European Network and Information Security Agency) is seriously considering the use of compliance-based regulations to supplement voluntary approaches (ENISA, 2012). In contrast, in the United States, the presumption is that what is required is a culture of voluntary good practice based on efficient information sharing. (See, for example, Blueprint for a Secure Cyber Future, Dept. of Homeland Security, 2011.)

Simple theoretical models of investment and risk mitigation in cyber security demonstrate that the circumstances under which the social and private incentives to appropriately invest in cyber security can be expected to differ. One surprising result is that even if the technological environment were to be modified to minimize conflicts between social and private incentives, the nature of attacker behaviour is itself likely to create an incentive for underinvestment in security and a consequent need for government regulation.

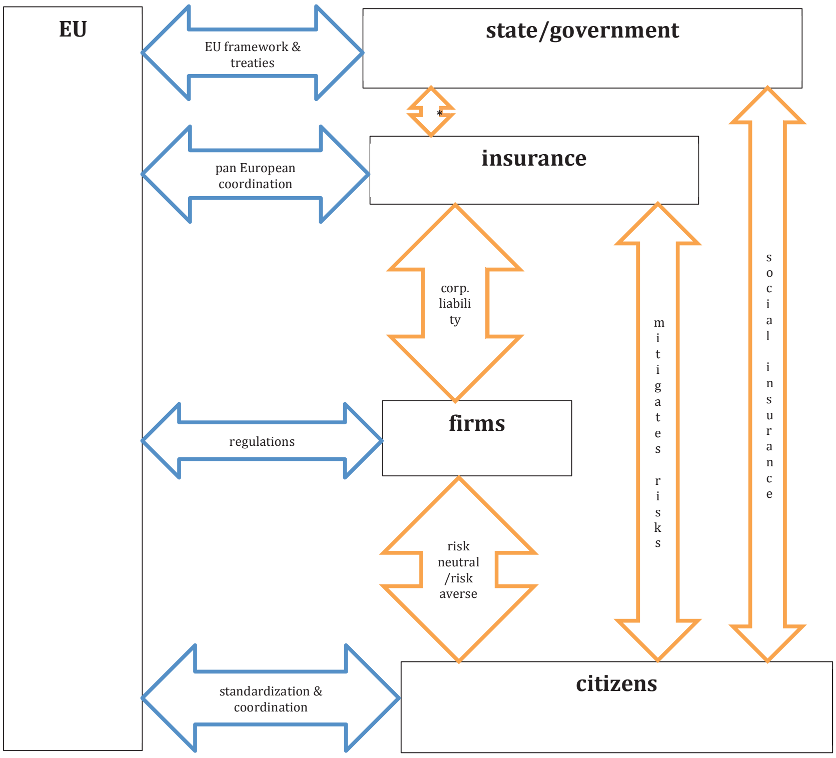
Externalities can occur because of the nature of the technological environment. For example, vulnerability in the software or hardware of one prospective target may create an avenue for attacks on other targets. In addition, when a successful attack is carried out on one target, this may create losses for others as well. For example, when one firms service is interrupted due to an attack, the firms customers may suffer losses that are not fully compensated by the firm in question.

There is evidence that attackers dynamically readjust their effort in response to the behaviour of attackers and potential targets. For example, Herley (2012) observe that one reason for email phishing attacks is to identify the email users who are most likely to fall prey to an attack. They observe that such behaviour is only sensible if an attack on all potential targets is too costly and attackers intend to focus attacks on more vulnerable targets. Baldwin et al. (2012) show that spikes in attacks on specific systems can lead to mutual excitement of attacks on other systems. Such behaviour suggests that attackers respond to an indication of a profitable opportunity (i.e., the initial attacks) by launching more attacks.

The conficker computer worm provides an example of the importance of externalities in the

Fig. 1. A socio-political structure of security.

Figure: A socio-political structure of security



Note: * state and insurance - data and information collector and aggregator of monetarizable risk
Source: Seconomics

context of cyber security. The conficker worm was first detected in 2008 and, at its peak in early 2009, had infected between 9 and 15 million computers. An interesting aspect of the conficker worm was that it posed relatively little danger to an individual infected machine, but turned this machine (usually in its down time) into a component of a larger botnet which was then used to mount attacks on larger computer systems via spam emails or denial of service style attacks.

One of the issues with combating such a worm was that many of the computers involved were commercial units housed in call centers and other large offices.4 Because of the relatively small level of damage to individual machines and the relatively high cost of defending against the conficker virus, the time taken to mitigate this worm was relatively slow. Nearly four years after the worms release, 1.7 million machines were still infected with the conficker worm.

Although the conficker worm was not particularly sophisticated technologically, it exploited in a sophisticated way the perverse economic incentives created by externalities. The cost of mitigation for large offices was higher than the risk adjusted cost to these offices of having the worm on their systems. Hence, many firms were slow to take action to remove it. There is a large literature in the field of economics on externalities and related topics such as public goods. See, for example, Varian (2010) for an introductory discussion and Cornes and Sandler (1996) and Laffont (2008) for more advanced treatments. However, there have been relatively few applications of the economic theory of externalities to the field of computer security.

There is a large literature in economics on externalities and related topics such as public goods. See, for example, Varian (2010) for an introductory discussion and Cornes and Sandler (1996) and Laffont (2008) for more advanced treatments. However, there have been relatively few applications of the economic theory of externalities to the field of computer security. Varian (2004) is one well-known example that considers how the nature of technological externalities may affect the level of investment in the reliability of information systems.

Varian (2004) considers how the nature of technological externalities may affect the level of investment in the reliability of information systems. Arora et al. (2008) models a single firms policy for disclosing and patching software vulnerabilities. The paper shows that the timing of disclosure depends on the fraction of the total cost of an attack, which is not borne by the firm. Kunreuther and Heal (2003a) and Kunreuther and Heal (2003b) model a group of agents each of which can choose between two alternatives that affect not only the agents own risk but also the risk of other agents. They consider factors that can cause the equilibrium actions of the group to tip from the high-risk to the low-risk alternative.

An early contribution to the literature on investment in information security by Straub and Welke (1998) outlines a model of threat and countermeasure that models risk as a combination of attacker and defender effort. Treating risk as a function of defender effort, Gordon and Loeb (2002) present a model of decreasing marginal returns to security investment. They propose a residual risk function that relates investment to the probability of a successful attack. Optimal investment in security is, therefore, a tradeoff between the risk-adjusted expected loss and the deterministic level of investment. Other threat models, such as Ioannidis et al. (2009, 2011, 2012); Chen et al. (2011) or Gordon et al. (2010), utilize a real options or portfolio optimization approach to model the defensive response of a firm. In the above papers, the behaviour of attackers is assumed to be exogenous in the sense that attackers do not respond to targets actions.

Some papers consider the interactions between attackers and defenders. For example, in Cavusoglu et al. (2008) a firms security manager must estimate an attacker effort function in order to compute the firms optimal expenditure on security. Florencio and Herley (2011) consider the relationship between the incentives of attackers to mount attacks and the observed volume of attacks. The papers by Cremonini and Nizovtsev (2010) and Fultz and Grossklags (2009) model the level of security in a computer network as the outcome of a strategic game between attackers and defenders. Png et al. (2006) model the response of a single attacker to the security efforts of a number of software users.

Public policy cannot be set without due reference to the reaction of firms. The policy makers objective function may be a standard utilitarian social planner seeking a Paretto efficient security allocation or an alternative. It is therefore instructive to delve deeper into how firms make a security investment decisions.

5. SECURITY AND THE FIRM

A standard tenant of the industrial organization literature is the separation of ownership and management of firms. When owners of capital can diversify investments across multiple firms, the decision making taken by those firms is usually assumed to be made in a risk neutral setting; whilst the corporate officers who make decisions for the firm would be expected to exhibit risk aversion. However, recent results from the financial literature indicate that the assumption of perfectly efficient capital market is not valid. We will show that in this case, firms decision-making will quite likely approximate the risk aversion of the corporate officers. Furthermore, when firms have variation in their risk preferences we can see that there will be a tension with the wider public policy mandate.

It is relatively easy to show that for a risk neutral firm, when security exhibits diminishing marginal returns to security investment, for a fixed level of attacking intensity, an increase in the discount rate leads to a net reduction in total security investment. Unfortunately, as security investment, in aggregate, is reduced, more attackers will be attracted to the targets, resulting in a greater increase than would be expected given a fixed level of attacking intensity. This attacker externality can only be overcome by a social planner mandating security investments across targets.

In public economics, the chosen discount rate is often referred to as the social discount rate or policy-maker time preference. The most commonly encountered public policy social discount rate is the base or policy rate dictated by central banks issuing fiat money. However, for most policy decisions requiring time preferences to be imposed either in legal structures or public investment the central bank policy rate is not used as it is deemed to be a basic rate, and therefore a premium is added or subtracted.

The fact that private discount rates diverge from social discount rates is the subject of extended discussion in the economics literature (e.g., Caplin and Leahy (2004); Lew (2011)). However, the broad consensus for firm private discount rates focuses on the financial economic viewpoint. Models such as the capital asset pricing model indicate that log linear preferences relative to a single representative risky rate can reduce the discount rate problem for private firms to a simple measurement of the covariance of asset valuations to the broader economic system. The fact that a risk premium exists for firms indicates that if targets are assumed to be firms and the steward is a public policy-maker, then the social discount rate will be required to be less than as risk premiums are always positive.

A more attractive way of thinking about discount rates is to derive the time horizon over which the majority of their value amortizes towards zero. This provides a baseline for the stewards time horizon in terms of managing externalities. Should the steward desire the externalities to be managed over a longer, more sustainable, time horizon, then his discount rate will be set lower than the representative rate determined by the individual firms. Larger scale ecosystems such as the internet are usually assumed to require longer term planning. Hence, stewards in this context might amortize expected losses from risks to the system over much longer periods. Therefore costs are imposed at rates that individual participants in the ecosystem may believe to be unjust given their own time preferences.

The very low social discount rate problem is an area of active debate in environmental economics and in particular the economics of climate change. The UK Governments Stern Review, see Stern (2006) sets time preferences with respect to a discount rate approaching zero. This has sparked substantial debate in the economics literature, as future losses from climate change impacts have not been discounted at rates markedly similar to public or private investments; see Nordhaus (2007); Weitzman (2007) for extended discussion. The issue is more acute here as losses from climate risks are generally assumed to be realized at a reasonable distance into the future. Therefore, even small discount rates have very little impact on the current cost-benefit analysis assessing risk mitigation.

For information security contexts, the impact of the time preference assumption is not so acute as investment horizons are much shorter (see for instance Ioannidis et al. (2012) for a model of investment horizons). However, the interaction of the externality with the differentiated discount rate between targets and the steward does indicate that this is an important issue for information ecosystems. Under certain measurements, targets may have very large discount rates amortizing information assets over periods as short as 12 to 18 months.

An example of the debate on choice and imposition of social discount rates is in the climate change literature Caplin and Leahy (2004), where the choice of discount rate is particularly acute as the forward horizons are over multiple decades and centuries and, in this context, exponential discounting reduces future losses toward zero after a finite number of years. However, the speed of discounting by firms of their information security assets can be very high suggesting that the rate could be as high as 40% per annum. It is unclear whether this discount rate also applies to future losses. If so, then the private discount rates would be expected to be very different from stated public discount rates that are normatively closer to 10% for developed countries (see US office of Management and Budget policy overview Lew (2011)).

6. SECURITY AND INSURANCE

Some risks can be compensated ex-post by an agreed monetary payment an insurance contract. In general, we will think about corporate liability insurance, which has the potential to act as a surrogate for an appropriate level of social protection by the state. Governments often find information collection and processing extremely difficult, this is why markets are commonly used to allocate productive capital. The action of making financial transaction places specific value on different outcomes. This inherently reduces information asymmetries and in theory should lead to welfare efficient outcomes; when markets function. The easies mechanism to place a specific value on security risks is in the market for insurance contracts protecting against liabilities and losses from cyber-attacks.

Cyber-insurance provides coverage in the event of a successful attack on the information infrastructure of an organization. This section outlines a game played between group(s) of targets that invest in defensive expenditure to reduce the risk of a successful attack by one or more attackers. Attackers are modelled as criminals with fixed costs who engage in a competition to infiltrate and expropriate valuable information from the targets. Into this mix we introduce a variety of insurance contracts and behavioural restrictions that the insurer can then impose on the target. Our main result is that neither monopolist insurers nor a fully competitive insurance market have the incentive to reduce inherent externalities within the market. In contrast, the monopoly insurer, acting rationally, would be positively assuaged towards inflating the cyber threat as long as they can identify the actuarially fair price of insurance risk and the maximum quote premium they can charge to risk averse targets.

The impact of moral hazard and adverse selection in the presence of insurance has a long history of investigation in economics, see for instance Pauly (1974); Shavell (1979, 1987); Cornes and Sandler (1996); Freeman and Kunreuther (1997) and Binmore (2005) for an eclectic set of examples that directly relate to the notions of public policy, liability sharing and insurance considered herein. In contrast, the impact of the presence of insurance and how the collective behaviour of victims of crimes can influence the aggregate behaviour of criminals generating the risks that are being insured, has far lower profile in the literature, although Cornes and Sandler (1996) provide a materially similar treatment to our cyber-insurance case. We can think of these effects as moral hazard and adverse selection effects that are once removed from the actions of the insuree, i.e. not the direct influence of target behaviour on target risk, with and without insurance, ceteris paribus, but the impact that changes in the aggregate behaviour of the pool of externally-insured or self insured have on the risk vectors generating the distribution of losses. Furthermore, adjustments in external environmental conditions have been shown to affect the distribution of insurance claims and their legitimacy, see for instance Dionne and Wang (2013) for auto insurance fraud. Our attacker externality shares several similarities with the incentive to defraud suggested in Dionne and Wang (2013). In that there is a systematic factor that varies across all targets. However, in our case this is fully endogenous, rather than driven by an external macroeconomic effect.

For a typical cyber-insurance contract a policy aimed at a small medium enterprise and is designed to limited liability up to 750,000 (approximately $1.2 million). The coverage includes direct losses of time and information assets as well as follow on legal costs incurred by the firm for loss of customer records and breaches of data protection. For instance Page 4 Clause (c) indicates that the insurance will pay for "... your unauthorized collection or misuse of any data concerning any customer or potential customer of yours which is either confidential or subject to statutory restrictions on its use and which you obtained through the internet or extranet".

The role of competitive and non-competitive insurance markets in economic decision making is extremely old, for instance in the Code of Hammurabi (1772 BC) liability between traders was covered by a premium with a deductible and aggregate choices of these traders in demanding particular safety features on boats affected the aggregate premium. However, the lineage of our modelling approach is more appropriately based around Arrow (1974) and Rothschild and Stiglitz (1976) who introduced the modern mathematical approaches to dealing with moral hazard and adverse selection when designing insurance contracts and Ehrlich and Becker (1972) that delineates the concept of self protection and self-insurance versus external insurance.

The presence of an appropriate level of deductible for insurance against losses when the risk generating function was solely a combination of the behavioural choices of the insured and the exogenous background risk is comprehensively addressed in Raviv (1979) and Schlesinger (1981) which now form the part of the canon of textbook treatments on this subject. However, the key focus has been on the impact of insurance on the individuals behaviour and the optimal contract design (deductible, screening, behavioural requirements, monitoring) and not on the background risk process generating the need for insurance. The differentiation between an individuals demand for insurance and a firms demand for insurance is addressed in Mayers and Smith Jr (1987). Corporate insurance differs from insurance for individuals as corporate stockholders (in diffuse ownership environments) can diversify away insurable risk. Therefore whilst owners maybe risk averse themselves, it does not explain the demand for corporate insurance for a value-maximizing firm. That corporate officers buy insurance to hedge against risks to their own positions is empirically investigated from a legal perspective in Baker and Griffith (2007), who uses the demand for corporate liability insurance as a predictor of firms corporate governance risk. Indeed, Griffith (2006), again from the law perspective, argues that the SEC should mandate insurance details of officers and directors liability insurance policies.

That firms demand in significant liability coverage cannot be in question. Specifically in the domain of Cyber liability insurance Lloyds of London in 2014 reported that 75 million of insurance premiums were paid by UK small medium enterprises. Furthermore, several media outlets have reported that the US market for cyber insurance may approach $1.5 billion in 2015.2 Caillaud et al. (2000) suggests that risk-neutral firms will demand insurance as they are induced to risk-aversion as the disclosure of accident-losses that deteriorate the profitability of projects is private to the firm and costly audit is needed to demonstrate that the accident was something that the agents managing the firm could not control. MacMinn and Garven (2000) also argue that the demand for corporate insurance stems from the firm choices mimicking the risk averse behaviour of its corporate officers.

Finally, Holmstrom and Tirole (2000) further the argument that the mix of contracting and agency costs can induce risk aversion in corporate decision-making. Given that a corporation is a diffuse entity, determining the correct typology of utility function to use in a quantitative analysis maybe difficult. In an important contribution Grossman and Hart (1982) deliberate on how corporate decision-making is driven by the characteristics of the corporate officers and their incentives. That corporate officers prefer not to take risk-neutral bets is a fairly well understood phenomenon. Transactions costs from moving jobs from one firm to another results in corporate officers who will seek to hedge bets as bankruptcy injects a shadow of the future problem for the officer; in Williamson (1989) the vertical structure of a firm and the inherent behavioural nature of financial decision making within are reviewed, along with a summary of empirical tests. Frictions are determined to be core drivers in how corporate decision making falls more into line with the behavioural characteristics of the firms officers than that of a strictly risk neutral entity working on behalf of highly diversified owners.

It is an interesting paradox that most decision making by individuals, in an experimental sense, is conducted on small lotteries where marginal changes in choices may lead to large fluctuations in the implied properties of the individuals utility function. Whilst most corporate decision making on the purchase of insurance and risk taking takes place for very large stakes well away from the problematic region, however our major theories of industrial organization suggest that firms should in fact, in general, be risk neutral. This problem of identification is a major concern of Rabin (2000), who demonstrates that standard utility theory is not useful in an experimental setting. However, empirical analysis in corporate decision making suggests that corporate officers are indeed risk averse.

Induced risk aversion in relation to firms demand for insurance, in particular for property, asbestos and pollution insurance (inherently demanded by corporations) appears to be a regular source of income for insurance companies. Financial Times reported that a ...well-regarded insurance analyst, who declines to be named, says [sic]: The ideal scenario this year is we have some hurricanes. The context of this comment is specifically in relation to the association between the realization of events, perceived increases in threat from liability claims and the subsequent demand for insurance, by firms with a need to satisfy firm level risk aversion.

Precautionary saving and the device of precautionary insurance is discussed at length in Eeckhoudt and Schlesinger (2006). Risk aversion, prudence and temperance are categorized by the preference for certain types of lotteries over other alternative lotteries. The measurement of prudence versus imprudence is based on the sign of the third and fourth derivatives. A mechanism that is analogous to the higher moments literature in asset pricing, see Scott and Horvath (1980). Unifying these two strands of the risk taking and utility theory literature is Deck and Schlesinger (2010), who show that individuals choosing between lotteries exhibit prudence and temperance in a laboratory setting. Unfortunately, the literature on industrial organization and corporate decision-making provides little insight into the translation of the personal preferences of corporate officers to the realized outcome in terms of the induced utility of the firm. Paulsson and Sproule (2002) provide some theoretical results on risk management for firms and allow the utility function that implies the preferences of the corporation, as an entity, to include the sign restrictions of Scott and Horvath (1980), implying both prudence (the third derivative being always positive) and temperance (the fourth derivative being negative), with imprudence and in-temperance having the opposite sign restrictions.

It therefore appears reasonable to consider both the prudent and imprudent cases when considering firm risk taking in particular. Given the demand for corporate liability insurance the concept of precautionary investment to mitigate first and second order risk factors appears to be reasonable for cyber insurance. Indeed, as Pal et al. (2013) very carefully illustrates, when cyber security vendors are able to use their population level knowledge of security risk versus potential targets and insurees limited knowledge, significant gains are possible for a vendor that can act as a quasi insurer, for instance by guaranteeing up-time or ensuring transactions. In this case it would appear that the targets optimal approach at the aggregate level would be to engage in carefully coordinated self-protection.

We will show that our results generally hold when losses for security events are large, as would be expected in the case of corporate security liabilities. However, as Rabin (2000) suggests, the choices made by individuals assumed to be a an expected-utility maximizer with a concave utility function can produce results that run counter to intuition and experimental results on choices of lotteries when the stakes are very small.

7. CITIZEN CHOICES AND SECURITY

Up to this point we have addressed decision making in regard to security risks in a structured legal setting, for instance in the contractual setting of a firm or the legal setting of public policy. We will now present the arguments surrounding the citizen provision of security. We will start off with self-protection before moving onto how natural, non-binding public institutions can arise in the absence of a binding regulatory framework.

We will first look at the concept of self-protection, as suggested in Ehrlich and Becker (1972). A self-protecting target is one that employs costly counter-measures against attack. A key assumption in the information security literature that also applies to the physical security literature is in regard to the rate of risk reduction for a give extra level of investment. The concept of diminishing marginal returns to security investment defines a pattern of risk reduction, where every extra unit of investment in security sees an ever-smaller reduction in the observed level of risk. When investments are linear, this results in a unique level of optimal investment that a target will engage in when no other means of buying protection (such as insurance are possible).

Of course this investment occurs over a given period and the losses that could potentially be suffered in the future will not have the same degree of impact as nearer term risks. These kinds of issues rely on the concept of discounting, how we value future outcomes relative to current investment. The discount factor dictates the degree by which future outcomes are valued in terms of current period consumption. By convention a discount factor is a quantity benchmarked around unity. When the discount factor is one, the value of an outcome is in equivalent current period valuation.

Discount factors are difficult to measure for individuals, however, in the presence of efficient markets providing capital for firms, we can take better measurements of a firms discount rate, from the measurable rate of return on invested capital. We will now begin to discuss this in the context of economies consisting of large numbers of firms, investing in their own security.

It is worth, at this juncture, to discuss briefly the rationality (or lack thereof) of attackers in this context. Typically we model attackers as maximizing a utility function that operates over a variety of consumption factors. These can include maximizing financial rewards, a utility of anarchy (e.g. a utility function that increases with the variation in the outcomes for others) and potentially a utility based on political motivations (for instance a terrorist who measures success in terms of targets killed). Rationality, in these cases is often difficult to model; however, many studies have indicated that once the consumption good for the attacker is identified, they typically exhibit the normal properties of diminishing marginal utility and risk aversion.

We have seen that whilst individuals can manage their risks without specific behavioural doctrines imposed on them either by government or conditions placed within an insurance contract. It is still maybe important to have collective action mechanisms in place to coordinate the security effort of individual citizens.

An approach to designing coordination mechanisms is the Institutional Analysis and Development (IAD) framework promoted in Crawford and Ostrom (1995) and Ostrom (1997). In Crawford and Ostrom (1995) three types of institutional statements are considered in a policy ‘action arena. A policy action arena is a domain of interactions, such as the regulation of cyber security. Within the action arena we observe rules, norms and strategies. In addition to these statements there is a syntax of institutional statements ADICO is a compression of Attribute, Deontic, aIm, Condition and Or else. For each of the three statements there is a subset. For rules the entire syntax is valid (ADICO), for norms only attribute, deontic, aim and condition apply, (ADIC). Strategies include only attribute, aim and condition (AIC).

How does this approach help in the design of appropriate policy institutions? The attribute is the individual or organization to which the policy institution statement applies. The deontic is the ‘prescriptive operator of an institutional statement that describes what is ideally permitted, obliged or forbidden (Crawford and Ostrom 2005, pp141–149). The aim describes the goal or action of the statements for which the corresponding deontic refers. The condition represents the operators denoting when and where the aim is appropriate. The Or else is the punishment action when a rule is not adhered to.

A game theoretic approach necessarily reduces the institutional problem to a simplified mathematical mechanism. Optimizing the mechanism is known as a ‘mechanism design problem. However, undertaking this task for a complex system provides a great number of challenges and in some cases is almost impossible to solve in a tractable manner. The IAD framework simplifies the institutional design narrative substantially and attempts to allocate a syntax that provides a practical approach to public policy design of legal structures and regulatory bodies.

The LINUX community is a good example of a non-contractually binding community group that provides complete operating systems and software for users. In this case, security is an investment by all users, who engage in open source production of tools. The open nature and continuous review of components is designed to ensure that security features are continuously updated. However, for certain important elements structured repository drives managed by well organised groups provide insurance to the wider user base. An important aspect of this community is the ability to choose the level of security risk to which user is exposed, in an informed way. Certain experimental software may not be available through the normal channels or contain higher risks. The user can gauge the risks and assessed cost-benefit analysis of using the software. Of course this style of citizen continuous involvement is applicable to certain domains.

8. CONCLUDING REMARKS

This chapter has addressed the interplay between public policy and social attitudes towards security risks. We have established security as a component of a utility function, both in terms of its impact on standard consumption goods and as a specific dimension of the utility function in its own right. We have then discussed the concept of a utilitarian social planner aggregating the utility of all individuals to find the optimal level of security investment. Using a series of examples we have look at supranational and national policy making for instance the interplay between the member states of the EU and European directives.

We showed that in the absence of a benign social planner, aggregate security will probably not attain a social optimum. The reason for this feature, is the public good aspect of security means that the consumption is non-excludable, however, individuals can benefit from not contributing by free-riding. We have then looked carefully at public policy and cyber security and the relationship between firms and a public policy maker. We then showed that firms might still consume corporate liability insurance, even if the standard theories of the firm suggest that decision-making should be taken under risk-neutrality (as owners are well diversified). We then discussed the public policy structures that may be used to manage the public good aspect of security. We then discussed one of the major approaches to designing community based public policy instruments, IAD and how this might be used in the future to address new and existing institutional mechanisms as new and innovative digital communities arise.

ACKNOWLEDGEMENTS

This work has been partly funded by the European Union’s 7th Framework Programme under grant agreement no 285223 - SECONOMICS (www.seconomics.org).

BIBLIOGRAPHY

Arora, A., R. Telang, and H. Xu (2008). Optimal policy for software vulnerability disclosure. Management Science 54(4), 642656.

Arrow, K. J.; Debreu, G. (1954). ”Existence of an equilibrium for a competitive economy”. Econometrica 22 (3): 265290.

Arrow, K. J. (1974). Optimal insurance and generalized deductibles. Scandinavian Actuarial Journal 1974(1), 142.

Arrow, K. J. and M. Priebsch (2011). Bliss, catastrophe, and rational policy. Environmental and Resource Economics, 119.

Asplund, M. (2002). Risk-averse firms in oligopoly.

International Journal of Industrial Organization 20(7), 9951012.

Baker, T. and S. J. Griffith (2007).Predicting corporate governance risk: Evidence from the directors & officers liability insurance market. The university of Chicago law review , 487544.

Baldwin, A., I. Gheyas, C. Ioannidis, D. Pym, and J. Williams (2012). Contagion in cybersecurity attacks. In R. Bohme (Ed.), Workshop on the Economics of Information Security 2012. WEIS.

Binmore, K. (2005). Natural Justice. Oxford University Press.

Binmore, K. (2007). Playing for Real; A Text on Game Theory. Oxford University Press.

Binmore, K. (2007). Playing for Real. Oxford University Press.

Caillaud, B., G. Dionne, and B. Jullien (2000).Corporate insurance with optimal financial contracting.

Cavusoglu, H., H. Cavusoglu, and J. Zhang (2008).Security patch management: Share the burden or share the damage. Management Science 54(4), 657670.

Chen, A., A. Pelsser, and M. Vellekoop (2011). Modeling non-monotone risk aversion using sahara utility functions. Journal of Economic Theory 146 (5), 20752092.

Chen, P., G. Kataria, and R. Krishnan (2011). Correlated failures, diversification, and information security risk management. Management Information Systems Quarterly 35(2), 397422.

Cornes, R. and T. Sandler (1996).Theory of Externalities, Public Goods, and Club Goods (2nd ed.). Cambridge University Press.

Cornes, R. and T. Sandler (1996). The Theory of Externalities, Public Goods, and Club Goods. Cambridge University Press. CoRR abs/1208.3994.

Cremonini, M. and D. Nizovtsev (2010). Risks and benefits of signaling information system characteristics to strategic attackers. Journal of Management Information Systems 26(3), 241274.

Deck, C. and H. Schlesinger (2010). Exploring higher order risk effects. The Review of Economic Studies 77 (4), 14031420.

Dionne, G. and K. C. Wang (2013). Does insurance fraud in automobile theft insurance fluctuate with the business cycle? Journal of Risk and Uncertainty 47 (1), 6792.

Eeckhoudt, L. and H. Schlesinger (2006). Putting risk in its proper place. The American Economic Review , 280289.

Ehrlich, I. and G. S. Becker (1972). Market insurance, self-insurance, and self-protection. The Journal of Political Economy , 623648.

Florencio, I. and C. Herley (2011). Where do all the attacks go? In B. Schneier (Ed.), Workshop on the Economics of Information Security 2011. WEIS.

Freeman, P. and H. Kunreuther (1997). Managing Environmental Risk Through Insurance. Kluwer Academic Publishing.

Fultz, N. and J. Grossklags (2009). Blue versus red: Towards a model of distributed security attacks. In

Gordon, L. A., M. P. Loeb, and T. Sohail (2010).Market value of voluntary disclosures concerning information security. Management Information Systems Quarterly 34(3), 567594.

Gordon, L. and M. Loeb (2002). The economics of information security investment. ACM Transactions on Information and Systems Security 5 (4), 438457.

Gordon, L. and M. Loeb (2002). The economics of information security investment. ACM Transactions on Information and Systems Security 5(4), 438457.

Griffith, S. J. (2006). Uncovering a gatekeeper: Why the SEC should mandate disclosure of details concerning directors and officers liability insurance policies. University of Pennsylvania Law Review, 11471208.

Grossman, S. J. and O. D. Hart (1982). Corporate financial structure and managerial incentives. In The economics of information and uncertainty, pp. 107140. University of Chicago Press.

Hemmer, T., O. Kim, and R. E. Verrecchia (1999). Introducing convexity into optimal compensation contracts. Journal of Accounting and Economics 28 (3), 307327.

Herley, C. (2012). Why do Nigerian Scammers say they are from Nigeria? In R. Bohme (Ed.), Workshop on the Economics of Information Security 2012. WEIS.

Holmstrom, B. and J. Tirole (2000). Liquidity and risk management. Journal of Money, Credit and Banking, 295319.

Ioannidis, C., D. J. Pym, and J. M. Williams (2013). Sustainability in information stewardship: Time preferences, externalities, and social coordination. In The Twelfth Workshop on the Economics of Information Security (WEIS 2013). Available at: http://weis2013.econinfosec.org/papers/IoannidisPymWilliamsWEIS2013.pdf.

Ioannidis, C., D. Pym, and J. Williams (2009). Investments and trade-offs in the economics of information security. In R. Dingledine and P. Golle (Eds.), Proc. Financial Cryptography and Data Security 09, Volume 5628 of LNCS, pp. 148166. Springer. Preprint available at `http://homepages.abdn.ac.uk/d.j.pym/pages/IoannidisPymWilliams-FC09.pdf`.

Ioannidis, C., D. Pym, and J. Williams (2011). Information security trade-offs and optimal patching policies. European Journal of Operational Research 216(2), 434444.

Ioannidis, C., D. Pym, and J. Williams (2012).

Fixed costs, investment rigidities, and risk aversion in information security: A utility-theoretic approach. In B. Schneier (Ed.), Economics of Security and Privacy III. Springer. Proceedings of the 2011 Workshop on the Economics of Information Security.

Johnson, S. D. (2014). How do offenders choose where to offend? Perspectives from animal foraging. Legal and Criminological Psychology 19 (2), 193210.

Kirwan, G. and A. Power (2013). Cybercrime: The Psychology of Online Offenders. Cambridge University Press.

Kunreuther, H. and G. Heal (2003a). Interdependent security. The Journal of Risk and Uncertainty 26(1), 231249.

Kunreuther, H. and G. Heal (2003b). You only die once: Managing discrete interdependent risks. Technical report, National Bureau of Economic Research. NBER Working Paper 9885.

Laffont, J. J. (2008). Fundamentals of Public Economics. MIT Press Books.

Lelarge, M. (2012). Coordination in network security games: a monotone comparative statics approach. MacMinn, R. and J. Garven (2000). On corporate insurance. In Handbook of insurance, pp. 541564. Springer.

Mayers, D. and C. W. Smith Jr (1987). Corporate insurance and the underinvestment problem. Journal of Risk and Insurance, 4554.

Moore, T., R. Clayton, and R. Anderson (2009). The economics of online crime. Journal of Economic Perspectives 23 (3), 320.

Motahari-Nezhad, H., B. Stephenson, and S. Singhal (2009). Outsourcing business to cloud computing services: Opportunities and challenges. Technical report, Hewlett Packard Laboratories Working Paper HPL-2009-23.

Myerson, R. (1991). Game Theory: Analysis of Conflict. Harvard University Press.

Pal, R., L. Golubchik, K. Psounis, and P. Hui (2013). On a way to improve cyber-insurer profits when a security vendor becomes the cyber-insurer. In IFIP Networking Conference, 2013, pp. 19. IEEE.

Paulsson, T. and R. Sproule (2002). Stochastically dominating shifts and the competitive firm. European Journal of Operational Research 141 (1), 107112.

Pauly, M. (1974). Over-insurance and public provision of insurance: The roles of moral hazard and adverse selection. Quarterly Journal of Economics 88 (1), 4462.

Pearson, S. (2009). Taking account of privacy when designing cloud-computing services. Technical report, Hewlett Packard Laboratories Working Paper HPL-2009-54.

Png, I., C. Tang, and Q.-H. Wang (2006). Hackers, users, information security. In Workshop on the Economics of Information Security 2006. WEIS.

R. Dingledine and P. Golle (Eds.), Proc. Financial Cryptography and Data Security 09, Volume 5628 of LNCS, pp. 167183. Springer.

Rabin, M. (2000). Risk aversion and expected-utility theory: A calibration theorem. Econometrica 68 (5), 12811292.

Raviv, A. (1979). The design of an optimal insurance policy. The American Economic Review, 8496.

Rothschild, M. and J. Stiglitz (1976).Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. The Quarterly Journal of Economics 90 (4), 629649.

Schlesinger,H.(1981). The optimal level of deductibility in insurance contracts. Journal of risk and insurance, 465481.

Schneier, B. (2001). Secrets and Lies: Digital Security in a Networked World. John Wiley and Sons.

Scott, R. C. and P. A. Horvath (1980). On the direction of preference for moments of higher order than the variance. The Journal of Finance 35 (4), 915919.

Shapiro, C. (1989). Theories of oligopoly behavior. In R. Schmalensee and R. Willig (Eds.), Handbook of Industrial Organization, Volume 1, Chapter 6. North-Holland.

Shavell, S. (1979). On moral hazard and insurance. Quarterly Journal of Economics 93 (2), 541562.

Shavell, S. (1987). Economic Analysis of Accident Law. Harvard University Press.

Straub, D. W. and R. J. Welke (1998). Coping with systems risk: Security planning models for management decision making. Management Information Systems Quarterly 22(4), 441469.

Tirole, J. (1988). The Theory of Industrial Organization. MIT Press.

Tirole, J. (1998). Theory of Industrial Organization. MIT Press.

Varian, H. (2004). System reliability and free riding. Available at: http://people.ischool.berkeley.edu/hal/people/hal/papers.html.

Varian, H. (2010). Intermediate Microeconomics: A Modern Approach. WW Norton & Co.

Williamson, O. E. (1989). Transaction cost economics. Handbook of industrial organization 1, 135182.

Zhang, D. and T. Zhang (2012). Optimal portfolio of corporate investment and consumption under market closure. International Journal of Business 17 (1), 25.

Cyberinsurance and Policy Response: Self-Protection and Insurance with Endogenous Adversaries

Fabio Massacci, Joe Swierzbinski and Julian Williams

The growth in corporate insurance contracts that provide liability coverage in the event of a security breach to the firms information systems has been marked. Lloyds of London reports that the US cyberinsurance market could be as large as $1.5 Billion or 2% of the corporate insurance market. The effect of the presence insurance on the behaviour of the individuals or firms purchasing coverage has been of considerable interest in the academic literature for more than four decades. However, 'cyberinsurance' has been heralded as a potential mechanism for efficiently valuing the cost of cyber attacks on corporations, an inherently difficult task and to act as a substitute social coordinator internalising the inherent externalities incumbent to the realm of information security. This paper outlines a one period model with heterogeneous firms, with induced risk aversion from their corporate officers facing losses from cyber attacks conducted by strategic adversaries. We demonstrate that whilst the presence of actuarially fair insurance increases the aggregate utility of target firms, the presence of insurance is not a substitute for a social planner coordinating security expenditure. Furthermore, we show that when insurance is provided by a monopolist mandating firms security expenditure (as has been proposed) aggregate security expenditure is predicted to fall dramatically.

KEY WORDS: Insurance, Risk Management, Cyber Attacks

1. INTRODUCTION

Cyberinsurance provides coverage in the event of a successful attack on the information infrastructure of an organization.[4] This paper outlines a game played between a group of targets who invest in defensive expenditure to reduce the risk of a successful attack by one or more attackers. Attackers are modelled as criminals with fixed costs who engage in a competition to infiltrate and expropriate valuable information from the targets. Into this mix we introduce a variety of insurance contracts and behavioural restrictions that the insurer can then impose on the target. Our main result is that neither monopolist insurers nor a fully competitive insurance market have the incentive to reduce inherent externalities within the market. In contrast, the monopoly insurer, acting rationally, would be

[1] University of Trento, Italy, Fabio.Massacci@unitn.it
[2] University of Aberdeen, UK, j.swierzbinski@abdn.ac.uk* (Professor Swierzbinski is a guest contributor on the policy paper compendium.)
[3] University of Durham, UK, julian.williams@durham.ac.uk

[4] For a typical cyberinsurance contract see for instance `http://www.hiscox.co.uk/shared-documents/Hiscox-Business-Insurance-E-risks-policy-wording-03-11-6076.pdf`, accessed September 2014. This is a policy aimed at a small medium enterprise and is designed to limited liability up to £750,000 (approximately $1.2 million). The coverage includes direct losses of time and information assets as well as follow on legal costs incurred by the firm for loss of customer records and breaches of data protection. For instance Page 4 Clause (c) indicates that the insurance will pay for "*... your unauthorized collection or misuse of any data concerning any customer or potential customer of yours which is either confidential or subject to statutory restrictions on its use and which you obtained through the internet or extranet*".

positively assuaged towards inflating the cyber threat as long as they can identify the actuarially fair price of insurance risk and the maximum quote premium they can charge to risk averse targets.

The impact of moral hazard and adverse selection in the presence of insurance has a long history of investigation in economics, see for instance Pauly (1974); Shavell (1979, 1987); Cornes and Sandler (1996); Freeman and Kunreuther (1997) and Binmore (2005) for an eclectic set of examples that directly relate to the notions of public policy, liability sharing and insurance considered herein. In contrast, the impact of the presence of insurance and how the collective behaviour of victims of crimes can influence the aggregate behaviour of criminals generating the risks that are being insured, has far lower profile in the literature, although Cornes and Sandler (1996) provide a materially similar treatment to our cyberinsurance case. We can think of these effects as moral hazard and adverse selection effects that are once removed from the actions of the insuree, i.e. not the direct influence of target behaviour on target risk, with and without insurance, ceteris paribus, but the impact that changes in the aggregate behaviour of the pool of externally-insured or self-insured have on the risk vectors generating the distribution of losses. Furthermore, adjustments in external environmental conditions have been shown to affect the distribution of insurance claims and their legitimacy, see for instance Dionne and Wang (2013) for auto insurance fraud. Our attacker externality shares several similarities with the incentive to defraud suggested in Dionne and Wang (2013). In that there is a systematic factor that varies across all targets. However, in our case this is fully endogenous, rather than driven by an external macroeconomic effect.

Some discussion has occurred in the insurance literature from theft, for instance if large numbers of households buy burglar alarms, the aggregate cost of being a burglar increases as the need to invest in more specialist cognitive skills to bypass the alarms or the search for vulnerable homes becomes more time consuming. As such, if households with insurance are either required or strongly incentivized to buy alarms then this externality may reduce aggregate costs of insurance as less burglars are in the market for burglaries. Theoretically, this would then decrease insurance premiums as the actuarial risk of a payout decreases. However, it is likely that factors linking rate of return on burglary to burglar job choice decisions is highly inelastic; there are high costs associated with changing career and the stigma of prior convictions can result in substantial abatement costs when choosing an alternative career path.

The background risk of being a victim of burglary will, in-the-main, be a function of security choices of the target and background exogenous factors such as the location of the property and regional and national crime trends. The latter components of the risk model are materially unaffected by the influence of insurer and insuree actions in way that affects the individual and aggregate behavior of the criminals generating the risks, for the current period of coverage for a standard insurance contract, usually one year.

The role of adjusting aggregate behavior by mandating behavior on the insuree is mostly the remit of the public policy maker. For instance, the policy maker can collect higher taxes to invest in higher levels of physical security and increase the costs for criminals so as to dissuade them from choosing to engage in criminal behavior that results in insurance losses.

From the perspective of fraud and theft activities against firms, the focus of this paper, cybercrime has a number of different characteristics to the more traditional forms of crime. For instance, the choice of a software engineer to work on either malicious software or software with a more legitimate business purpose is simply a matter or re-tasking code. Most cyber criminals are anonymous (indeed this is the name of a cyber criminal group), therefore the decision to work as a either a 'white-hat' or a 'black-hat' is simply a cost benefit analysis that assesses the opportunity costs, risk of detection and time investment between these two roles. There is recent evidence to suggest that the decision of a hacker to enter into criminal activity is fragile. Johnson (2014) illustrates that much of the online crime is based on spatial opportunity. Spatial in the cyber crime sense is in terms of ease of access to particular systems and opportunity to illicitly monetize that access for personal benefit.

This is further supported in Kirwan and Power (2013) who indicate that the 'consistency assumption', there are consistent features to crimes committed by a single individual, is supported by evidence from criminals who have conducted online fraud. However, the 'homology assumption', individuals exhibit similar personality traits across their 'normal' and criminal behaviors is less well supported. Indicating that the decision to switch to online criminal behavior may be quickly reversed. Therefore the impact of the 'herd' choices of targets on attacker behavior may well be very significant in determining the aggregate amount of 'cyber-risk', the produce of losses incurred from cyber attacks and their likelihood, that is determining the individual cost of insurance. As such the reaction of criminal attackers on a corporate network to changes in individual and collective security could play an important role in the level of cyber-risk to be priced into an insurance contract.

In this paper we provide a comprehensive analysis of the impact of this risk on the standard insurance setting. The current importance of this topic is difficult to understate. Currently, in the cyber-crime research community a strand of thinking indicates that the presence of cyberinsurance firms (or firms offering insurance-like contracts, see (Pal et al., 2013)) will lead to a 'better' appreciation of risks and provide public policy-makers with the information needed to appropriate value their enforcement and regulatory options, see Schneier (2001) Chapter 5 for an early commentary and Moore et al. (2009) for discussion from the software engineering community. An interesting observation is that in the cyber-security literature, see Pal et al. (2013) as a good example, there is already an understanding that well informed cyber security vendors acting as insurers can extract considerable surpluses.

The remainder of this paper is organized as follows §(2) provides a brief analysis of the relevant contemporary literature on insurance with endogenous risks. §(3) outlines the base-case assumptions for our targets and attackers in the absence of insurance. §(4) introduces an efficient insurance market providing actuarially fair insurance, whilst §(5) recomputes the equilibrium choices of targets and attackers when insurance is provided by a single monopolist extracting a surplus from the targets risk aversion. Finally §(6) provides some simple numerical examples mixed with real case-studies from the early period of cyberinsurance provision and §(6) provides some conclusions and future directions.

2. BACKGROUND AND RELATED WORK

The role of competitive and non-competitive insurance markets in economic decision making is extremely old, for instance in the Code of Hammurabi (1772 BC) liability between traders was covered by a premium with a deductible and aggregate choices of these traders in demanding particular safety features on boats affected the aggregate premium. However, the lineage of our modelling approach is more appropriately based around Arrow (1974) and Rothschild and Stiglitz (1976) who introduced the modern mathematical approaches to dealing with moral hazard and adverse selection when designing insurance contracts and Ehrlich and Becker (1972) that delineates the concept of self protection and self-insurance versus external insurance.

The presence of an appropriate level of deductible for insurance against losses when the risk generating function was solely a combination of the behavioural choices of the insured and the exogenous background risk is comprehensively addressed in Raviv (1979) and Schlesinger (1981) which now form the part of the canon of text-book treatments on this subject. However, the key focus has been on the impact of insurance on the individuals behaviour and the optimal contract design (deductible, screening, behavioural requirements, monitoring) and not on the background risk process generating the need for insurance.

The differentiation between an individuals demand for insurance and a firms demand for insurance is addressed in Mayers and Smith Jr (1987). Corporate insurance differs from insurance for individuals as corporate stockholders (in diffuse ownership environments) can diversify away insurable risk. Therefore whilst owners may be risk averse themselves, it does not explain the demand for corporate insurance for a value maximizing firm. That corporate officers buy insurance to hedge against risks to their own positions is empirically investigated from a legal perspective in Baker and Griffith (2007), who uses the demand for corporate liability insurance as a predictor of firms corporate governance risk. Indeed, Griffith (2006), again from the law perspective, argues that the SEC should mandate insurance details of officers and directors liability insurance policies.

That firms demand in significant liability coverage cannot be in question. Specifically in the domain of Cyber liability insurance Lloyds of London in 2014 reported that £75 million of insurance premiums were paid by UK small medium enterprises. Furthermore, several media outlets have reported that the US market for cyberinsurance may approach

$1.5 billion in 2015.[5] Caillaud et al. (2000) suggests that risk-neutral firms will demand insurance as they are 'induced to risk-aversion' as the disclosure of 'accident-losses' that deteriorate the profitability of projects is private to the firm and costly audit is needed to demonstrate that the 'accident' was something that the agents managing the firm could not control. MacMinn and Garven (2000) also argues that the demand for corporate insurance stems from the firm choices mimicking the risk averse behaviour of its corporate officers. Finally, Holmström and Tirole (2000) further the argument that the mix of contracting and agency costs can induce risk aversion in corporate decision making.

Given that a corporation is a diffuse entity, determining the correct typology of utility function to use in a quantitative analysis may be difficult. In an important contribution Grossman and Hart (1982) deliberate on how corporate decision making is driven by the characteristics of the corporate officers and their incentives. That corporate officers prefer not to take risk-neutral bets is a fairly well understood phenomena. Transactions costs in changing from one firm to another means that corporate officers will seek to hedge bets as bankruptcy injects a shadow of the future problem for the officer. In Williamson (1989) the vertical structure of a firm and the inherent behavioural nature of financial decision making within are reviewed, along with a summary of empirical tests. Frictions are determined to be core drivers in how corporate decision making falls more into line with the behavioural characteristics of the firms officers than that of a strictly risk neutral entity working on behalf of highly diversified owners.

Taken together the assumption that firms exhibit some form of hyperbolic-absolute-risk-aversion (HARA), either in the form of constant or relative risk aversion appears not only plausible, but probable. For compensation contracts Hemmer et al. (1999) argues that using HARA type utility to describe manager preferences provides a reasonable approximation of the true decision making function. More recently, Zhang and Zhang (2012) argues that HARA type utility functions are the appropriate choice for modelling corporate decision making for portfolio selection. Finally, Arrow and Priebsch (2011) presents and example of risk averse firms represented by HARA type utility functions in relation to the decision making of firms and public-policy makers.

The use of HARA type utility functions to describe the decision preferences of firms has been discussed in Asplund (2002), who conclude that risk aversion in managers inundates itself into corporate decision making due to transactions costs for managers shifting from one firm to another in the event of large losses.

However, in a more recent contribution Chen et al. (2011) suggests that for more complex decision making, non-monotone functions may be more useful when considering decision making under uncertainty for very large and very small lotteries. Given that corporate decision making is often diffuse this approach may be appealing when taking into account the induced utility function of the firm. The fact that most corporate liability insurance is on a large scale (e.g. for property, plant and intellectual assets) suggests that some of the critique of expected utility theory suggested in Rabin (2000) may not apply. The major concern of Rabin (2000) is in regard to decision made under expected utility maximization for lotteries with stakes that are valued near the origin. It is an interesting paradox that most decision making by individuals, in an experimental sense, is conducted near the origin where small variations in choices leads to large fluctuations in the implied properties of the individuals utility function. Whilst most corporate decision making on the purchase of insurance and risk taking takes place for very large stakes well away from the problematic region, however our major theories of industrial organization suggest that firms should in fact, in general, be risk neutral.

Induced risk aversion in relation to firms demand for insurance, in particular for property, asbestos and pollution insurance (inherently demanded by corporations) appears to be a regular source of income for insurance companies. Financial Times reported that a '...well-regarded insurance analyst, who declines to be named, says [sic]:' "*The ideal scenario this year is we have some hurricanes.*"[6] The context of this comment is specifically in relation to the association between the realization of events,

[5] See report by Ian Allison, September 29, IBTimes https://uk.news.yahoo.com/lloyds-london-reveals-heightened-cyber-risk-boosts-profits-084212923--finance.html#TH6gQgO for an overview.

[6] "Struggle to Increase Rates Hits Lloyds", Alistair Gray, Financial Times, March 28, 2012 http://www.ft.com/cms/s/0/6fbb6cf4-78ea-11e1-88c5-00144feab49a.html?siteedition=intl#axzz3EEWQiyES. Last accesses October 2, 2014.

perceived increases in threat from liability claims and the subsequent demand for insurance, by firms with a need to satisfy firm level risk aversion.

Precautionary saving and the device of precautionary insurance is discussed at length in Eeckhoudt and Schlesinger (2006). Risk aversion, downside risk aversion (or 'prudence') and 'temperance' are categorized by the preference for certain types of lotteries over other alternative lotteries. The measurement of prudence versus imprudence is based on the sign of the third and fourth derivatives. A mechanism that is analogous to the higher moments literature in asset pricing, see Scott and Horvath (1980). Unifying these two strands of the risk taking and utility theory literature is Deck and Schlesinger (2010), who show that individuals choosing between lotteries exhibit prudence and temperance in a laboratory setting. Unfortunately, the literature on industrial organization and corporate decision making provides little insight into the translation of the personal preferences of corporate officers to the realized outcome in terms of the induced utility of the firm. Paulsson and Sproule (2002) provide some theoretical results on risk management for firms and allow the utility function that implies the preferences of the corporation, as an entity, to include the sign restrictions of Scott and Horvath (1980), implying both prudence (the third derivative being always positive) and temperance (the fourth derivative being negative), with imprudence and intemperance having the opposite sign restrictions.

It therefore appears reasonable to consider both the prudent and imprudent cases when considering firm risk taking in particular. Given the demand for corporate liability insurance the concept of precautionary investment to mitigate first and second order risk factors appears to be reasonable for cyberinsurance. Indeed, as Pal et al. (2013) very carefully illustrates, when cyber security vendors are able to use their population level knowledge of security risk versus potential targets and insuree's limited knowledge, significant gains are possible for a vendor that can act as a quasi insurer, for instance by guaranteeing up-time or ensuring transactions. In this case it wold appear that the targets optimal approach at the aggregate level would be to engage in carefully coordinated self protection.

Cyberinsurance has only recently been made available widely. It is therefore useful to gain some insight into the likelihoods of losses and their magnitude. Table I provides a series of quotes from a major insurer for a range of different firms. Using this data we have made some illustrative calculations assuming that the firm exhibits preferences characterised by constant relative risk aversion, in this case an iso-elastic power utility function. By dividing the quoted premium by the coverage limit we can derive the probability of an incident, in one year, under the assumption that the quote is actuarially fair. We can see that the highest probability under this assumption is for financial and E-commerce firms as 3.7%. However, this is of course a misleading value as the insurance market in this area is, in most likelihood, very far away from being actuarially fair. We therefore further assume that the insurance company, as a near monopolist, can charge a monopoly price up to the break-even of expected utility for the firm versus the certain utility in presence of insurance.

We compute this numerically for risk aversion coefficients of 0.1 and 1 with losses are relative to annual total revenue, which is a standard approach for corporate liability insurance. We choose 0.1, as a check, the probability should be very close to the actuarially fair insurance (and this is the case for all organizations in this sample). For a risk aversion coefficient of 1, the probability of an event with a successful claim for the Financial and E-commerce sector only reduces the implied probability of attack very slightly. However, for several of the firms, the implied probability of a claim event decreases from 0.6% to 0.1%.

In a final experiment we take two industry reported payout ratios 10% and 50% on premiums and compute the minimum implied constant relative risk aversion for firms with iso-elastic power utility.[7] For the widely quoted 10% payout, all firms have a risk aversion coefficient above unity, however for a more reasonable 50% payout several firms have minim relative risk aversion coefficients close to a half.

We will show that our results generally hold when losses for security events are large, as would be expected in the case of corporate security liabilities. However, as Rabin (2000) suggests, the choices made by individuals assumed to be an expected-utility maximizer with a concave utility function can produce results that run counter to intuition and

[7] See: "Cyber insurance market tempts new participants" by Alistair Grey, Financial Times October 6, 2014. http://www.ft.com/cms/s/0/69db580c-4d37-11e4-8f75-00144feab7de.html#axzz3G9LENdhM.

Table I . Selection of purchased annual cyberinsurance contracts across a variety of commercial settings.

Industry	Revenue US$	Limit US$	Premium US$	Premium Divided by Coverage	Implied Probability of Loss CRRA(0.1)	Implied Probability of Loss CRRA(1)	Implied CRRA @10% Payout	Implied CRRA @50% Payout
Healthcare (HC)	25,000,000	1,000,000	12,900	0.01290	0.01287	0.01264	1.03917	1.03816
Education	25,000,000	1,000,000	6,000	0.00600	0.00599	0.00588	1.03916	0.97951
Financial	100,000,000	1,000,000	37,000	0.03700	0.03698	0.03682	1.02286	1.02265
Retail	50,000,000	1,000,000	26,000	0.02600	0.02597	0.02575	1.05891	1.05838
E-commerce	50,000,000	1,000,000	37,000	0.03700	0.03696	0.03664	1.05892	1.05845
Restaurant Chain	50,000,000	1,000,000	10,000	0.01000	0.00999	0.00990	1.00250	1.00204
Manufacturing	100,000,000	10,000,000	50,000	0.00500	0.00497	0.00475	1.07686	1.02014
HC IT Prov.	1,200,000	5,000,000	15,900	0.00318	0.00289	0.00078	1.16032	0.59353
HC SaaS Prov.	1,500,000	5,000,000	30,420	0.00608	0.00553	0.00149	1.16051	0.59411
HC Rec. Prov.	5,000,000	1,000,000	8,010	0.00801	0.00792	0.00718	1.13646	1.12896
Data Hosting	200,000	1,000,000	2,750	0.00275	0.00254	0.00101	1.16002	0.75832
HC IT Cons.	150,000	1,000,000	3,298	0.00330	0.00305	0.00121	1.16005	0.75849
HC Data Analysis	20,000	2,000,000	4,900	0.00245	0.00224	0.00075	1.17211	0.66403
e-Waste Co.	1,500,000	1,000,000	3,564	0.00356	0.00340	0.00215	1.14896	1.07566
Psych. Office	1,000,000	1,000,000	1,600	0.00160	0.00148	0.00059	1.15996	0.75795
Doctors Office	700,000	500,000	649	0.00130	0.00123	0.00073	1.13302	0.98006
Online Retailer	500,000	1,000,000	1,100	0.00110	0.00102	0.00040	1.09063	0.75725
Prof. Cons, Serv.	400,000	1,000,000	1,200	0.00120	0.00111	0.00044	1.09064	0.75728
Hospital	170,000,000	5,000,000	42,000	0.00840	0.00839	0.00828	1.04972	0.99632
Data Stor. Cent.	15,000,000	20,000,000	120,000	0.00600	0.00542	0.00111	1.15218	0.54269

Note: The first column identifies the industrial sector for which the corporate liability insurance has a specific cyberinsurance clause. The second column denotes the reported revenue (approximated by the insurance company) of the organisation in one year. The third column provides the level of coverage for each organisation while the fourth column reports the payable insurance premium. From this data we have computed the following: column five reports the ratio of the premium to the coverage (if the insurance was actuarially fair, then this would be the probability of an event, assuming one claim per year). Assuming that the quote is not actuarially fair, but the maximum chargeable to a risk averse target, we have computed the implied probability of an attack resulting in a claim for a constant relative risk averse target, when the coefficient of relative risk aversion (CRRA), computed by $R(w) = wU''(w)/U'(w)$ is 0.1 or unity, presented in columns six and seven respectively. This is numerically solved using a standard iso-elastic utility function. Finally, we compute the implied minimum coefficient of risk aversion assuming two payout ratios, first 10%, which is the reported pay-out ratio on cyber insurance and second 50% as a comparator. The quotes are for actual purchased coverage from by a major insurer.

experimental results on choices of lotteries when the stakes are very small.

3. THE SELF-PROTECTION MECHANISM

We will follow Caillaud et al. (2000) and assume that target firms exhibit 'induced' risk aversion that is generated by the convex preferences of the corporate officers of the firm. From this point onwards, when we refer to 'target-preferences' or 'target-utility' we are referring to the induced preferences exhibited by the firm reflecting the 'transferred' preferences from the actual corporate officers making the financial decisions to the revealed actions of the firm. We are, of course, not implying that firms have behavioral characteristics as individual entities.

3.1 Targets' Model

In our most general case we assume that there is a population of N targets heterogeneous in both risk preferences and endowments of assets, although all are strictly risk-averse.

Each target $i \in \mathbb{N}_+$ has an endowment of assets $W_i > 0$ and in the event of a successful cyber attack is subject to losses L_i, where $L_i > 0$. The utility function of the i target over a random consumption variable w, for wealth, is denoted by $U_i(w)$. Our model is a single period universe, where there are two outcome states: *target has been successfully attacked* incurring loss L_i and *a target has not been successfully attacked* incurring no loss. We denote by $U_i'(w)$ and $U_i''(w)$ the first and, respectively, the second derivatives of $U_i(\cdot)$ w.r.t. the wealth variable w.

Assumption 1: The 'induced' utility function of the $i \in \{1, \dots, N\}$ firms, $U_i(w)$ is an at least thrice differentiable von Neumann Morgenstern utility function where for all w it is $U_i'(w) > 0$ and $U_i''(w) < 0$. We place no specific sign restriction on the third derivative, however the domain of certain results depend on whether the targets third derivative exhibits prudent behavior i.e. for all w it is $U_i'''(w) > 0$.

The consequence of prudence is that the first order derivative of the utility function is upward convex, anti-monotone (downward sloping) and satisfies $\mathbb{E}[U_i'(w)] \geq U_i'(\mathbb{E}[w])$. Whilst the standard axioms of risk aversion are well known, the concept of prudence has been restricted in discussion to topics relating to precautionary saving and insurance. It is therefore worth quickly recapping the concept of prudence from a utility perspective. This is a broad class of functions that subsumes the HARA type utility functions where $U_i'(w) = -U_i''(w)/(aw+b)$, for non negative coefficients $a \geq 0$ and $b \geq 0$ where with $a = 0$, $b > 0$ we have a CARA utility function, whereas with $a > 0$, and $b = 0$ will yield a CRRA utility function[8].

Each target may make a security investment x_i, where $0 \leq x_i < L_i$ and is targeted by a number of attackers n_i where $0 \leq n_i < \infty$. The probability of a successful attack from the viewpoint of the target, is determined by the level investment and the number of attackers focussing on the particular target and is denoted by $\sigma_i(n_i, x_i)$. For each target there are only two outcome states, 'successfully-attacked' and 'unsuccessfully-attacked', the probability of each state is therefore σ_i and $1 - \sigma_i$ respectively.

The expression $\Delta U_i(x_i)$ denotes the difference in utility between the state when target i has not been successfully attacked and the state when it has been successfully attacked

$$\Delta U_i(x_i) = U_i(W_i - x_i) - U_i(W_i - x_i - L_i). \tag{1}$$

For a given level of wealth w and a potential loss L_i, we define the *locally risk-neutral loss* as the function $\mathcal{L}(w, L_i) \in [0, L_i]$ such that

$$U_i'(w - \mathcal{L}(w, L_i)) = \frac{U_i(w) - U_i(w - L_i)}{L_i} \tag{2}$$

The economic interpretation of the value $w - \mathcal{L}(w, L_i)$ is the level of wealth between $U_i(w)$ and $U_i(w - L_i)$ at which a risk averse target is locally risk neutral for small changes in wealth in the sense of Pratt (1964). This function characterizes the form of the utility function in response to a potential loss L_i and is independent from the success probability σ_i. For example, for the exponential CARA functions the value of $\mathcal{L}$ is constant for any value of wealth and only depends on the constant of absolute risk aversion and the loss L_i. The geometrical interpretation of $w - \mathcal{L}(w, L_i)$ is the point of the utility curve where the tangent to the utility curve is parallel to the risk neutral line between the points on the utility curve corresponding to w and the point $w - L_i$.

For a risk neutral target this function is not well defined as the right side of the equation is identically equal to 1 for all value of wealth. For a risk averse target[9] this function is uniquely defined by the Mean Value Theorem of the calculus.

We do not impose explicit externalities which directly link the investment x_i in one target i to the probability σ_j of a successful attack on another target j. The externalities in our approach are entirely driven by the aggregate reaction of attackers.

Early research in Gordon and Loeb (2002) has provided a general set of assumptions of the functional form of $\sigma_i(x_i, n_i)$, specifically that it is strictly decreasing in x_i and strictly increasing in n_i. However, for the purposes of our analysis we will impose several other assumptions to ensure tractability.

Assumption 2: The function $\sigma_i(\cdot)$ should be at least twice differentiable in x_i and n_i.

Assumption 3: $\sigma_i(\cdot)$ is continuous strictly increasing when the number of attackers against target i increases, for all n_i it is $\partial\sigma_i/\partial n_i > 0$, ceteris paribus. In the absence of attackers the probability of a successful attack is zero, therefore when $n_i = 0$, $\sigma = 0$, $\forall 0 \leq x_i < \infty$.

Assumption 4: $\sigma_i(\cdot)$ is continuous strictly decreasing with increasing investment by target i in

[8] This is a well known property of HARA utility functions see for instance Scott and Horvath (1980) for an application to portfolio selection or Paulsson and Sproule (2002) for an experimental economics view. Consider the first order derivative for both sides of the equation $U_i''(w) = -(\frac{U'''(w)}{aw+b} - a\frac{U''(w)}{(aw+b)^2})$. We can then apply transitive linearity to the functional form for absolute prudence such that $U'''(w) = -(aw+b)U_i''(w)(1 + \frac{a}{(aw+b)^2})$. Since $U_i''(w) < 0$, therefore by construction $U'''(w) > 0$.

[9] This is also true for a risk seeking target.

security investment, for all x_i it is $\partial\sigma_i/\partial x_i < 0$, ceteris paribus.

Assumption 5: The rate of reduction in $\sigma_i(\cdot)$ with increasing x_i, from A.4, is strictly decreasing with increasing defensive expenditure, for all x_i it is $\partial^2\sigma_i/\partial x_i^2 > 0$

Assumption A.3 simply indicates that the likelihood of at least one successful attack increases as the number of attackers, n_i, acting against that particular target increases. Assumption A.4 restricts the functional forms driving the effectiveness of investment in reducing the likelihood of at least one successful attack to only continuous functions. Whilst at the micro-level the choices of individual firms on investment in security controls are somewhat more atomic, for the population of the targets the continuous function assumption ensures tractability and ease of exposition.

Assumption A.5, 'diminishing-marginal-returns-to-security-investment', DMRSI, is a well understood property of for implementing security controls on complex information systems. The number of obvious flaws that an attacker can exploit falls rapidly with initial security investment. However, once the obvious controls are implemented the cost to the target for equivalent marginal reductions will increase rapidly. This is because 'high-security' controls have a larger number of elements to their aggregate costs (such as loss of continuous availability of key systems and soft controls in vetting staff administrator access as opposed to simply implementing software solutions that do not deteriorate the 'normal' or 'expected' operation of the firms information infrastructure.

3.2 Attackers' Model

Cyber attacks are conducted by a pool of attackers. In the event of a successful attack, the successful-attacker realizes a reward R_i. As with most insurance cases, the reward $R_i > 0$ is assumed to be far lower than the loss L_i, however as the unit of account for attackers and targets is assumed to be fundamentally different this assumption does not bind.

To enter the lottery for the reward R_i each attacker must spend a cost C_i, for tractability we assume that all attacks are independent[10].

We restrict our attention to non-cooperative attackers making single entry decisions.

Assumption 6: Attackers have fixed costs of entry c_i, are risk neutral and make binary attack or no-attack decisions. We will further assume that when looking at aggregate number of attackers per target $c_i = C$.

Assumption 7: Attacker-target matching is fully-degenerate. The probability matching of a given j attacker to the i target is random with a uniform distribution across targets (i.e. the probability is $1/N_T$ where N_T is the total number of targets).

Assumption 8: The scalar $R_i \geq 0$ is the reward for the *first-winner-takes-all* attacker to successfully attack target i. In the event of a successful attack on target i the successful attacker does not share this reward with other n_i attackers and at this point no further attack will generate any reward. We denote the *rate of return* on C for a given reward R_i for an attack on the i target as $R_i/C = \rho_i$.

Assumption A.6 is motivated by the current bimodal distribution of attacks as summarized by Tankard (2011): specialized targeted attacks and large scale, albeit specialized, attacks. Targeted attacks – also called Advanced Persistent Threats Mandiant (2013) — are typically carried by well-funded adversaries (in many cases national governments) for either cyber-espionage, intellectual property theft, or political motives (See Li et al. (2011) for a technical analysis of a latter). They are preceded by a phase of intelligence aimed at gathering information on specific individuals, and includes significant elements of social engineering and customization. In such cases discussing the role of cyberinsurance has no sense: what would be the coverage for the theft of the new Ferrari's car maker design? or the premium for insuring against losses due to illegal access to the email of a military contractor with secret clearance in Afghanistan? In contrast, large scale attackers may specialize in a "business sector" (such as bank or gaming companies) but would not specifically

[10] One can think of the cost C_i as being the total cost for the attacker of probing the targets defenses, developing or buying suitable exploits, infiltrate the targets information system by deploying the exploits, and successfully exfiltrate the information assets needed to generate the reward R_i. The costs, include opportunity costs and risk of punishments if caught, in addition to any fixed costs in terms of purchasing software or equipment.

target the individual. For example, the rock-phish gang, allegedly one of the most successful compact of cybercriminal specialized in phishing (See Moore and Clayton (2007) for a discussion), loaded their web server with a large number of bank websites (ten or more) possibly customized by locale (e.g. all Italian or all British). Hackers may customize their attack with spear-phishing as this would improve their click-through rate (See Paper (2011) for some sample estimates) but would still need to target a generic user at the technical level.

To understand this issue consider the case of spear-phishing Hong (2012), which is one of the largest concern for companies. In these kind of attacks a specific message is send to a user via an email or a social media post in order to trick him to click on a link (where an exploit kit has been deployed) or to open an attachment (where a malware exploit have been embedded). The customization of the message is dependent on the attempted target and may require a specialized investigation into social networks and therefore increase the costs of the *preparatory* activities which are the only one where attackers may speculate about possible rewards. Such customization does not change the *technical* execution of the attack: the attacker has in general no way to know on which system the PDF file will be opened, he can only plan that *if* the PDF is opened with a vulnerable system it will then compromise that machine. Once, and if, the machine has been compromised the attacker then needs to engage in costly effort for the *conversion* activities in order to extract revenue from the compromised machine, that is only rarely successful on the individual, but is successful on the aggregate.

While doing preparing the attacks, the social engineering effort is an additional costs for the hacker, but this compensated by the much larger click-through rate (i.e. number of victims that actually do click on the malicious link). Traditional click-through rate are smaller for non-targeted campaigns than for specialized business-sector specific campaign (See Paper (2011) and Kanich et al. (2008) for a comparison). Still, even if the attack is targeted other social or demographics factors may still reduce the chances of a click-through Sheng et al. (2010). Therefore, the attacker cannot strategically chose an individual target and entirely invest the effort on it as only large numbers guarantee a chance of success. This justify our assumption A.7.

In terms of costs (Assumption A.6) we have already discussed how the social engineering effort is balanced by the improved click-through ratio. In terms of technical effort, the key issue is that attackers do not know a priori that their chosen target is vulnerable to a particular exploit and therefore have to work on “expected” configurations in the hope that “on average” it will succeed. For instance, Google reports that exploit kits represent two thirds of the threats against end users, Rajab et al. (2011). The Contagio web sites list hundreds of them. However, a number of empirical studies of the successful exploit kit market in Russia Kotov and Massacci (2013); Allodi and Massacci (2014) have found that each kit only exploits a handful vulnerabilities, on the order of ten to fifteen vulnerabilities per kits. Exploit kit providers are even turning to the cloud in order to be more resilient and reduce costs Nappa et al. (2013). Even malware that is apparently designed for geo political purposes by national governments (e.g. Stuxnet and Duqu) has a limited number of exploits in order to provide a profile that escapes detection (See e.g. the study of Duqu Bencsáth et al. (2012)). We can infer from this that the cost of attacks will be essentially constant $c_i = C$ for a business sector. Each attacker needs to develop a general collection of tools and the rate of return will therefore be determined from the rewards on the target side $\rho_i = R_i/C$.

Assumption A.8 ensures that whilst the attacker dynamics are tractable they stay true, in terms of aggregate behavioral characteristics, to the effects observed from the limited empirical evidence, see for instance Moore and Clayton (2007); Holz et al. (2009); Cho et al. (2010). Qualitative research in hacking communities, predominantly in Russia, have found that whilst hacking resources are shared or exchanged for relatively low amounts of money, the rewards from attacks are not normally shared within the hacker community. This is mostly due to the difficulty of monetizing the same reward twice Herley and Florêncio (2010); Florêncio and Herley (2010). In the general case, the reward R_i may be a function of the type of security investment and the type of target as some mitigation measures may decrease the value of the reward (in the physical domain this would be the staining of stolen cash from ATM). In general, the link between L_i and R_i is invariably difficult to properly correlate and these are treated as being separate units of account. The relationship between the moments of the financial outcomes for attackers and defenders in this case is purely based on the probabilistic assessment of the likelihood of success and assumed to be independent.

3.3 Unregulated Markets

In the first instance, let x_i and n_i be given exogenously. The expected utility of target i for a given pair (x_i, n_i), denoted $\mathbb{E}[U_i(x_i, n_i)]$, is therefore described by the following

$$\begin{aligned}\mathbb{E}[U_i(x_i, n_i)] &= (1-\sigma_i(x_i, n_i))\, U_i(W_i - x_i) \\ &\quad +\sigma_i(x_i, n_i) U_i(W_i - x_i - L_i)\end{aligned} \tag{3}$$

To illustrate the properties of the model, we first consider the case of the risk neutral target. In the insurance literature (see Section.2) the probability of the unfortunate event is normally considered as an exogenous parameter. In our scenario, the probability σ_i is partly determined by the strategic decisions of the agents (x_i) and this by itself has a major impact.

If the target i is risk neutral and therefore $U_i(w) = w$, its expected utility for a given level of security expenditure is simply the expected net monetary value of its assets

$$\mathbb{E}\left[U_{\text{risk neutral } i}(x_i, n_i)\right] = W_i - x_i - \sigma_i(x_i, n_i)L_i \tag{4}$$

However, the function σ_i is convex in the security investment argument (Assumptions A.4 and A.5) and therefore *the risk neutrality curve is no longer a straight line.* We illustrate this phenomenon in Figure **??** where we plot the available wealth w_i, after the security investment, of the target on the horizontal axis[11] and the expected utility on the vertical axis. The plot is obtained from a concrete illustrative σ_i-function that we describe in Section.6.

The straight 45-degree line through A would have been the "exogenous" risk neutral line if σ_i were exogenous (e.g. $\sigma_0 = \sigma_i(0, n_i)$). The concave curve through the points A and B is the expected utility of the target $\mathbb{E}[U_i]$. The convex curve at the bottom of the figure through points C and D describes how the expected loss $\sigma_i L_i$ changes as the remaining available wealth changes (after the security investment): if the target has decided not not invest any money in self-protection (point C where $w_i = W_i$ and therefore $x_i = 0$) its expected losses would be maximal. The expected losses will decrease as x_i increases and therefore w_i decreases.

The optimal level of defensive expenditure for a risk neutral target is obtained by maximizing the expected utility (point B in the figure). Since W_i is constant, in presence of a given number of attacker, this value is obtained by setting to zero the usual first order condition which leads to the following equation

$$L_i \frac{\partial \sigma_i(x_i, n_i)}{\partial x_i} = -1 \tag{5}$$

The choice of the optimal expenditure $x^\star$ (represented by the point E as a level of remaining wealth $w_i^\star = W_i - x_i^\star$) has also an interesting geometric interpretation. The point D on the expected loss curve is the point where the tangent to the curve would be parallel to the "exogeneous" risk neutral curve. Point F also has an interpretation in the model. It represents the overall optimal expenditure for unregulated risk averse target in presence of actuarially fair insurance. We discuss this issue later in Sect.4.1.

From Assumptions A.6 and A.7 we know that the j attacker solves their binary decision to attack condition under risk neutrality. If the population of potential attackers is sufficiently large, then, in expectation, as the only attackers who gain rewards are the first to succeed, the difference between expected reward and deterministic cost will be zero. At this point further attackers will have a negative net expected reward and will cease to engage in attacks.

We denote the number of attackers against the i target that satisfies the zero expected profit condition as n_i^*, when x_i is exogenous and this is given by the following expression

$$\sum_{i=1}^{N} R_i \cdot \sigma_i(x_i, n^*) \;=\; \sum_{i=1}^{N} n^\star \cdot C_i \tag{6}$$

Competition between attackers is particular intense in this framework and this entry condition generates a strong bound on the number of attackers, which leads us to our first proposition.

Proposition 1: On the equilibrium path the number of attackers per target n^* is bounded between the worst rate of return of attacks times the average success probability and the best rate of return times the average success probability.

$$\rho_{-}\bar{\sigma}(n^*) \;\leq n^* \;\leq \rho^{+}\bar{\sigma}(n^*) \tag{7}$$

where $\bar{\sigma}(n^*)$ is the average over all x_i for the equilibrium value of n^*.

[11] To use wealth on the horizontal axis it is convenient to use the transformation $w_i = W_i - x_i$ and therefore the expected loss is $\sigma(W_i - w_i, n_i)L_i$ where w_i is the current level of wealth.

Lemma 1: *On the equilibrium path, the number of attackers per target n^* is equal to the rate of return ρ_i weighted average probability of at least one successful attack for target i*

$$n^* = \frac{1}{N}\sum_{i}^{N} \rho_i \sigma_i(x_i^*, n^*) \qquad (8)$$

The proof follows easily from Proposition 1 by setting $\rho = \rho^+ = \rho_-$.

When all targets and attackers are identical, it is plausible to restrict attention to symmetric equilibria. At equilibrium both attackers and targets correctly forecast the choices of other players.

Corollary 1: *If all targets are identical the number of attackers per target n^* at equilibrium is equal to probability of success times the rate of return ρ*

$$n^* = \rho\sigma(x^*, n^*) \qquad (9)$$

The proof follows easily from Proposition 1 by setting $\rho = \rho^+ = \rho_-$.

In a symmetric Nash equilibrium (x^*, n^*) each target i selects the same level of defensive expenditure x^* that optimizes the expected utility of the target whereas the number of attackers per target n^* is determined by the free entry condition from Lemma 1.

$$\begin{aligned} x^* &= \arg\max_x\{\mathbb{E}[U(x, n^*)]\} \\ n^* &= \rho\cdot\sigma(x^*, n^*) \end{aligned}$$

we will now show that the presence of a public policy acting as a benevolent social planner the socially optimal level of investment $x_i^\dagger$ will be greater than x_i^*.

The first order condition for the expected utility of the targets can be expanded as follows:

$$\frac{\partial\mathbb{E}[U_i]}{\partial x_i} = -\mathbb{E}[U_i'] - \frac{\partial\sigma_i}{\partial x_i}\Delta U_i \qquad (10)$$

This decomposition contains a positive term and a negative term which capture the interplay between the risk-aversion of the target and the marginal effectiveness of self-protection expenditures. The first term is negative and capture the unwillingness of the target to increase its spending to counter marginal increases in risk. The second term increases with the increase in effectiveness of security expenditures, and is amplified by the difference in utility between the normal situation and the case in which the attack is successful.

The risk-aversion of the targets makes it possible to establish a general (albeit not tight) bound on the value of the optimal investment. We consider the ratio λ_i of the marginal rate of utility in the best case scenario when no loss is present ($w = W_i$) and the worst case scenario where the target has spend L_i in self-protection and has been nonetheless hacked ($w = W_i - 2L_i$).

$$\lambda_i = \frac{U_i'(W_i)}{U_i'(W_i - 2L_i)} \leq 1 \qquad (11)$$

Lemma 2: *For a given value of the number of attackers n, the marginal loss due to a successful attack at the equilibrium x_i^* of unregulated risk averse targets is bounded as follows*

$$-\frac{1}{\lambda_i} \leq L_i \left.\frac{\partial\sigma_i}{\partial x_i}\right|_{x_i = x_i*} \leq -\lambda_i \qquad (12)$$

Proof. See Appendix.

3.4 A Benevolent Social Planner Mandating Security Investments

Before introducing the insurance market, it is useful to ascertain the optimal investment policy that a fully informed benevolent social planner would mandate. By ‘benevolent’ we will adhere to a classical utilitarian definition. Hence the social planner will have a utilitarian social welfare utility function that respects the preferences of the population of targets. The planner’s action is to ‘mandate’ security investments for each target denoted $x_i^\dagger$ and we will assume that $x_i^\dagger$ is binding, measurable and enforceable.

Let the social planner’s preferences regarding the risks of cyber-attacks be described by an aggregate von Neumann-Morgenstern utility function of the form

$$U_P = \sum_{i=1}^{N} \nu_i U_i \qquad (13)$$

This utility function is a utilitarian social welfare function since it consists of the sum of the utilities of the individual targets weighted by the values ν_i assigned by the policy makers to the different targets. Without loss of generality, we normalize the weights so that

$$\sum_{i=1}^{N} \nu_i = 1. \qquad (14)$$

In practice the social planner is not necessarily a public-policy maker.

The policy maker is able to specify the levels of defensive expenditure for targets (compulsory security standards). In the special case where targets are identical and the weights ν_i are equal for all targets, the preferences of the policy maker and the targets coincide.

3.5 The Stackelberg Planner

Let us assume that the policy maker moves first and that all other actors (targets and attackers) make their choices in a second stage after observing the choice of the policy maker. As a result of this assumption the expected utility of the policy maker is determined by

$$\mathbb{E}[U_P] \; = \; \sum_{i=1}^{N} \nu_i \; \mathbb{E}[U_i], \tag{15}$$

where $\mathbb{E}[U_i]$ is specified in (3).

Since the policy maker mandates the appropriate level of defensive expenditure, x_i, for each target in stage one, targets have no longer have a decision to make and, hence, are not active players in the game without insurance. However, the payoffs of the targets are still important since the policy maker is assumed to choose the levels of defensive expenditure in order to maximize the expected utility in (15). In stage two of the game, each potential attacker chooses whether or not to participate in attacks against the population of targets.

We model the outcome as a subgame-perfect equilibrium of the corresponding two-stage game. In a subgame-perfect equilibrium, the choices of the policy maker in the first stage of the game must be optimal given the strategies of the players in the second stage. If the strategies of the actors in the second stage are optimal only for the policy maker's equilibrium choice of defensive expenditures, then the resulting equilibrium will just be a Nash equilibrium. The strategy of each actor in the second stage must also be optimal for each possible security investment of the policy maker in the first stage and given the strategies of all other second-stage players. See Tirole (1988) and Binmore (2007) for a further discussion.

For the strategies of potential attackers to be part of a subgame-perfect equilibrium, it is sufficient for the equilibrium number of attackers per target, n_i^*, implied by these strategies to satisfy (6) for each set of feasible defensive expenditures. This guarantees that potential attackers respond to different levels of defensive expenditure in such a way that they are always indifferent between participating or not participating in attacks.

Since the equilibrium number of attackers per target, n^*, adjusts to changes in the levels of defensive expenditure (Proposition 1), the policy maker must take this adjustment into account when determining the optimal choice in the first stage of each game.

We assume that the optimal choice of the policy maker satisfies the usual first-order conditions: $\partial\mathbb{E}[U_P]/\partial x_i = 0$ for all i under the constraint represented by (6).

Solving the first order condition and dividing by ν_i (the i quota of the reaction of the expected utility of policy maker to changes in x_i) yield the following decomposition which illustrates how the incentives of the policy maker may differ from the incentives of individual unregulated targets.

$$\begin{aligned}\frac{1}{\nu_i}\frac{\partial\mathbb{E}[U_P]}{\partial x_i} \quad = \quad & \frac{\partial\mathbb{E}[U_i]}{\partial x_i} + \\ & + \frac{\partial\sigma_i}{\partial x_i}\left(\Delta U_i(x_i) - \frac{1}{\nu_i}\Delta U_P(x_i)\right) \\ & - \frac{\partial n^*}{\partial x_i}\left(\mathbb{E}\left[\frac{\partial U_i(x_i)}{\partial n^*}\right] - \frac{1}{\nu_i}\mathbb{E}\left[\frac{\partial U_P(x_i)}{\partial n^*}\right]\right)\end{aligned} \tag{16}$$

The divergence between the Nash equilibrium levels of defensive expenditure and the levels of expenditure deemed optimal by the policy maker in the subgame-perfect equilibrium occurs because the individual targets neglect the effect of their expenditure on the behavior of attackers. In particular, each target ignores the beneficial effect the target's expenditure has in reducing attacks on other targets. The term (16) captures the gap between the difference in utility $\Delta U_i(x_i)$ of the individual target i and its pro-quota contribution to the global utility of the policy maker $\frac{1}{\nu_i}\Delta U_P(x_i)$. The utility of the policy maker is always larger because there is a beneficial effects on the other target: the increase in expenditure of target i makes the overall number of attacker n^* lower and therefore improves the utility of other targets j. The term (16) capture the sensitivity of the utility to an increase in the number of attackers. Since U_j is weakly concave for all j, the difference between the expected sensitivity to attackers $\mathbb{E}\left[\partial U_i(x_i)/\partial n^*\right]$ of the utility of the individual target i is less steep than target's i pro-quota overall sensitivity of the policy maker $\frac{1}{\nu_i}\mathbb{E}\left[\partial U_P(x_i)/\partial n^*\right]$.

THEOREM 1: *In absence of insurance the security investment $x_i^\dagger$ mandated by a benevolent social planner to risk averse target i is larger than the optimal security investment x_i^* that the same target i would have chosen in an unregulated environment.*

Proof. See Appendix 7.

Discussion of Theorem 1: That the benevolent social planner mandating investments can 'improve', strictly in an individual welfare sense, the outcomes for all targets is relatively straightforward and a well understood effect. In the absence of the social planner the equilibrium level of expenditure is arrived at by the targets traversing the curve generated by the aggregate number of attackers per target. when the number of targets is large, no single target can, by altruistically and unilaterally raising their defensive expenditure, reduce aggregate number of attackers. As such, their unilateral increase has only the effect of shifting them from their optimal expenditure in the presence of the average number of attacker n_i^*. From assumptions A.1 to A.7 for a given n_i^*, the optimal expenditure x_i^* is unique and all deviations from this expenditure are sub-optimal. However, the social planner explicitly accounts for the attacker reaction and by mandating expenditure across *all* targets attains a higher overall utility for each target, than each individual target could do by acting alone.

4. INTRODUCING CYBERINSURANCE CONTRACTS

We now introduce an insurance market that provides targets with coverage against losses from cyber-attack. In the first instance we will assume that the market for insurance is perfectly competitive providing actuarially fair insurance. We will then move on to the opposite case when a single monopoly insurer can extract a full surplus from targets.

Let the i target have an available insurance contract described by a pair (q_i, ℓ_i), which specifies the premium, or quote, q_i paid by the target i in the event of a loss, and the amount of the deductible, or excess, $\ell_i \leq L_i$ that will be left to be paid by the target if the successful attack takes place. Following the standard assumptions, the premium is paid upfront and its cost is incurred in both of the outcome states.

4.1 Unregulated targets

In absence of the benevolent social planner from §§(3.5) individual targets can freely choose a level of defensive expenditure x_i the level of defensive expenditure as well as whether to purchase an insurance contract specified by the tuple (q_i, ℓ_i). As such the i target's expected utility is therefore given by:

$$\begin{aligned}\mathbb{E}[U_i(q_i, \ell_i, x_i, n_i)] &= (1 - \sigma_i(x_i, n_i))\, U(W_i - x_i - q_i) + \\ &\quad + \sigma_i(x_i, n_i) U(W_i - x_i - q_i - \ell_i).\end{aligned} \tag{17}$$

For the purposes of our analysis herein, it is useful to impose certain assumptions on the information set and range of actions available to a provider of insurance. We will now outline carefully our assumptions in this respect.

4.2 The Cyberinsurance Market

An insurer provider issues a cyberinsurance contract that pays out and amount $L_i - \ell_i$ in the event of loss from a cyber attack.

ASSUMPTION 9: cyberinsurance companies are profit maximizers with a risk-neutral break-even requirement for the issuance of coverage at an actuarially fair price.

ASSUMPTION 10: The security expenditure of all targets x_i are fully auditable, both ex-ante and ex-post, and the investment in defensive expenditure is made with commitment.

ASSUMPTION 11: The insurance company can identify the aggregate number of attackers per target for a given level of defensive expenditure across the population of targets. Therefore the provider of the insurance can fully determine the actuarially fair value of insurance for the i target in the presence or absence of insurance.

Assumption A.9 is fairly standard in the insurance literature. Each individual insurance company will provide coverage to a minimum price, the actuarially fair premium for a given deductible ℓ_i. When ℓ_i is zero, the insurer is providing full coverage of losses L_i. Assumption A.10 is possibly a unique case for cyber-security. Whilst The security features of the information systems used by the target firms are relatively straightforward to assess Marcella Jr and Greenfield (2002). For significant attacks for which uninsured losses were sufficient to bankruptcy of the attacked target forensic investigation can relatively easily uncover evidence of insufficient security measures. For examples of such events and services that are available to investigate potential cyberinsurance fraud see for instance Hoogstraaten et al. (gust) and the services provided by specialized security firms such as those covered in Mandiant (2013).

As discussed previously, the quantity $\sigma_i(x_i, n_i)$ represents the probability that target i incurs a loss from an attack. Hence, it also represents the probability that an insurer who insures target i for contract (q_i, ℓ_i) will pay out the amount $L_i - \ell_i$. This means that the profit of the insurer for target i would be

$$\Pi_i = q_i - \sigma(x_i, n_i)(L_i - \ell_i). \tag{18}$$

Since insurance markets are efficient, the insurer doe not make a profit on the equilibrium path and must therefore charge

$$q_i = \sigma_i(x_i, n_i)(L_i - \ell_i) \tag{19}$$

In the economic literature on insurance, the provision of insurance coverage at this price is commonly referred to as *actuarially fair insurance.* Consider now n_i as fixed exogenously, target i wishes to choose x_i and (q_i, ℓ_i) to maximize the expected utility $\mathbb{E}[U_i(x_i, q_i, n_i)]$. When the target has risk neutral induced preferences, the target is indifferent between buying insurance coverage or simply 'self insuring', even when this coverage is provided at an actuarially fair price. Hence a risk neutral target will always choose no insurance ($\ell_i = L_i$).

When targets are risk averse, their choices requires a more careful reasoning. At first we consider again a non-cooperative game among the players. The choices of attackers are unchanged w.r.t. their choices in absence of efficient insurance markets (see §3.3). In the absence of insurance, the strategy for a target at the equilibrium was simply the choice of defensive expenditure. When targets can also purchase actuarially fair insurance, target i's strategy involves two choices: (i) the level of deductibles l_i and (ii) the level of defensive expenditure, x_i. The premium will be determined by the deductible according (19) given our assumption about the efficiency of insurance markets.

Proposition 2: For a given number n_i of attackers per target, a risk averse target i which is offered insurance at an actuarially fair rate will always find it optimal to choose a level of coverage equal to the full loss ($\ell_i = 0$).

Proof. See Appendix 7.

Discussion of Proposition 2 The general solution of the Nash equilibrium requires the simultaneous solution of setting to zero of the partial derivative of the expected value in (17) with respect x_i and ℓ_i under the free entry constraint represented by (6). When $U_i(w)$ is weakly concave, so that target i is risk averse, it is convenient to solve for target i's optimal choice in two steps. At first, we calculate the optimal deductible $\ell_i(x_i)$, for each level of defensive expenditure x_i. Then, we calculate the optimal level of defensive expenditure x_i when ℓ_i is set to its optimal level, that is, when $\ell_i(x_i)$ is substituted for ℓ_i in (17).

Once target i has chosen to be fully insured ($\ell_i = 0$) Eq. (17) will reduce to a single state as the quote is payable in both periods and the coverage is complete as there is no deductible. This means that the i target will choose x_i to maximize the utility $U_i(W_i - x_i - \sigma_i(x_i, n_i)L_i)$. Since $U_i(w)$ is an increasing function, this corresponds to choosing x_i to maximize the expected net value of target i's assets, and namely $W_i - x_i - \sigma_i(x_i, n_i)L_i$. Unsurprisingly, a risk averse target who is able to offload the entire risk of a loss by the purchase of actuarially fair insurance chooses the same level of defensive expenditure as would be chosen by a risk neutral target.

For a given number of attackers, the optimal security expenditure of the risk-averse target who purchased actuarially fair insurance will therefore be identical to the security expenditure of the risk neutral target $x^\star$ and will be determined by Eq. (5). Yet, the target that purchases actuarially fair insurance will spend *more* than the target that opts for self-protection: for a given number of attackers per target n, the former will spend $x_i^\star - \sigma_i(x_i^\star, n)$, whereas the latter will spend $x_i^\star$. This can be immediately seen in Figure. **??** where point F is to the left of point E. This result is to be expected because risk averse target are indeed ... risk averse: they will chose to buy an insurance if they had the possibility. Thus, they end up spending the amount of money that a risk neutral target would spend ($x_i^\star$) and the additional cash equivalent to eliminate the risk ($\sigma_i(x_i^\star)L_i$).

The relation between $x_i^\star$ and the security expenditure of the risk neutral target x_i^* depends on the actual shapes of U_i and σ_i but it is possible to establish some general result for risk averse targets.

At first, we provide a general condition which determines when the availability of actuarially fair insurance is detrimental to the overall social welfare, i.e. a reduced security investments by all targets. Our condition is based on relation between the marginal rate of attack's success for the security investment for the risk neutral case x_i^{**} as determined by Eq. (**??**) and the normalized expected marginal utility of the risk averse target for the same level of investment.

Proposition 3: For a given number of attackers per target $n > 0$, let $x_i^\star$ be the optimal security investment for the risk averse target from Eq. (5) with actuarially fair insurance quote $q_i^\star = \sigma_i(x_i^\star, n_i)L$. The optimal security investment x_i^* of the risk averse target, in the absence of any available insurance contracts, will be lower than $x_i^\star$ if and only if the the expected marginal utility is smaller than the marginal utility of the local risk-neutral loss at $W_i - x_i^\star$. That is

$$x_i^\star < x_i^* \quad \text{iff} \quad \mathbb{E}[U_i'(x_i^\star)] < U'(W_i - x_i^\star - \mathcal{L}(W_i - x_i^\star, L_i))$$

Proof. See Appendix.

4.3 Proof of Proposition 3

Proof. First, we consider the case in which the targets have fair insurance available. We use Proposition 2 to determine that targets will chose full insurance and their utility function is therefore identical to $U_i(W_i - x_i - \sigma_i L_i)$. For any given number of attackers, the maximum value of the utility will be attained by setting the usual first order condition. The derivative of the utility function, in the presence of full insurance is the following

$$\tfrac{\partial U_i(W_i - x_i - \sigma_i L_i)}{\partial x_i} = U_i'(W_i - x_i - \sigma_i L_i)(-1 - \tfrac{\partial \sigma_i}{\partial x_i} L_i)$$

Since U_i is concave, the first factor is positive for all values of wealth, i.e. $U_i' > 0$. The first order condition can only be attained by setting the second factor to zero; this yields Eq. (5). So we denote with $x_i^\star$ be the value of the security investment for the insured target, which is equal to the optimal expenditure of the risk-neutral target.

For the no-insurance case the first order condition is given by

$$\begin{aligned}
& \tfrac{\partial \mathbb{E}[U_i(x_i)]}{\partial x_i} \\
= \quad & \tfrac{\partial(\sigma_i U_i(W_i - x_i - L_i) + (1-\sigma_i)U_i(W_i - x_i))}{\partial x_i} \\
= \quad & \tfrac{\partial \sigma_i}{\partial x_i} U_i(W_i - x_i - L_i) + \sigma_i U_i'(W_i - x_i - L_i)(-1) + \\
& -\tfrac{\partial \sigma_i}{\partial x_i} U_i(W_i - x_i) + (1 - \sigma_i) U_i'(W_i - x_i)(-1) \\
= \quad & -\sigma_i U_i'(W_i - x_i - L_i) - (1 - \sigma_i) U_i'(W_i - x_i) + \\
& -\tfrac{\partial \sigma_i}{\partial x_i}(U_i(W_i - x_i) - U_i(W_i - x_i - L_i)) \\
= \quad & -\mathbb{E}[U_i'(x_i)] - \tfrac{\partial \sigma_i}{\partial x_i} L_i \tfrac{U_i(W_i - x_i) - U_i(W_i - x_i - L_i)}{L_i} \\
= \quad & -\mathbb{E}[U_i'(x_i)] - \tfrac{\partial \sigma_i}{\partial x_i} L_i U_i'(W_i - x_i - \mathcal{L}(W_i - x_i, L_i))
\end{aligned}$$

Now, replace the value of $x_i^\star$ in $\frac{\partial \mathbb{E}[U_i]}{\partial x_i}$ and observe that at $x_i^\star$ it is $L_i \partial \sigma_i / \partial x_i = -1$. So we obtain the following value for the first order condition.

$$\left.\tfrac{\partial \mathbb{E}[U_i(x_i)]}{\partial x_i}\right|_{x_i = x_i^\star} = -\mathbb{E}[U_i'(x_i^\star)] + U_i'(W_i - x_i^\star - \mathcal{L}(W - x_i^\star, L_i))$$

Notice that U_i' and $\mathbb{E}[U_i']$ are both positive for all values of wealth. If $U_i'(W_i - x_i^\star - \mathcal{L}(W - x_i^\star, L_i)) < \mathbb{E}[U_i'(x_i^\star)]$ the overall value of $\partial \mathbb{E}[U_i(x_i)]/\partial x_i$ at $x_i^\star$ is negative.

Since the value of $\partial \mathbb{E}[U_i(x_i)]/\partial x_i$ at x_i^* is zero by definition of optimal expenditure for unregulated targets and it starts from a positive value at $x_i = 0$, the value of $x^\star$ must be reached after passing x^* and therefore $x_i^\star > x^*$.

Discussion of Proposition 3 This general condition is an *iff* condition so it does not *not* imply that the security expenditure (let alone the overall expenditure) of the actuarially fair insured risk averse target will always be lower than the security expenditures of the unregulated targets. The main reason for this behavior is that the interplay between σ_i and U_i can change the relative position of the optimal security expenditure for un-insured vs insured targets. Indeed, Lemma 2 states that for the uninsured targets the optimal level of expenditure x_i^* bounds the marginal expected loss as follows:

$$-\frac{1}{\lambda_i} \leq L_i \left.\frac{\partial \sigma_i}{\partial x_i}\right|_{x_i = x_i^*} \leq -\lambda_i$$

At the same time we have that

$$-\frac{1}{\lambda_i} \leq -1 = L_i \frac{\partial \sigma_i}{\partial x_i}\bigg|_{x_i = x_i^\star} \leq -\lambda_i$$

Therefore the marginal chances of a successful attacks for the optimal investment $x^\star$ of the insured target is essentially in the same narrow interval as the marginal chances for the optimal investment x_i^* for the un-insured target. The precise value of the cross-over point requires to specify a functional form for σ_i and U_i, and would therefore be less general. We explore such alternatives in Section 6.

The assumption that the outcome is a Nash equilibrium implies that the targets' forecasts of the number of attackers per target (i.e. the threats that they face). Similarly, it is assumed that potential attackers correctly infer the levels of defensive expenditure and, hence, the level of vulnerability of the population of targets.

The assumption that the choices of attackers and targets are made simultaneously implies that an individual target neglects the effect that a change in the target's level of defensive expenditure has on the incentives of potential attackers to mount attacks. Similarly, each potential attacker is assumed to neglect the effect that the attacker's decision might have on the overall level of threat perceived by the targets and, hence, on the targets' levels of defensive expenditure. These assumptions appear to be plausible approximations when the number of potential attackers and the number of targets is large. For in this case, a change in the choice of defensive expenditure by a single target is not likely to affect the overall expected reward from attacks by very much. Similarly, a change in the participation decision of a single potential attacker is not likely to have a significant effect on the number of attackers per target.

4.4 Policy Maker with Efficient Insurance Markets

To our efficient insurance markets case, we now introduce a benevolent social planner with the same utilitarian objectives as described in Section 3.4. The policy maker, the targets, the attackers and the insurers satisfy the same assumptions and objective functions of the previous sections. We again model the situation as a two stage game with a subgame-perfect equilibrium.

Under hypothesis of efficient insurance markets, the profits of the insurers are zero and we can ignore this component altogether in the utility function of the policy maker. Therefore, the objective function of the policy maker is determined by (15) where the utility of the targets is determined by (17) from Section 4.1 in place of (3) from Section 3.3.

The policy maker still chooses the level of defensive expenditure for each target in stage 1. In stage 2, each potential attacker also chooses whether or not to participate in attacks. The choices available to the targets are instead different from the no insurance case because each target i can decides the insurance contract, (q_i, ℓ_i). Each target chooses a level of coverage to maximize the expected utility given in (17) where $x_i^\dagger$ is determined exogenously by the policy maker. Under the assumption of insurance market efficiency the premium q_i will be determined by (19).

A strategy for the policy maker is simply the policy maker's choice of defensive expenditures for the targets. Strategies for stage 2 players are more complex. A strategy for each potential attacker is a specification of whether or not to participate in attacks for each possible set of choices by the policy maker in the first stage. A strategy for a target is a specification of the chosen level of coverage for each possible choice by the policy maker in the first stage.

In a subgame-perfect equilibrium the chosen ℓ_i, for each target i must be optimal for each set of defensive expenditures that could be chosen by the policy maker when the number of attackers per target is also given by the equilibrium level n_i^*. If a target is risk neutral, $\ell_i = L_i$ is, of course, always optimal. For the case where a target is strictly risk averse, Proposition 2) asserts that $\ell_i = 0$ is optimal for all levels of defensive expenditure x_i and all levels of n_i. In both cases, each target i always receives the expected utility value of its assets, $U(W_i - x_i - \sigma_i L_i)$.

Since the policy maker anticipates this outcome in stage one of the game, the policy maker's expected utility from (15) can be rewritten in the following simpler form.

$$\mathbb{E}[U_P] \;=\; \sum_{i=1}^{N} \nu_i U_i(W_i - x_i - \sigma_i(x_i, n^*(x_1, \ldots, x_N) L_i) \tag{20}$$

The function $n^*(x_1, \ldots, x_N)$ has also been substituted for each n_i in the expression for $\mathbb{E}[U_i]$ because the policy maker forecasts the response of potential attackers to different levels of defensive expenditure in the second stage of the game along Lemma 1.

The equilibrium levels of defensive expenditure determined in the first stage of the game are assumed to satisfy the usual first order conditions for optimality: $\partial \mathbb{E}[U_P]/\partial x_i = 0$ for all i. This condition can be decomposed into two components that clarify why the investment mandated by the policy maker is larger than the invested chosen by unregulated targets, even in presence of an efficient cyberinsurance market.

$$\frac{1}{\nu_i}\frac{\partial \mathbb{E}[U_P]}{\partial x_i} = \left.\frac{\partial \mathbb{E}[U_i]}{\partial x_i}\right|_{n=n^*} + \frac{1}{\nu_i}\frac{\partial n^*}{\partial x_i}\left.\frac{\partial \mathbb{E}[U_P]}{\partial n}\right|_{n=n^*} \tag{21}$$

Once again each target ignores the beneficial effect the target's expenditure has in reducing the number of attackers not only on itself but on the other targets as well. This phenomenon is captured by the second term of the decomposition which is the pro-quota variation to the social expectation of a reduction in the number of attackers ($\partial \mathbb{E}[U_P]/\partial n$) due to the reduction of this very number of attackers thanks to the increase in expenditure by the i-th target $\partial n^*/\partial x_i$.

Theorem 2: *In the presence of efficient markets for cyberinsurance for targets with a weakly concave utility function, the security investment $x_i^\ddagger$ mandated by a benevolent policy maker to each individual target i is larger than the optimal security investment x_i^* that the same target i would have chosen in an unregulated environment.*

Proof. See Appendix 7.

Discussion of Theorem 2. That the combination of efficient insurance market and benevolent social planner will reduce aggregate risk (by increasing overall expenditure) is not particularly surprising. The insurance market does not extract any surpluses from the targets and as such the risk averse targets will all choose complete coverage that is priced at an actuarially fair rate. The fully informed benevolent social planner then mandates security across all targets to eliminate the externalities generated by the attacker behavior.

5. MANDATED SELF-PROTECTION FROM A MONOPOLY CYBER INSURER

We now consider A.10 and A.11 where a single insurer who can set a required level of defensive expenditure as part of the insurance contract. This scenario ought to provide the best chance for the incentives of a profit-maximizing insurer to align with those of a "benevolent" policy maker. In contrast, the choice of defensive expenditure by the insurer will not generally be socially optimal.

For example, in credit card security the consortium of credit card companies play the mutual role of social policy makers and insurer provider. The PCI DSS standard (Williams and Chuvakin, 2012, Chap.3) precisely specifies the risk level of vulnerabilities that must be patched. Merchants whose software does not comply with the specification will be subject to hefty fines and will be responsible for any disowned purchase, and eventually lose possibility to accept credit cards from customers. At the same time, merchants that deploy the extra security measures mandated by the credit card consortium will benefits from a liability waiver in case of frauds and disowned purchases which will be absorbed by the credit card companies[12].

The outcome of the interaction between the insurer and the targets is modeled as a subgame-perfect equilibrium of a two-stage game. In the first stage of the game, the insurer makes its offer to each target. The offer consists now of a triple $(q_i, \ell_i, x_i^\sharp)$ where the first two arguments are respectively the premium and the deductible as in the previous section and $x_i^\sharp$ is the required level of defensive expenditure which an insured target must incur. The total profit obtained by an insurer would be the difference between the premium paid by the targets and expected losses that he must cover (minus the deductibles).

$$\Pi = \sum_{i=1}^{N} q_i + \sigma_i(x_i^\sharp, n^*(x_1^\sharp, \ldots x_N^\sharp)(\ell_i - L_i) \tag{22}$$

where $\ell_i \leq L_i$. Later in we will shorten the term $n^*(x_1^\sharp, \ldots x_N^\sharp)$ to $n^*(x_i^\sharp)$.

In the second stage of the game, targets and potential attackers make simultaneous choices. Each target must choose whether or not to accept the insurer's offer. If a target accepts, then no further choice is required. If a target rejects the offer, then the target must also choose the level of defensive expenditure x_i which it will incur. As in previous sections, each potential attacker must choose whether or not to participate in attacks on the targets.

The monopolist wishes to offer an insurance contract which all targets will be willing to purchase.

[12] http://usa.visa.com/merchants/grow-your-business/payment-technologies/verified-by-visa.jsp.

For the targets to be willing to accept the insurer's offer, they must be indifferent between accepting or not accepting and at the same time the insurer will try to extract the maximum possible rent from the target. Therefore, the following incentive compatibility constraint must hold

$$\begin{aligned}&\mathbb{E}[U_i(q_i,\ell_i,x_i^\sharp,n^*(x_i^\sharp))]\\ &\qquad\geq \mathbb{E}[U_i(0,L_i,x_i^*(n^*(x_i^\sharp)),n^*(x_i^\sharp))].\end{aligned} \tag{23}$$

Since the insurer holds a monopoly it will extract all possible surplus from each target and therefore we only consider the boundary form of the constraint.

The left-hand side of (23), as defined in (17) denotes the expected utility which target i obtains by accepting the insurance contract. In an equilibrium, where all targets are assumed to purchase insurance and choose the level of expenditure specified by the insurance contract, the appropriate forecast of the number of attackers is $n^*(x_i^\sharp)$, the equilibrium number of attackers per target. The right-hand side of (23) indicates the expected utility which a target obtains by rejecting the insurer's offer $q_i = 0$ and $\ell_i = L_i$ and making an optimal choice of defensive expenditure, ceteris paribus.

It is convenient to solve the insurer's problem in two steps. In Step 1, the insurer chooses $q_i(x_i)$ and $\ell_i(x_i)$ to produce the profit that satisfies the incentive compatibility constraint for each level of defensive expenditure, x_i. In Step 2, the insurer chooses the level of defensive expenditure that maximizes $\Pi(x_i)$.

When all x_i are held fixed, the right-hand side of (23) is a constant as is the quantity $\sigma_i(x_i, n^*(x_1, \dots x_i, \dots x_n)$ in the profit of the insurer from (22). Then the left side of (23) is always smaller then the utility of target when he chose the level of deductible ℓ_i.

$$U_i(W_i - x_i - q_i) \geq \mathbb{E}[U_i(q_i,\ell_i,x_i^\sharp,n^*(x_i^\sharp))] \tag{24}$$

As such, the insurer is going to always offer the insuree a full insurance contract $(q_i, 0)$ as this would maximize the value of the offer for the target. This is same that was happening when the target was receiving an offer for actuarially fair insurance as in Proposition.2. At the same time, since the insurer has a monopolist advantage the compatibility bound will hold as an equality. This provides the following implicit expression for q_i and x_i.

$$\begin{aligned}U_i(W_i - x_i - q_i) \quad = \quad & (1-\sigma_i)(x_i, n^*(x_i))U_i(W_i - x_i)\\ & +\sigma_i(x_i, n^*(x_i))U_i(W_i - x_i - L_i).\end{aligned} \tag{25}$$

Since U_i is invertible, by construction, the implicit function above is well defined for q_i.

In order to understand what happens on the equilibrium path, it is useful to consider how the incentive compatibility equation makes sure that a change in the mandated security expenditure x_i of target i affects *both* the quote q_i of the target i itself *and* the quote q_j of the the other targets j. The marginal adjustment in the quotes q_j and q_i is given by the following equations.

$$\begin{aligned}\frac{\partial q_j}{\partial x_i} &= -1 - \frac{1}{U_i'(W_i - x_i - q_i)}\frac{\mathbb{E}[U_i(x_i)]}{\partial x_i}\\ &+\frac{\partial n^*}{\partial x_i}\cdot\frac{\partial \sigma_i}{\partial n}L_i\cdot\frac{U_i'(W_i - x_i - \mathcal{L}(W_i - x_i, L_i))}{U_i'(W_i - x_i - q_i)}\\ \frac{\partial q_j}{\partial x_i} &= \frac{\partial n^*}{\partial x_i}\cdot\frac{\partial \sigma_j}{\partial n}L_i\cdot\frac{U_j'(W_j - x_j - \mathcal{L}(W_i - x_i, L_i))}{U_j'(W_j - x_j - q_j)}\end{aligned} \tag{26}$$

The first two components of the equation for $\partial q_i/\partial x_i$ are to beexpected. The first term is a constant negative marginal rate showing that, ceteris paribus, the quote must decreases linearly with the mandated security expenditures x_i since the latter directly competes with the quote q_i for a direct share of the overall wealth $W_i - x_i - q_i$ of the target. The second term has the opposite sign of the marginal expected utility of the uninsured target (calculated as in Equation 10) and takes into account the ability of the insurer to monetize the risk aversion of the target: the quote increases as the expected utility of the target decreases. This growth is controlled by a factor, the marginal utility of the target when he consider both the expenditure and the quote. Depending of the relative value, the partial derivative $\partial q_i/\partial x_i$ may be positive or negative and therefore there is a an optimal value for the insurer.

The first two terms of Equation (26) would not be sufficient to guarantee that the equilibrium condition would be reachable. If the monopolist sets the initial insurance at too expensive a value then targets will not diffuse to buy this expensive insurance. So the insurance company would need to specify initially some contract that ensures that targets are shifted to an equilibrium path and progressively rise the expenditure whereby $x^\sharp$ is the outcome and that no individual target can do better by rejecting this offer. Only after all targets are at $x^\sharp$ then the equilibrium will be maintained because any target is worse off by rejecting insurance. An alternative would be for the insurer to collude with a policy maker so that targets are forced to take on an initial insurance.

The need to lure targets into insurance is captured by the third term of Eq. (26) and the condition on the other partial derivative Eq. (26). While Equation (26) can change sign, the second equation is *always* negative. It is the product of three terms: the first term is the changes in number of attackers n^* at equilibrium as the security expenditure of x_i changes, the marginal changes in expected loss when the number of attackers increases (positive), and finally the ratio between the marginal utility at the locally risk neutral loss for the uninsured target and the marginal utility of the remaining wealth of the insured target (positive as well). By Lemma 1 we have that $\partial n^*/\partial x_i$ is proportional to $\partial\sigma_i/\partial x_i$ and the latter is negative by assumption. By increasing the security expenditure of a target, the insurer is forced to decrease the quote of the other targets.

This phenomenon will be absent without the strategic features of the target-attacker game: in absence of the strategic interaction of *both* targets and attackers ensembles, the insurer would simply mandate an exorbitant security expenditure to each target to decrease the expected loss and pocket the entire quote. However, if the insurer starts mandating too much expenditure to the insured targets, the overall number of attacker will diminish. Therefore, the other targets will have less incentives to join or keep the insurance scheme as the overall risk will diminish. As a result, the insurer has to lower the quote *and* to require a lower mandatory security expenditure to make insurance more appealing.

The insurer optimizes its profit function by taking the usual first order condition on the profit (22) under the implicit expression for $q(X_i)$ defined by (25) and the attackers entry condition.

$$\begin{aligned} \frac{\partial \Pi}{\partial x_i} &= 0 && (27)\\ U_i(W_i - x_i - q_i) &= \mathbb{E}[U_i]. && (28) \end{aligned}$$

The first order condition can then be decomposed as follows:

$$\begin{aligned} &\frac{\partial \Pi}{\partial x_i} = \\ &-1 + \frac{\mathbb{E}[U_i'(x_i)]}{U_i'(W_i - x_i - q_i)} + \\ &+\frac{\partial\sigma_i}{\partial x_i} L_i \big(\frac{U_i'(W_i - x_i - \mathcal{L}(W_i - x_i, L_i))}{U_i'(W_i - x_i - q_i)} - 1\big) \\ &+\frac{\partial n^*}{\partial x_i} \sum_{j=1}^{N} \frac{\partial \sigma_j}{\partial n} L_j \; \big(\frac{U_j'(W_j - x_j - \mathcal{L}(W_j - x_j, L_j))}{U_j'(W_j - x_j - q_j)} - 1\big) \end{aligned} \qquad (29)$$

Theorem 3: *For a given number of attacks per target $n > 0$, let the optimal offer of quote, deductible and mandatory security investment $(q_i^\sharp, \ell_i^\sharp, x_i^\sharp)$ of the monopolist insurer for a downside risk aver target determined by solving the equation below for $x_i^\sharp$*

$$\begin{aligned} \mathbb{E}[U_i'(x_i^\sharp)] &= U_i'(W_i - x_i^\sharp - \mathcal{L}(W_i - x_i^\sharp, L_i)) && (30)\\ q_i^\sharp &= \mathcal{L}(W_i - x_i^\sharp, L_i) && (31)\\ \ell_i^\sharp &= 0 && (32) \end{aligned}$$

The optimal security expenditure x_i^* chosen by the unregulated targets who cannot purchase insurance is larger than the above security expenditure $x_i^\sharp$ mandated by the monopolist insurer if the marginal change in the expected loss at $x^\sharp$ is smaller then negative unity i.e.

$$\frac{\partial \sigma_i(x_i^\sharp)}{\partial x_i} L_i < -1 \textit{ implies } x^\sharp < x^*$$

Proof. See Appendix 5.1.

5.1 Proof of Theorem 3

Proof. At first we prove the that the expectation of the target is always maximized by setting deductibles $\ell_i = 0$.

$$\begin{aligned} &\mathbb{E}[U_i(q_i, \ell_i, x_i^\sharp, n^*(x_i^\sharp))] = \\ &(1-\sigma_i)U_i(W_i - x_i - q_i) + \sigma_i U_i(W_i - x_i - q_i - \sigma_i \ell_i) \\ &\leq (1-\sigma_i)U_i(W_i - x_i - q_i) + \sigma_i U_i(W_i - x_i - q_i) \\ &= U_i(W_i - x_i - q_i) \end{aligned}$$

Then we derive (26) and (26) from the incentive compatibility equation (25). For the first equation we use the following derivation:

$$\frac{\partial U_i(W_i-x_i-q_i)}{\partial x_i} = \frac{\partial \mathbb{E}[U_i]}{\partial x_i}$$
$$U_i'(W_i - x_i - q_i)\frac{\partial(W_i-x_i-q_i)}{\partial x_i}$$
$$= \frac{\partial((1-\sigma_i(x_i,n^*(x_i)))U_i(W_i-x_i)+\sigma_i(x_i,n^*(x_i))U_i(W_i-x_i-L_i))}{\partial x_i}$$
$$U_i'(W_i - x_i - q_i)\left(-1 - \frac{\partial q_i}{\partial x_i}\right)$$
$$= \frac{\partial(1-\sigma_i(x_i,n^*(x_i))}{\partial x_i}U_i(W_i - x_j) + (1-\sigma_i)U_i'(W_i - x_i)(-1)$$
$$+\frac{\partial\sigma_i(x_i,n^*(x_i))}{\partial x_i}U_i(W_i - x_i - L_i)) + \sigma_i U_i'(W_i - x_i - L_i)(-1)$$
$$= -\frac{\partial\sigma_i(x_i,n^*(x_i))}{\partial x_i}\Delta U_i - \mathbb{E}[U_i'(x_i)]$$
$$= -(\frac{\partial\sigma_i}{\partial x_i} + \frac{\partial\sigma_i}{\partial n}\frac{\partial n^*}{\partial x_i})\Delta U_i - \mathbb{E}[U_i'(x_i)]$$
$$\frac{\partial q_j}{\partial x_i} = -1 + \frac{\mathbb{E}[U_i'(x_i)]}{U_i'(W_i-x_i-q_i)} +$$
$$+(\frac{\partial\sigma_i}{\partial x_i} + \frac{\partial\sigma_i}{\partial n}\frac{\partial n^*}{\partial x_i})\frac{\Delta U_i}{U_i'(W_i-x_i-q_i)}$$
$$= -1 + \frac{\mathbb{E}[U_i'(x_i)]}{U_i'(W_i-x_i-q_i)} + \frac{\partial\sigma_i}{\partial x_i}\frac{\Delta U_i}{U_i'(W_i-x_i-q_i)}$$
$$+\frac{\partial n^*}{\partial x_i}\frac{\partial\sigma_i}{\partial n}L_i\frac{\Delta U_i}{U_i'(W_i-x_i-q_i)}$$
$$= -1 - \frac{1}{U_i'(W_i-x_i-q_i)}\frac{\partial\mathbb{E}[U_i(x_i)]}{\partial x_i}$$
$$+\frac{\partial n^*}{\partial x_i}\frac{\partial\sigma_i}{\partial n}L_i\frac{\Delta U_i}{U_i'(W_i-x_i-q_i)}$$
$$= -1 - \frac{1}{U_i'(W_i-x_i-q_i)}\frac{\partial\mathbb{E}[U_i(x_i)]}{\partial x_i}$$
$$+\frac{\partial n^*}{\partial x_i}\frac{\partial\sigma_i}{\partial n}L_i\frac{U_i'(W_i-x_i-\mathcal{L}(W_i-x_i,L_i))}{U_i'(W_i-x_i-q_i)}$$

For the second equation we use the following derivation

$$\frac{\partial U_j(W_j-x_j-q_j)}{\partial x_i} = \frac{\partial\mathbb{E}[U_j]}{\partial x_i}$$
$$U_j'(W_j - x_j - q_j)\frac{\partial(W_j-x_j-q_j)}{\partial x_i}$$
$$= \frac{\partial((1-\sigma_j(x_j,n^*(x_i)))U_j(W_j-x_j)+\sigma_j(x_j,n^*(x_i))U_j(W_j-x_j-L_j))}{\partial x_i}$$
$$U_j'(W_j - x_j - q_j)\left(-\frac{\partial q_j}{\partial x_i}\right)$$
$$= \frac{\partial(1-\sigma_j(x_j,n^*(x_i))}{\partial x_i}U_i(W_i - x_j)$$
$$+\frac{\partial\sigma_j(x_j,n^*(x_i))}{\partial x_i}U_j(W_j - x_j - L_j))$$
$$= -\frac{\partial\sigma_j(x_j,n^*(x_i))}{\partial x_i}\Delta U_j$$
$$= -\frac{\partial\sigma_j}{\partial n}\frac{\partial n^*}{\partial x_i}\Delta U_j$$
$$= -\frac{\partial n^*}{\partial x_i}\frac{\partial\sigma_j}{\partial n}L_j\, U_j'(W_j - x_j - \mathcal{L}(W_j - x_j, L_j))$$
$$\frac{\partial q_j}{\partial x_i} =$$
$$\frac{\partial n^*}{\partial x_i}\frac{\partial\sigma_j}{\partial n}L_i\ \frac{U_j'(W_j-x_j-\mathcal{L}(W_j-x_j,L_j))}{U_j'(W_j-x_j-q_j)}$$

Now the insurer will optimize its profit function by taking the usual first order condition.

$$\frac{\partial\Pi}{\partial x_i} = \frac{\partial\sum_{j=1}^N q_j - \sigma(x_j,n^*(x_i))L_j}{\partial x_i}$$
$$= \frac{\partial(q_i-\sigma(x_i,n^*(x_i))L_i)}{\partial x_i} + \sum_{j\neq i}\frac{\partial q_j-\sigma(x_j,n^*(x_i))L_j}{\partial x_i}$$
$$= \frac{\partial q_i}{\partial x_i} - \frac{\partial\sigma(x_i,n^*(x_i))}{\partial x_i}L_i + \sum_{j\neq i}(\frac{\partial q_j}{\partial x_i} - \frac{\partial\sigma(x_j,n^*(x_i))L_j}{\partial x_i})$$
$$= -1 - \frac{1}{U_i'(W_i-x_i-q_i)}\frac{\partial\mathbb{E}[U_i(x_i)]}{\partial x_i}$$
$$+\frac{\partial n^*}{\partial x_i}\frac{\partial\sigma_i}{\partial n}L_i\frac{U_i'(W_i-x_i-\mathcal{L}(W_i-x_i,L_i))}{U_i'(W_i-x_i-q_i)}$$
$$-(\frac{\partial\sigma_i}{\partial x_i} + \frac{\partial n^*}{\partial x_i}\frac{\partial\sigma_i}{\partial n})L_i+$$
$$+\sum_{j\neq i}\frac{\partial n^*}{\partial x_i}\frac{\partial\sigma_j}{\partial n}L_j\ \frac{U_j'(W_j-x_j-\mathcal{L}(W_j-x_j,L_j))}{U_j'(W_j-x_j-q_j)}+$$
$$-\sum_{j\neq i}\frac{\partial n^*}{\partial x_i}\frac{\partial\sigma_j}{\partial n}L_j$$
$$= -1 + \frac{\mathbb{E}[U_i'(x_i)]}{U_i'(W_i-x_i-q_i)} + \frac{\partial\sigma_i}{\partial x_i}\frac{\Delta U_i}{U_i'(W_i-x_i-q_i)}$$
$$+\frac{\partial n^*}{\partial x_i}\frac{\partial\sigma_i}{\partial n}L_i\frac{U_i'(W_i-x_i-\mathcal{L}(W_i-x_i,L_i))}{U_i'(W_i-x_i-q_i)}$$
$$-(\frac{\partial\sigma_i}{\partial x_i} + \frac{\partial n^*}{\partial x_i}\frac{\partial\sigma_i}{\partial n})L_i+$$
$$+\sum_{j\neq i}\frac{\partial n^*}{\partial x_i}\frac{\partial\sigma_j}{\partial n}L_j\ \frac{U_j'(W_j-x_j-\mathcal{L}(W_j-x_j,L_j))}{U_j'(W_j-x_j-q_j)}+$$
$$-\sum_{j\neq i}\frac{\partial n^*}{\partial x_i}\frac{\partial\sigma_j}{\partial n}L_j$$
$$= -1 + \frac{\mathbb{E}[U_i'(x_i)]}{U_i'(W_i-x_i-q_i)} + \frac{\partial\sigma_i}{\partial x_i}L_i\frac{U_i'(W_i-x_i-\mathcal{L}(W_i-x_i,L_i))}{U_i'(W_i-x_i-q_i)}$$

$$+\frac{\partial n^*}{\partial x_i}\frac{\partial\sigma_i}{\partial n}L_i\frac{U_i'(W_i-x_i-\mathcal{L}(W_i-x_i,L_i))}{U_i'(W_i-x_i-q_i)}$$
$$-(\frac{\partial\sigma_i}{\partial x_i} + \frac{\partial n^*}{\partial x_i}\frac{\partial\sigma_i}{\partial n})L_i+$$
$$+\sum_{j\neq i}\frac{\partial n^*}{\partial x_i}\frac{\partial\sigma_j}{\partial n}L_j\ \frac{U_j'(W_j-x_j-\mathcal{L}(W_j-x_j,L_j))}{U_j'(W_j-x_j-q_j)}+$$
$$-\sum_{j\neq i}\frac{\partial n^*}{\partial x_i}\frac{\partial\sigma_j}{\partial n}L_j$$

We can now start to group terms appropriately:

$$\frac{\partial\Pi}{\partial x_i} = -1 + \frac{\mathbb{E}[U_i'(x_i)]}{U_i'(W_i-x_i-q_i)}+$$
$$+\frac{\partial\sigma_i}{\partial x_i}L_i(\frac{U_i'(W_i-x_i-\mathcal{L}(W_i-x_i,L_i))}{U_i'(W_i-x_i-q_i)} - 1)$$

$$+\frac{\partial n^*}{\partial x_i}\sum_{j=1}^N\frac{\partial\sigma_j}{\partial n}L_j\ (\frac{U_j'(W_j-x_j-\mathcal{L}(W_j-x_j,L_j))}{U_j'(W_j-x_j-q_j)} - 1)$$

Now we can consider a special value for $q_i = \mathcal{L}(W_i - x_i, L_i)$ for all targets i. By replacing the value in the first order condition for Π and putting the value to zero we obtain the following constraint that must additionally be satisfied:

$$\begin{aligned}\mathbb{E}[U_i'(x_i)] &= U_i'(W_i - x_i - q_i)\\ &= U_i'(W_i - x_i - \mathcal{L}(W_j - x_j, L_j))\end{aligned}$$

To determine whether such solution is at all possible we observe that for downside risk averse targets, Jensen's inequality is reversed for the convex function U_i' and we have $U_i'(W_i - x_i - \sigma_i L_i) \leq \mathbb{E}[U_i'(x_i)]$. Therefore it is possible to find a value of q_i such that $U_i'(W_i - x_i - \sigma_i L_i) \leq U_i'(W_i - x_i - q_i) \leq$

$\mathbb{E}[U_i'(x_i)]$. This yields the following derivation

$$\begin{aligned}&U_i'(W_i - x_i - \mathcal{L}(W_j - x_j, L_j)) = \mathbb{E}[U_i'(x_i)]\\&\geq U_i'(W_i - x_i - \sigma_i L_j)\\&W_i - x_i - \mathcal{L}(W_j - x_j, L_j) \leq W_i - x_i - \sigma_i L_j\\&\mathcal{L}(W_j - x_j, L_j) \geq \sigma_i L_j\end{aligned}$$

Therefore *if* the equation above admits a solution for a particular value of x_i this is a feasible solution for the insurer since he will make a positive profit. Let $x_i^\sharp$ be such value. Replace now the value of $x^\sharp$ in the Equation 10.

$$\begin{aligned}&\tfrac{\partial \mathbb{E}[U_i]}{\partial x_i}\\&= -U_i'(W_i - x_i^\sharp - \mathcal{L}(W_i - x_i^\sharp, L_i)) -\\&\tfrac{\partial \sigma_i}{\partial x_i} L_i U_i'(W_i - x_i^\sharp - \mathcal{L}(W_i - x_i^\sharp, L_i))\\&= (-1 - \tfrac{\partial \sigma_i}{\partial x_i} L_i) U_i'(W_i - x_i^\sharp - \mathcal{L}(W_i - x_i, L_i))\end{aligned}$$

If $(-1 - \frac{\partial \sigma_i}{\partial x_i}(x_i^\sharp) L_i) > 0$ then the marginal changes in the expected utility is positive at $x_i^\sharp$ and started from a positive value at $x_i = 0$. Such value must eventually increase to reach zero and thus $x^* > x_i^\sharp$. Therefore $\frac{\partial \sigma_i}{\partial x_i}(x_i^\sharp) L_i) < -1$ implies $x^\sharp < x^*$.

Discussion of 3

6. QUANTITATIVE EXAMPLE

It is useful to visualize the examples outlined in §(4) and §(5) using a worked case. The major choices of HARA type utility function usually coalesce on constant absolute risk aversion (CARA) as opposed constant relative risk aversion (CRRA). Whilst CRRA preferences are usually deemed to provide a more complete picture of risk preferences; CARA functions often provide a useful testbed for a model as solutions can usually be described in terms of simple analytic functions. A further motivation for using CARA preferences as a test case is that the baseline level of wealth W_i does not need to be specified. We should add that for each of the cases outlined below and equivalent case can be specified with a CRRA type function, solved numerically.

Exponential type functional forms for CARA preferences are very tractable and provide simple and intuitive forms useful in illustrating the major facets of the model.

Let $\sigma = \sigma_i$, $L = L_i$, $n = n_i$, $R = R_i$, hence $\rho = \rho_i = R/C$ and $x = x_i$. Recall, that for the simultaneous Nash equilibrium targets are assumed to be ex-ante identical and as such choose symmetric identical actions. We assume that the CARA utility we will use is an exponential utility function of the standard form:

$$U(w) = -1/\gamma e^{-\gamma w} \tag{33}$$

where γ is the coefficient of absolute risk aversion. For the probability of a successful attack as a function of security investment x and attacking intensity n we choose:

$$\sigma = e^{-\alpha x} n^\beta \tag{34}$$

where $\alpha > 0$ and $0 < \beta < 1$ are positive scalar parameters. Notice, that σ is bounded in n, in order to be interpreted as a probability.

The locally risk neutral loss function is independent on the level of wealth and is equal to

$$\mathcal{L}(w, L) \quad = \quad \frac{1}{\gamma} \log\left(\frac{e^{\gamma L} - 1}{\gamma L}\right) \tag{35}$$

As such we can only consider equilibrium cases for this functional form when $0 \leq n < \exp(\frac{\alpha x}{\beta})$. When σ exhibits exponential decay in x, risk neutral targets will exhibit 'log-optimal' behavior. This implies that the asymptotic behavior of the optimal investment function is $O(\log(L))$, in the absence of insurance and with fixed attacking intensity. Validation of this functional form has been the subject of extensive study in the security literature, see Ioannidis et al. (2013) and Lelarge (2012) for typical examples in non-insurance settings.

For analytical tractability we will only consider cases when $\alpha > \gamma$. Following the steps in Corollary 1, the equilibrium expenditure on investments for risk averse targets with utility function from (33) is as follows

$$\begin{aligned}x^* \quad = \quad & \frac{\beta}{\alpha}\log(\rho) + \frac{1-\beta}{\alpha}\log\left(\frac{\alpha}{\gamma} - 1\right)\\&+\frac{1-\beta}{\alpha}\log\left(e^{\gamma L} - 1\right)\end{aligned} \tag{36}$$

for comparison, when all targets are risk neutral, the symmetric equilibrium is given by

$$x^\ddagger = \frac{\beta}{\alpha}\log(\rho) + \frac{1-\beta}{\alpha}\log(\alpha L) \tag{37}$$

and the number of attackers in both cases is given by $n^\clubsuit = \rho \exp(\frac{1}{1-\beta}\alpha x^\clubsuit)$, where $\clubsuit\in\{*,\dagger\}$. Derivations for (36) and (37) may be found in Appendix **??**. Notice that both $x^\dagger \geq x^*$ and $x^* > x^\dagger$ are possible under this derivation. However, notice that by setting $\alpha > \gamma$, $x^\dagger$ is $O(\log(L))$ whereas x^* is $O(L)$, therefore for large L, the asymptotic behavior is that $x^* > x^\dagger$, as such for arbitrarily large L the presence of actuarially fair insurance reduces aggregate investment and hence increases n. When $\alpha \leq \gamma$ the model looses analytical tractability as the expression inside the logarithm in the second term of (36) less than or equal to zero for all values of L. In the $\alpha \leq \gamma$ case the solution needs to be numerically solved from the first derivative of expected utility.

Let us now consider the case when a benevolent utilitarian social planner internalizes the attacker reaction function and sets mandatory investment across all targets. The first order condition that the policy maker solves for each ex-ante identical target is

$$\begin{aligned} & \frac{1}{\gamma} e^{-\gamma W+\alpha(-x)+\gamma x} \\ & \times \left((\alpha-\gamma)\rho^{\frac{\beta}{1-\beta}} \left(e^{\gamma L}-1\right) e^{\frac{\alpha\beta \mathrm{x}}{\beta-1}} - \gamma e^{\alpha x} \right) = 0 \end{aligned} \tag{38}$$

subject to $x > 0$ and this is given by

$$\begin{aligned} x^{\dagger} &= \frac{\beta}{\alpha}\log(\rho) \\ &+ \frac{\beta-1}{\alpha}\log\left(\frac{(1-\beta)\gamma}{\left(e^{\gamma L}-1\right)\left(\alpha+(\beta-1)\gamma\right)} \right). \end{aligned} \tag{39}$$

Notice that the term within the logarithm is always positive, however it is receding towards zero at a rate proportional to $\exp(\gamma L)$, in L by inspection we can see that asymptotically the optimal investment function is $O(L)$.

When actuarially fair insurance is available, the optimal target investment under actuarially fair insurance is the solution in x to

$$\begin{aligned} & \frac{\partial \mathbb{E}[U_P]}{\partial x} = -e^{\alpha x} \\ & - \frac{\alpha}{\beta-1} L \left(\rho^{\frac{\beta}{1-\beta}} e^{\frac{\alpha\beta x}{\beta-1}} \right) e^{\gamma L} e^{-\alpha x} \left(\rho^{\frac{\beta}{1-\beta}} e^{\frac{\alpha\beta x}{\beta-1}} \right) \\ & -\gamma W + x(\gamma-\alpha) = 0. \end{aligned} \tag{40}$$

using simple algebraic manipulation we can see that the resulting expression is nearly the same as in that in (37) for $x^{\ddagger}$ with a minor adjustment

$$x^{\ddagger\ddagger} = \frac{\beta}{\alpha}\log(\rho) + \frac{\beta-1}{\alpha}\left(\log(1-\beta) - \log(\alpha L)\right) \tag{41}$$

We now turn to the case when the insurance provision is in the form of a monopoly. First, for a given level of defensive expenditure x and attacking intensity n the maximum quote q that can charged to a target, for full insurance $\ell = 0$, before the target rejects is given by

$$q^{\dagger} = \frac{1}{\gamma}\log\left(e^{\gamma L} n^{\beta} - n^{\beta} + e^{\alpha x}\right) - \frac{\alpha x}{\gamma}. \tag{42}$$

The monopolist insurers profit function from (22) for this case is given by $\Pi = N(q - \sigma L)$. As all targets are identical in this example the profit per target is representative of the overall profit. Therefore, using our expression for (42) and substituting the optimal attacker behavior for n, the monopolists profit is therefore

$$\begin{aligned} \Pi(x) &= L\rho^{\frac{\beta}{1-\beta}} \left(-e^{\frac{\alpha x}{\beta-1}} \right) \\ & - \frac{1}{\gamma}\log\left(\rho^{\frac{\beta}{1-\beta}} \left(e^{\gamma L}-1\right) e^{\frac{\alpha\beta x}{\beta-1}} + e^{\alpha x} \right) + \frac{\alpha x}{\gamma} \end{aligned} \tag{43}$$

solving $\partial\Pi(x)/\partial x = 0$ with respect to x, yields the monopolists optimal mandated security investment

$$\begin{aligned} x^{\sharp} &= \frac{\beta}{\alpha}\log(\rho) + \frac{(\beta-1)}{\alpha}\log\left(e^{\gamma L} - 1 - \gamma L\right) \\ & - \frac{(\beta-1)}{\alpha}\log\left(\gamma L\left(e^{\gamma L}-1\right)\right). \end{aligned} \tag{44}$$

As $L \rightarrow 0$, this is subject to the rationality constraint that you would not spend more on mitigating a loss than the loss itself, as such $L - x^{\sharp} - q^{\dagger}(x^{\sharp}) > 0$. Notice, that depending on the size of L, the second and third terms in (44) will switch signs, however, one or other will always be positive. We can also see that the rationality condition $W - q(x^{\sharp}) - x^{\sharp}$ will bind for $L \rightarrow 0$.

6.1 An Insurance Trap

The CARA example provides us with some interesting predictions regarding the introduction of cyberinsurance contracts. Let us consider the following scenario, for our group of ex-ante identical risk averse targets we introduce actuarially fair insurance that. Two scenarios can then occur. First, case one, if L is sufficiently close to the origin, then the presence of actuarially fair insurance results in targets wanting to fully insure, by setting their deductible to $\ell = 0$ and they will increase their expenditure relative to the equilibrium in the absence of insurance. Second, case two, if L is sufficient large, targets will still take insurance, but will reduce their overall expenditure on protection as the targets can substitute coverage for risk reduction relatively inexpensively.

Let us now assume that the insurance market is no longer actuarially fair, for instance there is significant consolidation activity in the sector leaving only one monopolist, and this insurance company now begins to mandate that their insurees (the covered targets) set specific minimum levels of investment. Alternatively, there may only have been one insurance provider and a regulatory mechanism designed to prevent monopolistic behaviour is relaxed.

In the mandated quote investment combination results in an expected utility lower than that achieved by the targets in the absence of insurance, under certain circumstances targets can gradually reject the insurance contract and the group of targets will return to the equilibrium in the absence of insurance as the targets move up the Expected utility gradient by rejecting the insurance contracts. However, this will not be the case for sufficiently large L as an individual target will need to reduce their expected utility by rejecting the insurance contracts and rely on other firms to also exit the market.

The reason for this trap is as follows. For case one, the degree of attacking intensity faced by an individual target who rejects the unfair insurance contract is lower than the Nash equilibrium. So the individual target can maximize their expected utility by choosing no insurance and solving for their optimal investment, as more targets reject the unfair insurance, the aggregate investment is reduced, but the path to the equilibrium without insurance is always increasing, in expected utility for a given target rejecting an unfair contract.

In contrast for case two, if the expected utility for the equilibrium for no insurance is higher than that attained by the monopolist insurance company mandating investments and setting quotes higher than are actuarially fair, the targets can only continuously move up to a higher expected utility if they all simultaneously rejected the insurance contract. This is because the aggregate investment in security will be substantially lower under the monopolist. Assuming that a single targets increased investment choice alone cannot materially reduce the attacker intensity, an individual target would need to increase security investment substantially to obtain a similar expected wealth than with the unfair insurance contract. As long as the quote for the insurance contract obeys the incentive compatibility constraint in (42) the firm will have to lower expected utility by rejecting the contract and hope that other firms reject the contract and their combined increase in the security investment results in the expected utility from self-protection being higher than the utility obtained from the monopolist insurer.

7. CONCLUSIONS

This paper has outlined a series of models relating to the impact of strategic behaviour by attackers on the standard axioms of risk in a utility maximising set-up. The paper presents a general model based on simple behavioural assumptions for both attacker, target, insurance company and policy maker. The game is solved with the attacker and target in a sub-game and the policy maker and target engaging in a second level game. We show that the inclusion of a insurer providing actuarially fair insurance increases aggregate welfare for targets, but reduces aggregate security investment. We then show that a monopolist insurer extracting the maximum rent from the collection of targets has no incentive to mandate higher levels of defensive expenditure, as this does not achieve the insurers maximised profit. The overarching conclusions are that public policy is needed as the attacker generates an inherent externality by the presence of their reward maximising effort, however, the delegation to an insurance company (in particular a monopolist) does not provide any reduction in aggregate risk.

ACKNOWLEDGEMENTS

This work has been partly funded by the European Union's 7th Framework Programme under grant agreement no 285223 - SECONOMICS (www.seconomics.org).

REFERENCES

Allodi, L. and F. Massacci (2014). Comparing vulnerability severity and exploits using case-control studies. *ACM Trans. Inf. Syst. Secur. 17*(1), 1:1–1:20.

Arrow, K. J. (1974). Optimal insurance and generalized deductibles. *Scandinavian Actuarial Journal 1974*(1), 1–42.

Arrow, K. J. and M. Priebsch (2011). Bliss, catastrophe, and rational policy. *Environmental and Resource Economics*, 1–19.

Asplund, M. (2002). Risk-averse firms in oligopoly. *International Journal of Industrial Organization 20*(7), 995–1012.

Baker, T. and S. J. Griffith (2007). Predicting corporate governance risk: Evidence from the directors'& officers' liability insurance market. *The university of Chicago law review*, 487–544.

Bencsáth, B., G. Pék, L. Buttyán, and M. Félegyházi (2012). Duqu: Analysis, detection, and lessons learned. In *ACM European Workshop on System Security (EuroSec)*, Volume 2012.

Binmore, K. (2005). *Natural Justice.* Oxford University Press.

Binmore, K. (2007). *Playing for Real.* Oxford University Press.

Caillaud, B., G. Dionne, and B. Jullien (2000). Corporate insurance with optimal financial contracting. *Economic Theory 16*(1), 77–105.

Chen, A., A. Pelsser, and M. Vellekoop (2011). Modeling non-monotone risk aversion using sahara utility functions. *Journal of Economic Theory 146*(5), 2075–2092.

Cho, C. Y., J. Caballero, C. Grier, V. Paxson, and D. Song (2010). Insights from the inside: A view of botnet management from infiltration. In *Proceedings of the 3rd USENIX Conference on Large-scale Exploits and Emergent Threats: Botnets, Spyware, Worms, and More (LEET'10)*, pp. 2–2. USENIX Association.

Cornes, R. and T. Sandler (1996). *The Theory of Externalities, Public Goods, and Club Goods*. Cambridge University Press.

Deck, C. and H. Schlesinger (2010). Exploring higher order risk effects. *The Review of Economic Studies 77*(4), 1403–1420.

Dionne, G. and K. C. Wang (2013). Does insurance fraud in automobile theft insurance fluctuate with the business cycle? *Journal of Risk and Uncertainty 47*(1), 67–92.

Eeckhoudt, L. and H. Schlesinger (2006). Putting risk in its proper place. *The American Economic Review*, 280–289.

Ehrlich, I. and G. S. Becker (1972). Market insurance, self-insurance, and self-protection. *The Journal of Political Economy*, 623–648.

Florêncio, D. and C. Herley (2010). Phishing and money mules. In *IEEE International Workshop on Information Forensics and Security (WIFS'10)*, pp. 1–5. IEEE.

Freeman, P. and H. Kunreuther (1997). *Managing Environmental Risk Through Insurance*. Kluwer Academic Publishing.

Gordon, L. and M. Loeb (2002). The economics of information security investment. *ACM Transactions on Information and Systems Security 5*(4), 438–457.

Griffith, S. J. (2006). Uncovering a gatekeeper: Why the SEC should mandate disclosure of details concerning directors' and officers' liability insurance policies. *University of Pennsylvania Law Review*, 1147–1208.

Grossman, S. J. and O. D. Hart (1982). Corporate financial structure and managerial incentives. In *The economics of information and uncertainty*, pp. 107–140. University of Chicago Press.

Hemmer, T., O. Kim, and R. E. Verrecchia (1999). Introducing convexity into optimal compensation contracts. *Journal of Accounting and Economics 28*(3), 307–327.

Herley, C. and D. Florêncio (2010). Nobody sells gold for the price of silver: Dishonesty, uncertainty and the underground economy. In *Economics of Information Security and Privacy*, pp. 33–53. Springer.

Holmström, B. and J. Tirole (2000). Liquidity and risk management. *Journal of Money, Credit and Banking*, 295–319.

Holz, T., M. Engelberth, and F. Freiling (2009). Learning more about the underground economy: A case-study of keyloggers and dropzones. In *Computer Security ESORICS 2009*. Springer.

Hong, J. (2012). The state of phishing attacks. *Communications of the ACM 55*(1), 74–81.

Hoogstraaten, H., R. Prins, D. Niggebrugge, D. Heppener, F. Groenewegen, J. Wettinck, K. Strooy, P. Arends, P. Pols, R. Kouprie, S. Moorrees, X. van Pelt, and Y. Z. Hu (August, July). Black tulip (report of the investigation into the diginotar certificate authority breach). Technical Report PR-110202, Fox-IT BV (on behalf of the Ministry of the Interior and Kingdom Relations - NL).

Ioannidis, C., D. J. Pym, and J. M. Williams (2013). Sustainability in information stewardship: Time preferences, externalities, and social co-ordination. In *The Twelfth Workshop on the Economics of Information Security (WEIS 2013)*. Available at: `http://weis2013.econinfosec.org/papers/IoannidisPymWilliamsWEIS2013.pdf`.

Johnson, S. D. (2014). How do offenders choose where to offend? perspectives from animal foraging. *Legal and Criminological Psychology 19*(2), 193–210.

Kanich, C., C. Kreibich, K. Levchenko, B. Enright, G. M. Voelker, V. Paxson, and S. Savage (2008). Spamalytics: An empirical analysis of spam marketing conversion. In *Proceedings of the 15th ACM conference on Computer and communications security*, pp. 3–14. ACM.

Kirwan, G. and A. Power (2013). *Cybercrime: The Psychology of Online Offenders*. Cambridge University Press.

Kotov, V. and F. Massacci (2013). Anatomy of exploit kits. In *Engineering Secure Software and Systems (ESSOS'2013)*, pp. 181–196. Springer.

Lelarge, M. (2012). Coordination in network security games: a monotone comparative statics approach. *CoRR abs/1208.3994*.

Li, F., A. Lai, and D. Ddl (2011). Evidence of advanced persistent threat: A case study of malware for political espionage. In *Malicious and Unwanted Software (MALWARE), 2011 6th International Conference on*, pp. 102–109. IEEE.

MacMinn, R. and J. Garven (2000). On corporate insurance. In *Handbook of insurance*, pp. 541–564. Springer.

Mandiant (2013, February). APT1 (exposing one of Chinas cyber espionage units). Technical report. Available on the web at `http://intelreport.mandiant.com/Mandiant_APT1_Report.pdf`.

Marcella Jr, A. and R. S. Greenfield (2002). *Cyber forensics: a field manual for collecting, examining, and preserving evidence of computer crimes*. CRC Press.

Mayers, D. and C. W. Smith Jr (1987). Corporate insurance and the underinvestment problem. *Journal of Risk and Insurance*, 45–54.

Moore, T. and R. Clayton (2007). Examining the impact of website take-down on phishing. In *Proceedings of the anti-phishing working groups 2nd annual eCrime researchers summit*, pp. 1–13. ACM.

Moore, T., R. Clayton, and R. Anderson (2009). The economics of online crime. *Journal of Economic Perspectives 23*(3), 3–20.

Nappa, A., M. Z. Rafique, and J. Caballero (2013). Driving in the cloud: An analysis of drive-by download operations and abuse reporting. In *Proceedings of the 10th International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment (DIMVA'13)*, pp. 1–20. Springer-Verlag.

Pal, R., L. Golubchik, K. Psounis, and P. Hui (2013). On a way to improve cyber-insurer profits when a security vendor becomes the cyber-insurer. In *IFIP Networking Conference, 2013*, pp. 1–9. IEEE.

Paper, C. S. W. (2011, June). Email attacks: This time its personal. Technical report. Available on the web at `http://www.cisco.com/c/dam/en/us/products/collateral/security/email-security-appliance/targeted_attacks.pdf`.

Paulsson, T. and R. Sproule (2002). Stochastically dominating shifts and the competitive firm. *European Journal of Operational Research 141*(1), 107–112.

Pauly, M. (1974). Overinsurance and public provision of insurance: The roles of moral hazard and adverse selection. *Quarterly Journal of Economics 88*(1), 44–62.

Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica: Journal of the Econometric Society 32*(1/2), 122–136.

Rabin, M. (2000). Risk aversion and expected-utility theory: A calibration theorem. *Econometrica 68*(5), 1281–1292.

Rajab, M., L. Ballard, N. Jagpal, P. Mavrommatis, N. P. D. Nojiri, and L. Schmidt (2011, July). Trends in circumventing web-malware detection. Technical report, Google.

Raviv, A. (1979). The design of an optimal insurance policy. *The American Economic Review*, 84–96.

Rothschild, M. and J. Stiglitz (1976). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *The Quarterly Journal of*

Economics 90(4), 629–649.

Schlesinger, H. (1981). The optimal level of deductibility in insurance contracts. *Journal of risk and insurance*, 465–481.

Schneier, B. (2001). *Secrets and Lies: Digital Security in a Networked World.* John Wiley and Sons.

Scott, R. C. and P. A. Horvath (1980). On the direction of preference for moments of higher order than the variance. *The Journal of Finance 35*(4), 915–919.

Shavell, S. (1979). On moral hazard and insurance. *Quarterly Journal of Economics 93*(2), 541–562.

Shavell, S. (1987). *Economic Analysis of Accident Law.* Harvard University Press.

Sheng, S., M. Holbrook, P. Kumaraguru, L. F. Cranor, and J. Downs (2010). Who falls for phish?: a demographic analysis of phishing susceptibility and effectiveness of interventions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 373–382. ACM.

Tankard, C. (2011). Advanced persistent threats and how to monitor and deter them. *Network security 2011*(8), 16–19.

Tirole, J. (1988). *The Theory of Industrial Organization.* MIT Press.

Williams, B. R. and A. Chuvakin (2012). *PCI Compliance: Understand and implement effective PCI data security standard compliance.* Syngress Elsevier.

Williamson, O. E. (1989). Transaction cost economics. *Handbook of industrial organization 1*, 135–182.

Zhang, D. and T. Zhang (2012). Optimal portfolio of corporate investment and consumption under market closure. *International Journal of Business 17*(1), 25.

APPENDIX

Proof of Proposition 1

Proof. At first we prove the lower bound inequality:

$$\begin{array}{rcl}\min_i\{R_i\}\cdot\sum_{i=1}^N\sigma_i(x_i,n^*) & \leq & \sum_{i=1}^N R_i\cdot\sigma_i(x_i,n^*)\\ & = & \sum_{i=1}^N n^*\cdot C_i\\ & \leq & n^* N\max_i\{C_i\}\end{array}$$

Therefore we get

$$\min_i\{R_i\}\cdot\sum_{i=1}^N\sigma_i(x_i,n^*) \quad\leq\quad n^* N\max_i\{C_i\}$$

by dividing both terms for $N\cdot\max_i\{C_i\}$ and rearranging the terms we get the desired result $\rho_-\bar{\sigma}\leq n^*$.

Then we prove the upper bound inequality by the same reasoning

$$\begin{array}{rcl}\max_i\{R_i\}\cdot\sum_{i=1}^N\sigma_i(x_i,n^*) & \geq & \sum_{i=1}^N R_i\cdot\sigma_i(x_i,n^*)\\ & = & \sum_{i=1}^N n^*\cdot C_i\\ & \geq & n^* N\min_i\{C_i\}\\ \max_i\{R_i\}\cdot\sum_{i=1}^N\sigma_i(x_i,n^*) & \geq & n^* N\min_i\{C_i\}\end{array}$$

By dividing both terms for $N\cdot\min_i C_i$ and rearranging the fractions we get $\rho^+\bar{\sigma}\geq n^*$.

Proof of Lemma 2

Proof. We consider the first order condition of the expected utility of the target in Eq. 10 and set it to zero for a given value of n_i. We obtain $\frac{\partial\sigma_i}{\partial x_i}=-\mathbb{E}[U_i']/\Delta U_i$

At first we observe that because of the concavity of U_i we have that for all $x_i\in[0,L_i]$

$$\begin{array}{rcl}U'(W_i-x_i)\leq & \mathbb{E}[U_i'] & \leq U'(W_i-x_i-L_i)\\ -\frac{U'(W_i-x_i-L_i)}{\Delta U_i}\leq & -\frac{\mathbb{E}[U_i']}{\Delta U_i} & \leq -\frac{U'(W_i-x_i)}{\Delta U_i}\end{array}$$

The concavity of U_i also implies that $\Delta U_i\leq U_i'(W_i-x_i-L_i)L_i$ and therefore $\frac{1}{\Delta U_i}\geq\frac{1}{U_i'(W_i-x_i-L_i)L_i}$ and by multiplying both terms for the negative factor $-U_i'(W_i-x_i)$ we get $-\frac{U_i'(W_i-x_i)}{\Delta U_i}\leq-\frac{U_i'(W_i-x_i)}{U_i'(W_i-x_i-L_i)L_i}$

The shape of U_i also implies that $\Delta U_i\geq U_i'(W_i-x_i)L_i$ and therefore $\frac{1}{U_i'(W_i-x_i)L_i}\geq\frac{1}{\Delta U_i}$. By multiplying both terms for $-U'(W_i-x_i-L_i)$ we get $-\frac{U'(W_i-x_i-L_i)}{U_i'(W_i-x_i)L_i}\leq-\frac{U'(W_i-x_i-L_i)}{\Delta U_i}$.

By joining the derived inequalities with the previous derivation we have

$$\begin{array}{l}-\frac{U'(W_i-x_i-L_i)}{U_i'(W_i-x_i)L_i}\leq-\frac{U_i'(W_i-x_i-L_i)}{\Delta U_i}\\ \leq-\frac{\mathbb{E}[U_i']}{\Delta U_i}\leq-\frac{U_i'(W_i-x_i)}{\Delta U_i}\leq-\frac{U_i'(W_i-x_i)}{U_i'(W_i-x_i-L_i)L_i}\end{array}$$

by multiplying all terms for the positive factor L_i and replacing the expression $-\mathbb{E}[U_i']/\Delta U_i$ at the center of the inequalities for the partial derivative of σ_i we obtain the desired result.

Proof of Theorem 1

Proof. We start by showing how to decompose $\partial\mathbb{E}[U_P]/\partial x_i$ into the three components (16), (16) and (16). Let 'not attacked' denote the state of the universe when +no successful attack has occurred against target i and 'attacked' denote the state of the universe when one attack has been successful and L_i has been suffered and a reward R_i exfiltrated.

$$\begin{array}{rcl}\frac{\partial\mathbb{E}[U_P]}{\partial x_i} & = & \sum\limits_{j=1}^N\nu_j\frac{\partial\mathbb{E}[U_j]}{\partial x_i}\\ & = & \sum_{j=1}^N\nu_i\frac{\partial\left(\sigma_i U_j(i\text{ attacked})+(1-\sigma_i)U_j(i\text{ not attacked})\right)}{\partial x_i}\\ & = & \nu_i\frac{\partial\mathbb{E}[U_i]}{\partial x_i}+\\ & & +\sum\limits_{j\neq i}\nu_j\frac{\partial\left(\sigma_i(U_j(i\text{ attack})-U_j(i\text{ not attacked}))\right)}{\partial x_i}\\ & & +\sum\limits_{j\neq i}\nu_j\frac{\partial U_j(i\text{ not attacked}))}{\partial x_i}\\ & = & \nu_i\frac{\partial\mathbb{E}[U_i]}{\partial x_i}+\\ & & +\frac{\partial\left(\sigma_i\sum_{j\neq i}\nu_j\left(U_j(i\text{ attack})-U_j(i\text{ not attacked})\right)\right)}{\partial x_i}\\ & & +\frac{\partial n^*}{\partial x_i}\frac{\partial\left(\sum_{j\neq i}\nu_j U_j(i\text{ not attacked})\right)}{\partial n^*}\end{array}$$

$$
\begin{aligned}
&= \nu_i \tfrac{\partial \mathbb{E}[U_i]}{\partial x_i} + \\
&+ \tfrac{\partial \sigma_i}{\partial x_i} \textstyle\sum_{j \neq i} \nu_j \left(U_j(i \text{ attacked}) - U_j(i \text{ not attacked}) \right) \\
&+ \sigma_i \tfrac{\partial \left(\sum_{j \neq i} \nu_j (U_j(i \text{ attacked}) - U_j(i \text{ not attacked})) \right)}{\partial x_i} \\
&+ \tfrac{\partial n^*}{\partial x_i} \tfrac{\partial \left(\sum_{j \neq i} \nu_j U_j(i \text{ not attacked}) \right)}{\partial n^*} \\
&= \nu_i \tfrac{\partial \mathbb{E}[U_i]}{\partial x_i} + \\
&+ \tfrac{\partial \sigma_i}{\partial x_i} \textstyle\sum_{j \neq i} \nu_j \left(U_j(i \text{ attacked}) - U_j(i \text{ not attacked}) \right) \\
&+ \sigma_i \tfrac{\partial n^*}{\partial x_i} \tfrac{\partial \left(\sum_{j \neq i} \nu_j (U_j(i \text{ attack}) \right)}{\partial n^*} + \\
&- \sigma_i \tfrac{\partial n^*}{\partial x_i} \tfrac{\partial \left(\sum_{j \neq i} \nu_j U_j(i \text{ not attacked}) \right)}{\partial n^*} \\
&+ \tfrac{\partial n^*}{\partial x_i} \tfrac{\partial \left(\sum_{j \neq i} \nu_j U_j(i \text{ not attacked}) \right)}{\partial n^*} \\
&= \nu_i \tfrac{\partial \mathbb{E}[U_i]}{\partial x_i} + \\
&+ \tfrac{\partial \sigma_i}{\partial x_i} \textstyle\sum_{j \neq i} \nu_j \left(U_j(i \text{ attack}) - U_j(i \text{ not attacked}) \right) \\
&+ \tfrac{\partial n^*}{\partial x_i} \sum_{j \neq i} \nu_j \left(\sigma_i \tfrac{\partial U_j(i \text{ attack})}{\partial n^*} + (1 - \sigma_i) \tfrac{\partial U_j(i \text{ not attacked})}{\partial n^*} \right)
\end{aligned}
$$

Now we add and subtract the term $\frac{\partial \sigma_i}{\partial x_i} \left(U_i(i \text{ attack}) - U_i(i \text{ not attacked}) \right)$ and then by simple manipulation we get

$$
\begin{aligned}
&= \nu_i \tfrac{\partial \mathbb{E}[U_i]}{\partial x_i} + \\
&- \nu_i \tfrac{\partial \sigma_i}{\partial x_i} \left(U_i(i \text{ attacked}) - U_i(i \text{ not attacked}) \right) \\
&+ \tfrac{\partial \sigma_i}{\partial x_i} \left(\textstyle\sum_j \nu_j \left(U_j(i \text{ attacked}) - U_j(i \text{ not attacked}) \right) \right) + \\
&\tfrac{\partial n^*}{\partial x_i} \left(\textstyle\sum_{j \neq i} \nu_j \left(\sigma_i \tfrac{\partial U_j(i \text{ attacked})}{\partial n^*} + (1 - \sigma_i) \tfrac{\partial U_j(i \text{ not attacked})}{\partial n^*} \right) \right).
\end{aligned}
$$

Now if we divide the left and right sides of this expression by ν_i, replace the difference in utility of the second and term bring the negation in front for the third term and add and subtract the term $\frac{\partial n^*}{\partial x_i} \left(\sigma_i \frac{\partial U_i(i \text{ attack})}{\partial n^*} + (1 - \sigma_i) \frac{\partial U_i(i \text{ not attacked})}{\partial n^*} \right)$, and aggregate the positive one with the fourth term, we obtain the following result

$$
\begin{aligned}
&\tfrac{1}{\nu_i} \tfrac{\partial \mathbb{E}[U_P]}{\partial x_i} = \tfrac{\partial \mathbb{E}[U_i]}{\partial x_i} + \\
&+ \tfrac{\partial \sigma_i}{\partial x_i} \Delta U_i(x_i) \\
&- \tfrac{1}{\nu_i} \tfrac{\partial \sigma_i}{\partial x_i} \left(\textstyle\sum_j \nu_j \left(U_j(i \text{ not attacked}) - U_j(i \text{ attacked}) \right) \right) \\
&- \tfrac{\partial n^*}{\partial x_i} \left(\sigma_i \tfrac{\partial U_i(i \text{ attack})}{\partial n^*} + (1 - \sigma_i) \tfrac{\partial U_i(i \text{ not attacked})}{\partial n^*} \right) \\
&+ \tfrac{1}{\nu_i} \tfrac{\partial n^*}{\partial x_i} \textstyle\sum_j \nu_j \\
&\times \left(\sigma_i \tfrac{\partial U_j(i \text{ attack})}{\partial n^*} + (1 - \sigma_i) \tfrac{\partial U_j(i \text{ not attacked})}{\partial n^*} \right)
\end{aligned}
$$

We can now observe that the third term is the delta of utility of the decision maker $\Delta U_P(x_i)$ as it is the weighted average with all ν_j of the individual differences for the individual targets. The last term is also the expected value of the derivative wrt n^* of the utility function of the decision maker because it is a ν_j-weighted sum. By re-arranging the equations we obtain the desired results.

$$
\begin{aligned}
&\frac{1}{\nu_i} \frac{\partial \mathbb{E}[U_P]}{\partial x_i} = \frac{\partial \mathbb{E}[U_i]}{\partial x_i} + \\
&+ \frac{\partial \sigma_i}{\partial x_i} \left(\Delta U_i(x_i) - \tfrac{1}{\nu_i} \Delta U_P(x_i) \right) + \\
&- \frac{\partial n^*}{\partial x_i} \left(\mathbb{E}\left[\tfrac{\partial U_i(x_i)}{\partial n^*} \right] - \tfrac{1}{\nu_i} \mathbb{E}\left[\tfrac{\partial U_P(x_i)}{\partial n^*} \right] \right)
\end{aligned}
$$

By Lemma 1 and (8) we observe that $\partial n^* / \partial x_i$ it proportional to $\partial \sigma_i / \partial x_i$ and therefore the sign of the term is negative. Further, $\Delta U_i(x_i) - \frac{1}{\nu_i} \Delta U_P(x_i)$ has negative sign as well because there is the minimal albeit not negative contribution of the remaining targets $j \neq i$. Therefore, the first term (16) has a positive sign. The key issue is now to identify the sign of the term $\mathbb{E}\left[\frac{\partial U_i(x_i)}{\partial n^*} \right] - \frac{1}{\nu_i} \mathbb{E}\left[\frac{\partial U_P(x_i)}{\partial n^*} \right]$ because it is multiplied by a factor that is positive. Clearly $\partial U_i(x_i) / \partial n^*$ is negative because both individual and overall social welfare diminish by increasing the number of attackers. The overall social welfare diminishes faster that the individual welfare and therefore the overall result is positive because of the double negation. Hence the overall term (16) is also positive.

Therefore, the value of $x_i^\dagger$ of the security investment of the policy maker happens at a point where $\partial \mathbb{E}[U_i] / \partial x_i |_{x_i = x_i^\dagger} < 0$ whereas the security investment of the unregulated target happens at the place x_i^* where $\partial \mathbb{E}[U_i] / \partial x_i |_{x_i = x_i^*} = 0$. Since U_i is weakly concave then $x_i^* \leq x_i^\dagger$.

Proof of Proposition 2

Proof. First, it is useful to show that risk neutral targets are indifferent to insurance. When target i is risk neutral, so that $U_i(w) = w$, the right-hand side of (17) reduces to the quantity:

$$
\begin{aligned}
\mathbb{E}[U_i(q_i, \ell_i, x_i, n_i)] &= (1 - \sigma_i(x_i, n_i))(W_i - x_i - q_i) + \\
&\quad \sigma_i(x_i, n_i)(W_i - x_i - q_i - \ell_i) \\
&= W_i - x_i - q_i - \sigma_i \ell_i \\
&= W_i - x_i - \sigma_i L_i
\end{aligned}
$$

As in the case of no insurance discussed previously, the target's choice of defensive expenditure minimizes the expected monetary loss. Hence, $q_i = 0$ is optimal for a risk neutral target and $L_i = \ell_i$. When target i is risk averse (17) reduces to the quantity:

$$
\begin{aligned}
&\mathbb{E}[U_i(q_i, \ell_i, x_i, n_i)] = (1 - \sigma_i) U_i(W_i - x_i - \sigma_i(L_i - \ell_i)) + \\
&\sigma_i U_i(W_i - x_i - \sigma_i(L_i - \ell_i) - \ell_i) \\
&= (1 - \sigma_i) U_i(W_i - x_i - \sigma_i L_i + \sigma_i \ell_i) + \\
&\sigma_i U_i(W_i - x_i - \sigma_i L_i + \sigma_i \ell_i - \ell_i) \\
&\leq U_i \left((1 - \sigma_i)(W_i - x_i - \sigma_i L_i + \sigma_i \ell_i) + \right. \\
&\left. \sigma_i (W_i - x_i - \sigma_i L_i + \sigma_i \ell_i - \ell_i) \right) \\
&= U_i \left(W_i - x_i - \sigma_i L_i \right) \\
&= \mathbb{E}[U_i(q_i, \ell_i = 0, x_i, n_i)])
\end{aligned}
$$

Proof of Theorem 2

Proof. At first we derive (21) for the first order condition.

$$\begin{aligned}
&\tfrac{\partial\mathbb{E}[U_P]}{\partial x_i} = \textstyle\sum_{j=1}^{N} \nu_i \tfrac{\partial U_j(W_j - x_j - \sigma_j(x_j, n^*(x_1,\ldots,x_N)L_i)}{\partial x_i}\\
&= \nu_i \tfrac{\partial U_i(W_j - x_j - \sigma_i(x_i, n^*(x_1,\ldots,x_N)L_i)}{\partial x_i} +\\
&+ \sum_{j\neq i} \nu_i \tfrac{\partial U_j(W_j - x_j - \sigma_j(x_j, n^*(x_1,\ldots,x_N)L_j)}{\partial x_i}\\
&= \nu_i U_i'(W_j - x_j - \sigma_j(x_j, n^*)) \left(-1 - \tfrac{\partial \sigma_i(x_i, n^*(x_1,\ldots,x_N)L_i)}{\partial x_i}\right) +\\
&- \sum_{j\neq i} \nu_j U_j'(W_j - x_j - \sigma_j(x_j, n^*)) \tfrac{\partial \sigma_j(x_j, n^*(x_1,\ldots,x_N)L_i)}{\partial x_i}\\
&= -\nu_i U_i'(W_j - x_j - \sigma_j(x_j, n^*)) +\\
&-\nu_i U_i'(W_j - x_j - \sigma_j(x_j, n^*))\\
&\times L_i \left(\tfrac{\partial \sigma_i(x_i, n)}{\partial x_i}\Big|_{n=n^*} + \tfrac{\partial \sigma_i(x_i, n)}{\partial n} \tfrac{\partial n^*}{\partial x_i} \right) +\\
&- \sum_{j\neq i} \nu_j U_j'(W_j - x_j - \sigma_j(x_j, n^*))\\
&\times L_j \tfrac{\partial \sigma_j(x_j, n)}{\partial n} \tfrac{\partial n^*}{\partial x_i}\\
&= -\nu_i U_i'(W_j - x_j - \sigma_j(x_j, n^*)) +\\
&-\nu_i U_i'(W_j - x_j - \sigma_j(x_j, n^*)) L_i \ \tfrac{\partial \sigma_i(x_i, n)}{\partial x_i}\Big|_{n=n^*}\\
&- \tfrac{\partial n^*}{\partial x_i} \sum_{j} \nu_j U_j'(W_j - x_j - \sigma_j(x_j, n^*)) L_j \tfrac{\partial \sigma_j(x_j, n)}{\partial n}
\end{aligned}$$

Using the chain rule in reverse the derivatives for the first and second term over ∂x_i and for last term over ∂n we obtain the following equivalence

$$\begin{aligned}
&= \nu_i \ \frac{\partial U_i}{\partial x_i}\bigg|_{n=n^*}\\
&\frac{\partial n^*}{\partial x_i} \sum_{j} \nu_j \ \frac{\partial U_j}{\partial n}\bigg|_{n=n^*}\\
&= \nu_i \ \frac{\partial \mathbb{E}[U_i]}{\partial x_i}\bigg|_{n=n^*} +\\
&\frac{\partial n^*}{\partial x_i} \ \frac{\partial \mathbb{E}[U_P]}{\partial n}\bigg|_{n=n^*}
\end{aligned}$$

The final result is obtained by dividing both left and right side of the equation for ν_i.

Now we must establish the sign of the second term of the decomposition. At first, the term $\partial n^*/partial x_i$ is negative because by Lemma 1 it is proportional to σ_i and therefore it diminishes when x_i increases. The next term is the variance of the overall utility of the policy maker as the number of attackers increases. This is clearly is negative because both individual and overall social welfare diminish by increasing the number of attackers.

Therefore, the value of $x_i^{\ddagger}$ of the security investment of the policy maker happens at a point where $\partial\mathbb{E}[U_i]/\partial x_i|_{x_i = x_i^{\ddagger}} < 0$ whereas the security investment of the unregulated target happens at the place x_i^* where $\partial\mathbb{E}[U_i]/\partial x_i|_{x_i = x_i^*} = 0$. Since U_i is weakly concave then $x_i^* \leq x_i^{\ddagger}$.

Agency Problems and Airport Security: Quantitative and Qualitative Evidence on the Impact of Security Training

Martina de Gramatica, Fabio Massacci,, Woohyun Shim,, Uğur Turhan, and Julian Williams

In this paper we analyze the issue of agency costs in aviation security by combining results from of a simple quantitative economic model with a series of semi-structured interviews with key stakeholders for an airport operating in a relatively high risk state, Turkey. We tailor two quantitative models for this purpose, the first is a standard utility based setup where agents maximize utility in respect to effort, demonstrating the classic results obtained in previous Principal-Agent models that in the absence of perfect monitoring agents, security personnel rationally choose to reduce costly cognitive and physical security effort. Our second model incorporates non-monetary welfare in the utility function in addition to potentially transferable value and hence a deepening of employee human capital. To provide context and evidence for the trade-offs elucidated in the quantitative model we have undertaken an extensive interview process with regulators, airport managers, security personnel and those tasked with training security personnel. We find that the preferred form of aligning incentives, training, may have mixed results. Indeed, taken together our results indicate that empirical determination of the relative marginal effects of transferable skills and 'intrinsic' or 'emotional' buy-in from changes in training regimes may prove challenging. As such empirical risk modeling based on historical data correlating incidents and staff skills will prove unreliable.

KEY WORDS: Semi-structured interviews; principal–agent models; public policy; mixing qualitative and quantitative analysis; security risk

1. INTRODUCTION

Airport security and the potential for the systemic failure of airport security has been a central policy question. Moreover, the appropriate training of security personnel is commonly seen as an important policy instrument. This article seeks to complement this literature by using a quantitative economic model, adapting the approach of Bernabou and Tirole (2003), combined with semi-structured interviews of key stakeholders tasked with designing and implementing security policies *in-situ* in relatively high risk environment. Our interviews are conducted following the approach of Bloom and Van Reenen (2010) and provide an open ended discussion focusing on agency costs and the role of training. The methodological contribution of this paper is to provide a quantitative treatment of risk analysis when empirical evidence is limited or unavailable.

Our theoretical model focuses on the principal-

[1] University of Trento, Italy
[2] University of Trento, Italy
[3] University of Trento, Italy
[4] Anadolu University, Turkey
[5] University of Durham, UK

agent (P–A) problem between staff, the agents, implementing security policies and a principal acting on behalf of society. The occurrence of agency problems in risk management situations is not new. However, it is instructive to note that despite the fact that incentive incompatibility is a well understood economic concept and, furthermore, that the design of optimal contracts to mitigate agency costs to principals has an equally long track record, instances of risks being substantially exacerbated by agency problems are strikingly common; we will review some of these prior examples herein. In most cases identification of the P–A problem occurs *ex post* of a significant negative event primarily for the reason that many agency costs are hidden and accrue slowly over time and it is only a significant event that uncovers them.

Quantifying the risks to travelers from terrorist attack and the porosity of boarders to illicit transport of drugs or other contraband has generally been analyzed from a technological view combined with historical data analysis of evidence to provide empirically driven probabilistic models of risk. In contrast, our objective is to combine the theoretical understanding of agency from our models and tailored to aviation security, with semi-structured interviews conducted with members of the Turkish airport security regulator, security personnel and those charged with training security personnel. This represents a significant level of access to relatively secretive group. Our key finding is that training, a pressing policy topic, that does not provide transferable skills and build human capital is ineffective in motivating staff. However, agency problems, whilst certainly present may be overcome by use of appropriate policy strategies.

P–A problems can occur throughout the hierarchies of security within the air-transport domain. We will demonstrate how classic economic models of P–A can be re-tasked to understanding the incentives for agents throughout the security management structure from public policy to firm-level and team-level agents. We will then provide clear qualitative evidence from semi-structured interviews to support the trade-offs suggested in the quantitative model and provide a series of domains for which different outcomes are anticipated. We have been provided unprecedented access to an airport and conducted extensive interviews with the various agents entrusted with security there.

A simple question arises as to why model potential P–A problems in airport security? A stated goal of airport security provision usually outlined by the agencies entrusted with this task is managing risk.[6] However, the implementation of policies on the ground is performed by staff that is usually paid at or below the national average for their respective countries. The statistics for salary and turn-over are most readily available for the US and these indicate that staff turnover is relatively high and salaries are comparatively low.[7] The complete statistics for our case study country Turkey are not widely published, but interview evidence indicates that that the comparative situation is not materially different to that of the US, see §(6).

Anecdotal evidence from the UK indicates that after the foiled 2006 transatlantic terror plot involving water bottles the banning of water bottles reduced the successful detection rates for knives and other contraband. The reason for this being that the security personnel were given targets on detections and water bottles were the ‘low-hanging-fruit’ and hence detection of these items was being used to fulfill the performance related element of the security officers renumeration.[8] However, the postulated issue is a classic P–A style problem and one that we will address in detail in our model and will feature extensively in the interviews we have conducted.

The remainder of this paper is organized as follows: §(2) provides details on the geo-political and institutional arrangements for airport security provision in Turkey and some background on our specific airport and its unique importance to Turkish aviation security. In §(4) we present a series of simple P–A models that are specifically attuned to the aviation security setting and outline the important trade-offs that factor into the creation of explicit and implicit incentive compatible contracts. §(5) carefully outlines the objectives, methodology and setting for our semi-structured interviews and §(6) then proceeds to integrate the results from these interviews with the results of the theoretical model to illustrate the agency problems and the nature of the

[6] See for instance the Transport Security Administration “About TSA” document at `http://www.tsa.gov`.

[7] For pay rates see the US governments Federal employee pay guide at the “Office of Personnel Management”, `http://www.opm.gov`, the TSA main bands lie between 1 and 5, attaining a maximum step 10 pay of just over $39,000 in 2014. The national median pay in the United States in 2013 was $51,300.

[8] For some discussion on this see: Wall Street Journal, “Why airport security if broken and how to fix it” by Kip Hawley, April 15, 2012.

Table I . Terrorist Incidents 2000–2012

Year	All Incidents			Airport		
	World	US	Turkey	World	US	Turkey
2000	167	0	1	12	0	0
2001	52	5	3	12	5	1
2002	130	5	4	16	1	1
2003	164	2	7	7	2	1
2004	234	0	13	6	0	1
2005	65	0	2	1	0	0
2006	83	1	3	4	0	1
2007	91	2	3	8	0	0
2008	92	0	3	4	0	1
2009	101	5	1	5	1	0
2010	113	1	0	6	0	0
2011	91	5	2	3	0	1
2012	187	3	2	3	0	0

Note: Number of Terrorist incidents from ITERATE database. Location codes are 640 (Turkey), 646 (Kurdistan). Incidents codes for airports are 5 (Aircraft), 8 (Embarkation area). Multiple simultaneous attacks are treated as separate incidents (e.g. The September 11, 2001 attacks on the continental US).

trade-offs facing the policy maker. Finally, in §(7) we provide some commentary on the complementarity of this type of approach when combined with frequentist empirical analysis and some final general concluding remarks and opportunities for future research.

2. AIRPORT SECURITY IN TURKEY

In recent years Turkish citizens and visitors have been the victims of several terrorist related activities and the need to protect citizens and visitors using airports is a pressing need for public policy-makers in Turkey. Table I displays the frequencies related to terrorist incidents worldwide and in Turkey from 2000 to 2012.

At a broader geopolitical level, as of 2014, Turkey has been in membership negotiations with the European Union (EU). As part of the accession a significant number of political reforms have needed to be undertaken and the variation in policy approaches to security between Turkey and the EU has required particular attention in regard to conforming with the Acquis Communtairé and other elements of EU constitutional law. In §(6) this topic is intertwined with the operation policy considerations and some commentary is provided.

In Table II we focus on some examples of major terrorist incidents in Turkey over the 2005 to 2013 period. Whilst none have directly attacked an airport, several have either been mass casualty or near miss mass casualty events of the type commonly associated with attacks on transportation networks (e.g. bomb on public transport). The broad geographical distribution of the attacks (the distance measures are relative to Istanbul in the northwest corner of Turkey) illustrates that terrorist activity should not be generalized as being isolated in specific pockets of the country, for instance on the southern border and Istanbul in exclusivity.

In Turkey, the government agency which is responsible for aviation security is the Directorate General of Civil Aviation (DGCA). The DGCA has a dedicated unit responsible for training, education, research and inspection specifically relating to airport terminal, airside and ground security. In general we will refer to aviation security as the superset of all security issues relating to air transport and 'airport security' as the subset of security practice relating specifically to terminal, airside and ground security of passengers and planes.

The DGCA is comprised of personnel from various government agencies and the state police. The DGCA is one of the main providers of security training programs for airport security staff via one of the Turkish Universities (Anadolu University), who are not normally government employees, in contrast to, for instance, the US TSA who are federal employees. A key focus of our analysis will be the impact of training as this forms a very significant component of the overall security budget. In terms of contractual liability for security incidents, the arrangements are complex only partially documented and often legally untested. This is discussed, albeit anecdotally, in some detail in §§(6.2).

Part of the DGCA's remit is in the provision of security training. All airport security personnel in Turkey are mandated to undertake a specific curriculum containing a series of three training modules, these are in Turkish and publicly available. The first module is general security training and is made available to all staff with aviation related roles. The second and third modules contain role specific training and are further sub divided into several subject specific areas, such as correct X-ray machine operational procedures. We will note in the interviews that training remit of training two general categories arise, 'general' training and specific 'technical' training. We will later show that there is a strong consensus differentiating the relative value of these two training typologies.

Our case study interviews encompass stakeholders from both the DGCA, staff trainers from Anadolu

Table II . Spatial Location of Selected Terrorist or Similar Incidents In Turkey from 2005 to 2013.

Year	Description	Location	Distance	Type of attack
2005	Kuşadası minibus bombing	Kuşacası	380km	Bombing of public transport.
2007	Ankara bombing	Ankara	350km	Suicide bombing of public space.
2007	Zirve massacre	Malatya	850km	Multiple, coordinated, stabbing incident.
2008	Diyarbakır bombing	Diyarbakır	1178km	Car bomb in public space.
2010	Hakkâri bus bombing	Geçitli, Hakkâri	1330km	Bomb attack on local public transport.
2013	Reyhanlı bombing	Reyhanlı, Hatay	850km	Mass casualty coordinated car bomb attack on public space.
2013	US embassy bombing	Ankara	350km	Targeted car bomb attack on foreign government building.

To provide a scale for the broad geographical range of terrorism incidents we have added the distance from Istanbul (located in the northwestern corner of Turkey).

University (the DGCA main training centre for airport security staff in Turkey), members of the security staff from Anadolu Airport in Eskişehir. Our interviews were conducted at Anadolu airport and it is useful to provide some context on why this airport is a useful case for study.

In addition to training security personnel Anadolu airport is the training centre for Turkish air-traffic controllers and provides training and accreditation for staff across the airport domain in Turkey. Moreover, it is a functioning airport operating within the town of Eskişehir in the region of Central Anatolia. The airport serves mainly as a hub for the town for the university students resident there. It is worth noting that the need for a reasonable sized airport in Eskişehir is due to the large number of transient students that Anadolu University supports. The university has around 23,000 students locally resident and nearly two million undertaking distance learning. The reason for this large number stems from Anadolu University's status as the primary national distance learning centre in Turkey.

Over the course of their education students are sometimes required on site and with such a large student population the turnover of travelers means that the airport of the University is the 42nd busiest airport in the second most populous state in Europe, with 50,000 passengers traveling through it in 2013. The airport itself is part of the University and provides a practical test-centre for vocational courses on all aspects of the operations of airports including security, whilst actually operating as an airport itself.

We will now provide a brief overview of research relating to the main concepts covered later on in §(4) and §(6).

3. RELATED WORK

Our work aims at linking incentive issues in the critical infrastructure security with an economic model. As with other industry sectors, critical infrastructure such as electricity, transportation and telecommunications involves the multifaceted interactions among various internal and external parties in the security environment. Therefore, security systems of critical infrastructure require a wide array of actions of the parties.

One of the main issues in this situation is that the actions taken by the participating parties might not be easily observed and monitored. In economics, this type of problems has been commonly analyzed using P–A explanations. According to Eisenhardt (1989), P–A theory provides valuable tools for studying situations in which the information is asymmetrically distributed among actors, such as the principal and collection of agents, and the actor's goals are in conflict with those of others (i.e., misaligned incentives). The theory therefore allows us to answer a question on how the principal can design a contract that make the agent behave in the best interest of the principal. The theory suggests that, if both the principal and the agent have the same information, i.e., if the agent's action is perfectly verifiable, the agent is more likely to behave in the principal's interest. In such cases, paying the agent based on his action would provide sufficient incentive to act in a way the principal expected. In contrast, if the information is not symmetrically distributed between them, i.e., if the principal cannot verify the agent's action, the agent might not behave in a way the principal expected (i.e., moral hazard). In such cases, the principal often makes a payment to the agent based on the outcome, and this would increase the agent's risk (Eisenhardt (1989); Milgrom and Roberts (1992)). In applying the P–A theory, our work is particularly related to a number of studies in the fields of economics of information and critical infrastructure security information security, and supply chain management.

In the information security literature Anderson et al. (2007) argue that even when there is more spending on information security, security breaches cannot be avoided as long as moral hazard and adverse selection caused by misaligned incentives exist. This may occur when an agent (i.e. an individual or organization) who is responsible for system security is not directly exposed to losses resulting from a security accident and monitoring is incomplete. Therefore, without proper liability sharing regimes, these P–A problems might arise and jeopardize security of the systems in part or in whole.

Our work also builds on the literature on supply chain security with multiple agents (e.g.,Atallah et al. (2006); Bakshi and Kleindorfer (2009); Bakshi and Gans (2010)). Using mathematical models and simulation, Atallah et al. (2006) discuss incentive issues in developing secure protocols with a collaborative business process between supply chain partners. They show that collaborative action can only be conducted without disclosing any private information of the partners. Bakshi and Kleindorfer (2009) demonstrate how a first-best outcome in supply chain security with asymmetric information can be achieved, when supply chain partners make some security investment. They further illustrate that, even if the retailer cannot observe the supplier's action, 'buy out' contracts can lead to a first-best outcome. Subsequently, Bakshi and Gans (2010) explore a game-theoretic model that takes into account incentive and security issues, and identified an optimal security contract encompassing the U.S. Bureau of Customs and Border Protection, the trading firms and terrorists. In particular, they discuss moral hazard issues in the context of the port security, where an important finding is that a properly designed customs-trade partnership program can provide an incentive for trading firms to join the partnership program, and makes it possible to transfer some of the government's security burden to trading firms.

In addition, our work is also related to the literature on 'intrinsic motivation', which has received attention in various research fields over the years. For instance, Murdock (2002) models the agent's incentive structure with intrinsic motivation. He particularly examines the role of an intrinsic incentive that has no direct effect on the agent's effort, and argues that intrinsic incentives and implicit contracts are complements. Bernabou and Tirole (2003) provide a formal model to discuss how explicit incentive may undermine the agents' motivation in the long run and how intrinsic motivation can improve the agents' performance. Casadesus-Masanell (2004) presents a P–A framework only taking into account a fixed payment, and shows how intrinsic motivation can promote trust in the P–A relationship. Canton (2005) also examines the power of intrinsic motivation particularly in public organizations, and identifies cases where material incentives lead to crowding-in or -out of intrinsic motivation.

4. THE MODEL

We structure this section in two parts. At first we first present a standard model of agency, in the spirit of Holmstrom and Milgrom (1987), then we take into account the effects of intrinsic incentives and training on the performance of airport employees regarding security, partly based on Casadesus-Masanell (2004) and Canton (2005).

4.1 Basic Model

We focus on a P–A interaction where the principal and the agent are both on the security provision side. Games where attackers react to the choices of the principal and the agent are possible. Recent research in this direction have indicated that the only effect of this inclusion is to significantly magnify the issues that we raise herein, by ensuring that the penalty for agency problems is even greater than when the effect of attacking effort is exogenous (e.g., Kunreuther and Heal (2003); Pym et al. (2013)).

In this study, we consider the principal as a government agency representing the social planner (hereinafter, referred to as 'the-government') and the agent as a worker conducting security on the principal's behalf.[9] As for the agent, we therefore consider both the police who are hired by the government, and security staff (e.g., security guard and X-Ray screener) who are hired by an airport to meet the goals of the government (hereinafter, we refer both of them as 'the-employee'). Security staff hired by an airport implies that an airport hires them on behalf of the government. For the US case

[9] Following the convention, we refer to the principal as 'she' and the agent as 'he'.

Table III . Description of Model Parameters and Choices.

Principal's parameters	
α	Incentive wage rate.
β	Fixed wage.
Agent's choices and parameters	
a	Employee's choice of action.
r	Employee's level of risk aversion.
Welfare effect parameters	
ρ	Emotional satisfaction/'Feeling' of responsibility.
δ	Employee's feeling of 'burden'.
γ	Marginal rate of transferability of effort to future income.
Environmental parameters	
σ^2	the variance of the shock ϵ and for convenience we set $k := r\sigma^2$.

this is very applicable as the Transportation Security Administration is a federal agency that operates within airports. For European countries as well as Turkey there is a mix of approaches and in many cases the airport directly employs the security staff and as such is an intermediary agent. Similarly to the arguments regarding reactive attackers, the presence of an added layer of agents is most commonly found to amplify the underlying P–A problems (e.g., Patacconi (2009)).

In order to model the interaction between the government and the employee, we consider that the employee needs to comply with various security rules to avoid any penalty, but his action to comply with these rules is costly to him: he is adverse to taking action. The mechanism itself and its parameters, are designed to be as simple as possible in order to focus on the behavioral issues. For reference, a summary of the model parameters and their intuition used throughout in the study is provided in Table III .

Let a be the employee's action of compliance with security rules and x be the observable informative signal (i.e., outcome) from the action, a, and an exogenous shock, ϵ: i.e., $x = a + \epsilon$. For example, we can think of a as the level of care the employee takes for ensuring security and x as the airport security level achieved by his effort. We further assume that the shock is $\epsilon \sim \mathcal{N}(0, \sigma^2)$.

Following Holmstrom and Milgrom (1987), the reward function is defined as $s(x) = \alpha x + \beta = \alpha(a + \epsilon) + \beta$ after informative signal $a + \epsilon$ has been realized. This implies that the employee has to bear some uncertainty associated with α.

Following Bernabou and Tirole (2003), the employee's cost of action is considered to be quadratic, $c(a) = a^2$, and hence is a strictly convex function with increasing marginal cost of action (i.e., $c'(a) > 0$ and $c''(a) > 0$). The employee's monetary rent from carrying out a can therefore be denoted as $\alpha(a + \epsilon) + \beta - a^2$. We also assume the risk averse employee since it is hard for him to bear any short-time financial losses with his limited resources. The employee's corresponding utility is defined as

$$u_a = -e^{-r(\alpha(a+\epsilon)+\beta-a^2)} \tag{1}$$

where r is a coefficient of constant absolute risk aversion (hereinafter, the subscript a is used to denote the agent).

We now consider the government. We assume that the government is risk-neutral since she can diversify her security portfolios by employing various security measures. The principal's risk neutrality is a well accepted assumption in the context of employment contracts (e.g., Holmstrom and Milgrom (1987)).

The government's random net benefit can be defined as $u_p = x - s(x) = (1 - \alpha)x - \beta$ (hereinafter, the subscript p is used to denote the principal). Hence we derive the certainty equivalents for both government and employee. The certainty equivalent of the government is

$$\pi_p = (1 - \alpha)a - \beta, \tag{2}$$

and the employee's certainty equivalent is

$$\pi_a = \alpha a + \beta - a^2 - \tfrac{1}{2} r\sigma^2\alpha^2. \tag{3}$$

The last term of π_a is the risk premium of the employee, and implies that, if the variability of the shock, σ^2, and/or a level of risk aversion, r, increase, the employee will feel greater risk (e.g., Casadesus-Masanell (2004)). Therefore, $r\sigma^2$ shows the employee's perceived risk.

As a starting point, suppose that the employee's action can be fully observable without costs. In this symmetric information case, the first-best contract is attainable.

Proposition 1: If the employee's action is fully observable, the optimal contract and joint surplus are as follows:

$$\alpha^\dagger = 0, \quad a^\dagger = \tfrac{1}{2}, \quad \beta^\dagger = \tfrac{1}{4}, \quad \pi_p^\dagger + \pi_a^\dagger = \tfrac{1}{4} \tag{4}$$

Proof. See Appendix: Proof of Proposition 1 □

In reality, the employee's action is largely unobservable. Therefore, while the government wants to maintain more than a certain level of security, the employee may shirk his responsibilities if he can do this without being discovered and if the expected net gains from shirking are higher than those from exerting due care. We refer this model as a benchmark model, since it will be compared with the extended model presented in the following subsection (hereinafter, superscript ‡ will be used to denote a benchmark model). In this case, the problem for identifying an optimal contract can be solved by maximizing joint surplus $\pi_a^\ddagger + \pi_p^\ddagger$ subject to the incentive compatibility constraint: i.e.,

$$\max_\alpha \pi_a + \pi_p \quad \text{subject to} \quad a \in \arg\max \pi_a. \tag{5}$$

By solving this problem, the optimal contract and joint surplus can be expressed in terms of exogenous parameters.

Proposition 2: The optimal contract and joint surplus, when the principal is unable to observe the agents effort, are as follows:

$$\begin{aligned} \alpha^\ddagger &= \frac{1}{(1+2r\sigma^2)}, && (6) \\ a^\ddagger &= \frac{\alpha}{2} \equiv \frac{1}{2(1+2r\sigma^2)} && (7) \\ \beta^\ddagger &= \tfrac{1}{4}\alpha^2\left(-1+2r\sigma^2\right) = \frac{2r\sigma^2-1}{4(1+2r\sigma^2)^2}. && (8) \\ \pi_p^\ddagger + \pi_a^\ddagger &= \frac{1}{4(1+2r\sigma^2)}. && (9) \end{aligned}$$

under the constraint that $r\sigma^2 \geq \frac{1}{2}$.

Proof. See Appendix: Proof of Proposition 2 □

From (6) and (7), $\alpha^\ddagger$ and $a^\ddagger$ are strictly decreasing in the employee's perceived risk, $r\sigma^2$. This implies that, for instance, if the employee's perceived risk becomes sufficiently large, it may push him away from exerting due effort, and thus the contract may not be maintained. The constraint stem from the observation that a worker will not work for a negative salary ($\beta \geq 0$).

Comparing Proposition 2 with Proposition 1, we see that unobservable effort and subsequent moral hazard results in the decrease in the fixed payoff and the increase in the incentive rate from 0 to $1/(1+2r\sigma^2)$. Furthermore, it decreases the employee's action level and the government's overall net benefits.

4.2 Training, responsibility, burden and human capital

There is a growing literature that indicates that an employee's payoff might be a function of intrinsic preferences such as job satisfaction and peer recognition in addition to the direct monetary rewards captured in the simplest utility frameworks (see for instance Murdock (2002); Bernabou and Tirole (2003); Casadesus-Masanell (2004); Canton (2005)). For example, Huselid et al. (1997) show that employee education and training might be able to increase the employee's intrinsic motivation, thereby increasing his effort level and reducing a moral hazard issue. Furthermore, training affords the opportunity for the agent to increase their value in the labour market by signaling the value of their human capital.

In the followings, we incorporate various intrinsic factors affecting the employee's utility, and study how the optimal contract $\langle \alpha^\ddagger, \beta^\ddagger, a^\ddagger \rangle$ identified in the benchmark model changes. We assume that these intrinsic factors depends on the employee's motivation and training. A natural utility function for the employee under these circumstances would be as follows:

$$u = -e^{-r(\alpha+\rho)(a+\epsilon)+\beta+\gamma\delta a-(1+\delta)a^2} \tag{10}$$

The utility function in (10) includes several additional parameters, together with the terms directly relating to the monetary rewards. In detail, ρ represents the level of emotional satisfaction that is fostered by the employee's sense of responsibility or altruism. According to Casadesus-Masanell (2004) and Schmidt (2007), the agent who develops this feeling might care the principal's well-being and be willing to act in the principal's best interest. This will in turn increase the agent's emotional satisfaction, ρ.

The parameter δ can be interpreted as the employee's feeling of burden within his job. For example, as for the airport employee, one of the burdens might be from the psychological or cognitive effort that is entailed with undertaking the job. However, a high δ also indicates that the job, in all likelihood, requires substantial human capital, hence the higher delta, the more potential there is for effort a to be 'transferred' into human capital.

We now consider the term, $\gamma\delta a$. γ represents the transferability of effort a and burden δ in additional human capital and can be thought of as a 'rate of forward transferability of effort' (hereinafter, referred to as 'transferability'). When $\gamma = 0$ there is no transfer from effort to forward looking human capital.

Our conjecture is that training when combined with effort has a transferable value , $\gamma > 0$ for the agent, by increasing forward employability and providing certification and evidence of effort to become a qualified trainee. Evidence from the interviews suggest that the harder the employee works the more the training is valuable for future career pathways and this factors strongly in the agents stated decision making. However, if the training does not have a forward transferable value, it would zero.

From (10), the employee's certainty equivalent payoff is now given as

$$\begin{aligned}\pi_a &= (\alpha+\rho)a+\beta+\gamma\delta a-(1+\delta)a^2 \\ &\quad -\tfrac{1}{2}r(\rho+\alpha)^2\sigma^2.\end{aligned} \tag{11}$$

Following the same procedure, as before, with the benchmark model, we can identify the optimal contract and joint surplus.

Proposition 3: The optimal contract and joint surplus, taking into account motivation, burden and transferable human capital are as follows:

$$\alpha^* = \frac{1-2r\rho\sigma^2(1+\delta)}{1+2r\sigma^2(1+\delta)} \tag{12}$$

$$a^* = \frac{\alpha+\gamma\delta+\rho}{2(1+\delta)} \equiv \frac{\gamma\delta+\frac{1+\rho}{1+2r(1+\delta)\sigma^2}}{2(1+\delta)} \tag{13}$$

$$\beta^* = \tfrac{1}{2}r(\alpha+\rho)^2\sigma^2-\frac{(\alpha+\gamma\delta+\rho)^2}{4(1+\delta)} \tag{14}$$

$$\pi_p^*+\pi_a^* = \frac{(1+\gamma\delta+\rho)^2+2r\gamma\delta(1+\delta)(2+\gamma\delta+2\rho)\sigma^2}{4\left(1+\delta+2r(1+\delta)^2\sigma^2\right)} \tag{15}$$

Proof. See Appendix: Proof of Proposition 3 □

The most notable finding from the result is that when the employee's feeling of burden, δ, is very high, the employee's effort level, a^*, depends only on the level of transferability of effort to human capital, γ (i.e., $\lim_{\delta\to\infty} a^* = \gamma/2$). This implies that, for example, even if the employee's feeling of burden from training is very high, he will exert effort as long as γ is positive.

We summarize below how the optimal solution is affected by the changes in the various parameters:

Claim 1 As either r or σ^2 increases, i.e., as the employee's risk perception increases, the power of the incentive scheme, α^*, decreases. This in turn reduces the employee's effort level and total surplus, driving them further away from the first best outcome.

Claim 2 The increase in the emotional satisfaction, ρ, also results in the reduction of the power of the incentive scheme, α^*, since the monetary incentive can be substituted by the emotional satisfaction. Yet, it raises the effort level, thereby increasing total surplus.

Claim 3 The transferability rate, γ does not impact the incentive rate α^*. However, as γ increases, the effort level and total surplus also rise closer toward the fist best outcome since the size of the return depends on the effort level.

Claim 4 Lastly, the increase in the feeling of a burden, δ, reduces the strength of the incentive scheme.

Appendix 7 provides formal explanation of the optimal α and a for the presented models with various scenarios for the parameters. For completeness we provide three additional claims on the equilibrium contract useful in elucidating the model outcomes.

5. QUALITATIVE RESEARCH METHODS

Calibration and validation of a P–A based model represents a significant challenge in part due to the variety of factors affecting the multifaceted relationships between the various actors. Unfortunately, agency costs have usually been identified ex-post after some significant event has uncovered their existence, see for instance Hausken (2002); Garber and Paté-Cornell (2012); Paté-Cornell and Cox (2014).

Traditionally, empirical studies using regression analysis are the preferred method of choice for fitting linearized P–A models to data (e.g., Knoke and Kalleberg (1994); Van Herpen et al. (2005); Fitoussi and Gurbaxani (2012)). Empirical studies based on qualitative methods for analyzing P–A relationship, are less frequently used in the literature, however several recent studies have attempted to identify incentive structures from first principles in a similar manner to our own approach (e.g., Lin and Chang (2008); Kantor and Boros (2010); Ellig and McLaughlin (2012))

Several prior studies have indicated that when an appropriate statistical model is difficult to implement, for instance the theoretical model does not have a tractable Markovian representation for econometric identification (in essence ensuring parametric models parameters can be formulated as a fair experiment in relation to the assumed data generating process) or data is simply not available, the combination of a quantitative model with qualitative evidence may be the best alternative available to decision makers. For instance, a pure frequentist approach to risk modeling without due deference to the conditions under which the data was generated may lead to inappropriate policies being enacted. Kaufmann (2004) and Cramer and Thrall (2009) both identify the problem of threat inflation in the interpretation of frequentist data on terrorist attacks. In contrast, and equally problematic, Brown and Cox (2011) argue that, without proper conditioning of attack data against appropriate controls, the very fact that the decision to attack is endogenous to the target choices means that a probabilistic risk assessment will be unable to provide meaningful insight for forward looking policy. The first issue leads to over-investment in the presence of threat inflation and the second, may increase the chances of a catastrophic security failure.

Our contribution is to fill this gap by mapping the results of a quantitative model to the on-the-ground experience of key stakeholders in aviation security through a series of semi-structured interviews. In designing a qualitative study of this type, Yin (2010) considers three features: a topic, a data collection method, and a source of data.

Firstly, we started identifying promising general issues from the thematic analysis of preliminary data we collected during several meetings with aviation industry experts and workshops with airport stakeholders, with the support of introductory interviews and exploratory questionnaires, properly designed to arouse broad subject matters (Maxwell (2009)). This first collecting phase allowed us to narrow the focus of the research into role of security staff in airport and the interplay between regulations, employment strategies, types of training and effective security.

Second, in order to find evidence of the correctness and pertinence of the model developed, we conducted a series of interviews with security-related airport stakeholders at different levels in the Turkish airport field, providing empirical data supporting the hypothesis that training can properly be considered as a mitigation measure fitting these organizational issues. Specifically, we interviewed 11 individuals, among them airport security managers, private airport security contractors and government regulators. In Table IV we provide details on the roles of the interviewees and their institutions. We do not provide their names in order to protect their anonymity. For each quote we incorporate in the text we provide a number in square brackets to reference the person, for instance [# 1] to represent the director of airport safety management systems.

We selected the focused interview method outlined in Merton et al. (1990) focusing on a topic of conversation determined in advance, with the attempt to collect reactions and interpretations in a relatively open form. Focused interviews were conducted in a semi-structured form and in a conversational mode, starting each interview using the so called ‘grand tour’ questions as discussed by Brenner (2006). We had a list of 6-7 questions, depending on the interviewees, that was circulated to the participants one week before the interviews in order to make them aware of the type of questions that will be asked. These questions are reported at the end of the Appendix.

Third, for data sources, we chose interviewees by judgmental or purposive sampling (Maxwell (2009)), in order to capture the variety of roles and activities related to aviation security. A ‘gate–keeper’, in Yin’s terminology (Yin (2011)), working at Anadolu University provided the introductions and background details for the interviews. The interviews took place aside two different workshops organized by the University of Anadolu for civil aviation security stakeholders and have been carried out in separate rooms by the same interviewers. The sampling for the interviews had been set in advance with the support of the gatekeeper. Interviews lasted approximately 30-40 minutes and sometimes a translator attended an interview, ensuring better comprehension by interviewer and interviewee. The interviews were audio recorded with the permission of the interviewees and subsequently transcribed. In parallel, hand notes have been taken during the conversation, to collect details and information about feeling, perceptions and preliminary reflections of the interviewers.

6. EVIDENCE FROM INTERVIEWS

This section initially provides an overview of how our stakeholders’ perceived risk and the impact of heterogeneity of risk perceptions between participants impacts airport security in Turkey. We

Table IV . Background of the participants in the Semi-Structured Interviews.

#ID	Role	Institution	Interview Date
1	Executive director responsible for safety	Airport	Nov 15, 2013
2	Board member for operations and regulation	Airport	Nov 15, 2013
3	Executive director responsible for safety	Airport	Nov 15, 2013
4	Board member for operations and regulation	Airport	Nov 15, 2013
5	Senior manager in charge of training programs	Civil Aviation Authority	Feb 27, 2014
6	Senior manager in charge of training programs	Civil Aviation Authority	Feb 27, 2014
7	Chief of Security Operations	Private Security Contractor	Feb 28, 2014
8	Chief of Security Operations	Civil Aviation Authority	Feb 28, 2014
9	Senior airport manager	Airport	Nov 15, 2013
10	Senior airport manager	Airport	Nov 15, 2013
11	Senior manager in charge of training programs	Airport	Nov 15, 2013

then provide answers to questions that illustrate some of the specific channels of agency costs that we have quantified in our model alongside the representative parameters and their domains. Where we have summarized general points put forward by one or more of the interviewees we reference them using square brackets, for instance [# 1]. Specific quotes are reported in italics with the attribution placed before the quote, once again square brackets marked with a colon.

6.1 General Information and Risk Perception

The complex geo-political situation in Turkey is perceived to have an impact on the airport security domain. This is reported by some interviewees stating that sacrifice of democratic principles, such as freedom and privacy, is needed in order to mitigate the concern towards security objectives. An important characteristic that the interviewees exhibited was a high level of 'philosophical' alignment with the overarching policy objectives of the principal. The following extracts relate to the institutional and societal factors that can affect security effectiveness.

[# 2]: *"Turkish people are used to be checked with x-ray, even to enter into a mall they are X-ray checked. We want to keep this security measure. [Interviewer: isn't it very expensive?] Sure, but if something bad happens, then it will be more expensive. [...] I do not want anything bad happens. If you want to travel, you are checked and that is all. If you do not want, you do not travel."*

Most interviewees strongly supported the implementation of a wider detection system in strict collaboration with the intelligence, hoping that

[# 3]: *"Once you arrive at the airport, everything should be already done."*

The perceptions of risk displayed by the stakeholders was somewhat heterogeneous. Prior research on qualitative evaluation of risk perception, the suggest that one of the factor that shapes the attitude toward risk is the trust expressed for the security rules (e.g., Peters et al. (1997); Viklund (2003)). In the course of our interviews, we noted the interviewees' general dissatisfaction on the current security regulations for the airport security, perceived as weak and incomplete support in fighting risks and threats, rather than as a useful guideline to improve security.

[# 1]: *"In the (airport security) regulations, there are few things about practice [that] matters. They are based on regulatory compliance. If you are compliant with a regulation, the government think you are a secure one. [...] For example, [the government inspects whether] you use the tools that are requested. Yes or no, black or white? But what about the other things?"*

A consensus of the interviewees perceived the regulations to be a list of mandatory duties that managers were required to adhere to without substantive added value to the overall level of security.

Another important factor that affects the risk perception is the relationships between the authorities designated for the application of the security rules (e.g., Peters et al. (1997); Viklund (2003)).

The majority of interviewees expressed opinions on the poor cooperation between the various actors involved in airport security, particularly between security staff and police officers.

[# 10]: *"They (police officers) think that the whole department is belonging to them. [sic] They are out of training, they do not have specific info on airport security. [Interviewer: What happens if something happens?] Police takes responsibility on this. [Interviewer: Would* *it better to have only private guards?] No, police is really needed, but educated police."*

In the following subsections, we explore various agency problems experienced by Turkish airports and apply the results of our model to investigate the effects of the employee's intrinsic motivations.

6.2 Agency Costs, Employment Rules and Roles

The DGCA regularly conduct inspections and security audits on an airport [# 1, 4, 5, 6, 8, 11] for the purposes of monitoring. However, all of the participants indicated that monitoring was incomplete, the DGCA would not be able to observe actions of the participants including airport employees perfectly. A typical P–A explanation suggests that a principal will attempt to design a contract that ensures that an agent bears, in whole or in part, the expected costs of shirking. In airport security, however, this might not be feasible. Cumulative impacts of shirking effort can be very high, however the marginal impact on security for lack of effort by single members of staff is, in the main, very low.

Risks related to a terrorist events have high impact and occur with with very low probability of occurrence (Belzer and Swan (2011)). Once a terrorist event occurs, the security personnel might not be able to pay the damage, and hence the security risk cannot be transferred to the agent from the principal. Gross dereliction of duty notwithstanding. The precise chain of events leading to a successful terrorist attack are usually very difficult to precisely reconstruct (Enders and Sandler (2011)). The ability of an attacker to gain the information needed for a successful attack may have been collected weeks earlier by observing other agents not correctly performing their task. Together with imperfect monitoring, therefore, this can result in a sub-optimal contract, from the viewpoint of incentive compatibility, between the principal and the agent.

The Turkish job market is very difficult and many people will accept a job even if they can if the salary places their reservation utility at zero. This also implies the employee, particularly the security staff, feels high level of perceived risk in the job market (i.e., $r\sigma^2$ is high). Additionally, the interviewees stated that an employment contract for security staff in an airport is based on a fixed wage contract (i.e., $\alpha \approx 0$) [# 3, 7], and generally attracts workers who only have limited job alternatives (i.e.,$\gamma \approx 0$) [# 1, 7], this finding is similar to case studies in supply chain security (e.g., Belzer and Swan (2011))

[# 7]: *"Payment (for security staff) is very low. For this reason, a lot of [sic] person change job, security persons do not think that this is a very important job. They just come, work little time and then they leave."*

Additionally, interviewee [# 3] presented an argument that most of the security staff do not aware the importance of their role and not feel responsibility or motivation (i.e., $\rho \approx 0$). The low wages for security staff and quality of employees results in high employee turnover rate [# 1, 4].

[# 1]: *"[G]uys working for these security companies (in an airport) have no other choices for working so they have to work there if they want to earn money, but the problem is that they are not motivated enough."*

This is consistent with the interpretation of (13), where low levels of monetary and non-monetary incentives and high level of the employee's perceived risk result in a low level of effort.

As previously stated the airport as an organization and the DGCA have a complex liability sharing arrangements. Interviewees [# 6] discusses that an airport operator is responsible for any damage from a security related event. The operators then will shift some of their risks to their employees. However, if the employees are compensated with low wages, society at large will be liable for the whole costs of a security failure (Belzer and Swan (2011)). The P–A problem is often found to increase when opportunities to switch employment as the degree of human capital invested by the agent in his position is far lower.

It should be noted that, in spite of low wages, many security activities (e.g., liquid detection) are relatively easy to monitor as technology has automated many of the processes and staff therefore only need to respond to an alarm, rather than engage in costly cognitive effort to mange the ongoing operations. Other security activities (e.g., X–Ray screening), in contrast, require cognitive and physical effort and in many cases monitoring the employees' actions is expensive expensive or, indeed, logistically impossible.

6.3 Motivation and Security Training

There is substantial evidence from prior literature that compensation, both sunk and performance related is a driving factor in the effort exerted by employees (e.g., Quinlan and Wright (2008); Belzer and Swan (2011)) For example, when the participants in the aviation security try to employ security staff, those who have high opportunity cost (i.e., highly qualified workers) will only be attracted by high incentives. The corresponding evidence can also be found in a report published by United States. General Accounting Office (2000). According to the report, one of the main reasons that airport screeners do not perform their work properly is partly because of a low level of compensation which prevents an airport from hiring and retaining qualified workers with high intrinsic motivation.

The perception of interviewees in the Turkish case, [# 1, 3, 7], indicated that the level of pay, compared to other service worker, will not change significantly in the near future, although some modest increases have been noted recently. However, some previous studies have indicated that motivating employees by increasing their intrinsic preferences can improve the gap in optimal effort perceived between the principal and the agent (e.g, Murdock (2002); Bernabou and Tirole (2003); Canton (2005)).

In this subsection, we present the details of the interview results focusing on the employee's intrinsic preferences, and link them with our extended model in the previous section to explore whether the employee's intrinsic preferences can foster his motivation appropriately to make him exert due effort.

The interviewees identified differences in intrinsic motivation between the airport staff and the police, recalling the earlier comments on cooperation. The cultural role of the police within Turkish civil society was indicated to be an important driver of this sense of civic responsibility and hence reduced the agency costs we have previously identified. This however, appeared to have a negative impact on the security staff who perceived a degree of exclusion from this culture.

[# 1]: *"Security staff just help the police and they only have limited responsibility because the responsibility is taken from the state security department (i.e., the police). Since the department do not have enough police officers, airport security staff are needed as well."*

Uniformly, the stakeholders noted that the police are working directly for the government and follow a different statute and culture [# 3, 7]. While the police have more power and responsibilities [# 3], airport operators do not have the right to audit them because they belong to the state [# 7]. Yet, Airport police officers are not specifically trained for airport security [# 2, 3], and do not have a security training program specifically designed for them.

[# 2]: *"Police officers working near the Syrian boundaries have to be really very careful about possible terrorist attacks; they work there and then after 3 years they come to our airport and they behave the same. This is not good because the context has changes, is really different...They read the regulations we have but they do not know which is the difference between should/would/must/could."*

On the other hand, airport police officers have more responsibilities than airport security staff. When there is a security event, security staff need to report to the police and the police 'have the final responsibility' paraphrased from comments by [# 1, 3, 7, 10]. They seem to feel responsible for airport security. Furthermore, the police have a higher fixed wage than the security staff [# 7], which can attract better qualified and motivated applicants. Taken these together, we can infer that, in our motivation and training model, the burden of training δ is very low whereas the police's feeling of responsibility ρ would be higher than zero.

As a result, even if their salary is based on a fixed wage (i.e., $\alpha \approx 0$), the police might exert a positive level of effort which can mitigate some moral hazard problem. More specifically, from (13), their optimal effort level can be regarded as $a = \frac{\rho}{2}$ with $\alpha = 0$). An alternative explanation, that does not support the culture and intrinsic responsibility for increased effort by the police is that their effort maybe linked to transferable value from effort. Police often change duties and agglomerate experience and know-how. This may prove valuable in their future career and as such exhibiting greater effort provides direct utility to them via the standard rational utility maximization mechanism, $\gamma > 0$.

[# 7]: *"The problem is that they change, they do not know what airport security is. Sometimes in 6 months they change role twice. They change job position very often, they are not trained on the civil aviation security. In 6 months it could happen that they have to change 3 times their job."*

A common feature of the interviewees responses to interaction with law enforcement was the perception that the police's expertise and sometimes their motivation was very low.

[# 2]: *"They should have an appropriate and suitable training to do the security at the airport, and this training* *is different from the training required for the Syrian boundaries."*

As a result, it is unclear whether the increased effort level due to higher ρ can effectively increase the social surplus. The capacity to generate a tangible greater-than-zero γ for airport staff is very hard to gauge and we shall address this in more detail.

For security staff, motivation is essentially conducted through security training (Huselid et al. (1997)). The argument being, that intrinsic motivation can be a mechanism for reducing the P–A effort gap.

The interviewees provide further details on training programs which use two different types of approaches: 'strategic' and 'technical'. From the interviewees, training using a strategic approach aims at providing efficiency that ensure the achievement of a firm's general business objectives (henceforth, referred to as 'general training'), while training with a technical approach focuses on shaping a wide range of technical and professional practices (hereinafter, referred to as 'technical training').

In Turkey, most of the security training programs are designed and provided by DGCA: while there are also private agencies which provide training programs, particularly for more specific security technologies, training programs provided by the state is the main source to educate the employees in an airport. According to the interviewees, whereas the quality of private training programs is better than the programs offered by the state, these are less widely used since this is a more costly provision [# 7, 10]. Therefore, we focus our exploration only on training programs offered by DGCA.

The interviewees indicated that the security training for all airports, is effectively the same no matter the size.

[# 6]: *"We have training for all people involved in the airport security, as this personnel could be a potential threat to the security of the airport. [sic] In airport every person has a role and a duty in aviation security, so we need to train all of them in order to provide total security. We have to train them in aviation security procedures, national and international as well."*

During the interviews we identified that there are three modules for security training developed and mandated by the regulator [# 6, 11]. Specifically, Module 1 is security awareness training that is mandated for all attendants, staff and managers in an airport. Modules 2 and 3 are for training security staff including X-ray and metal detector operators and cabin crew.

In line with the publicly stated training curriculum by the DGCA the interviewees noted the three compulsory training modules developed and mandated by the regulator [# 6, 11]. While Module 1 is for transferring general security knowledge and increasing security awareness, Modules 2 and 3 focus more on transferring specific knowledge for a certain security work. According to interviewee [# 11], Modules 2 and 3 are compulsory and are more specific compared to Module 1. Every airport has to follow the procedures for Modules 2 and 3, very precisely and these are the necessary application mandated by the rules. It was further noted that Modules 2 and 3 require more resources and information for training. Interviewee [# 5] also stated that, while DGCA does not have any different implementation procedures for security awareness training, it does have specific training programs for educating the personnel working in different roles.

[# 5]: *"Security awareness training is for everyone in the airport because it is an indispensable part of airport security. On the other hand, training implementation has to be different for different roles; you cannot implement the same rules for cabin crew and ground service people or screening staff in security check points."*

Therefore, airport employees have different training depending on what their duties are. Training is differentiated from person to person, and from department to department. The differentiation is highly dependent on the department for whom it is designed. Interviewees stated that most staff have little security experience and need to be trained from scratch.

[# 1]: *"We have good security devices. However, there are not enough security training agencies in Turkey particularly specialized in aviation security. They are not efficient, so even if we had more money to invest, it would be difficult to find a good training. Training is mandated but not enough. We have to pay for further training…[It is] very difficult to train them. This is a general problem in Turkey, they do not earn a lot of money but they do a very critical job."*

Other interviewees also have a similar view on the security training. There was unanimity if regarding security training as the main issue in Turkish airport security. Arguing that personalized training schedules should be implemented and specifically planned for both general and specific security training. From the perspective of the quantitative analysis, a personalized record of training permits the agent to 'deepen' their personal human capital, $\gamma > 0$ and increase motivation $\rho > 0$.

As previously indicated, the payment scheme for security staff is based on a fixed salary β and the incentive rate for exerting effort is quite weak (i.e., $\alpha \approx 0$). Whilst the principals in this arrangement appear to identify with this issue, their approach is specifically in raising fixed salaries. The primary driver behind their rationale is that with higher salary their is positive movement in the agents motivation (i.e., $\rho > 0$) and the employees will subsequently exert more effort.

[# 7]: *"[...] security personnel has a big responsibility. So last year, we decided to raise their salary and now we pay them more. The situation now is better."*

However, many airports in Turkey are not able to afford the additional costs associated with the increased salary, and tend to depend on security training offered by DGCA attempting to raise the employee's intrinsic motivation. The interview results indicate that most of the interviewees believe that security training can mitigate a P–A problem to an extent.

[# 3]: *"You cannot easily change the physical environment but you can change people. So we have to improve training (and) people's vision [...] If you are more trained you feel more confident [...]"*

We will initially investigate whether a general training program (i.e., Module 1) can effectively incentivize the employee to exert due effort, and reduce moral hazard. General training programs are provided in the classroom environment. Several interviewees [# 10,11] stated that classroom training programs are boring and trainees were not motivated to follow it.

Linking this with our motivation and training model, it implies that training for general knowledge transfer of security might only incur a burden on the employees (i.e., $\delta > 0$) and will not provide the employees with the recognition of their role in ensuring airport security (i.e., $\rho \approx 0$) as indicated by interviewee [# 11].

General security training does not provide a specific certification to a qualified trainee and does not require an exam. Employees only need to retake a training program once in every 3 years. This implies that the general security training does not provide any information on the employee's repute and not increase his level of employability (i.e., $\gamma \approx 0$). Consequently, general security training might not be helpful to increase employees' motivation and to reduce moral hazard (i.e., $a \approx 0$) — indicated in Claim 7 in the Appendix. Indeed, this is common impression some of the interviewees have expressed about general security training.

We now consider the effectiveness of a training program aiming at transferring specific technical knowledge. In the interview with a training manager [# 11], he explained that Modules 2 and 3 are carried out by on-the-job training and practical exercises as well as classroom lectures. This approach was deemed to be very effective in motivating trainees and in attaining skills (i.e., $\rho > 0$) while these cause higher burden on the trainees than a general training program (i.e., $\delta > 0$).

The other facet of specific technical training is the mandatory renewal of employees certification and the possible loss of the job due to the failure during this renewal process [# 5, 6, 11]. In Turkey, security staff who need to take the Module 2 and 3 training programs to renew their certification every 2 years. This is accomplished through an examination conducted by the Training Department of Aviation Security. If they cannot pass the exam, their certification is canceled and they cannot no longer work for an airport.

[# 7]: *"(When we hire new employees,) we evaluate them before they start working in the airport. We use some procedures. We check their experience, if they have been working for at least 3 years, and then we evaluate them.[. . .] When we are selecting persons, we use a lot of criteria. For example, we need to know whether X-ray operators are able to use that technology, so we need to have an examination, [sic] because probably they have no experience."*

We can interpret technical training as providing a degree of transferable value from effort. This type of training provides certification that can be used for later employment and normally entails some managed supervision *in situ*.

[# 7]: *"There is a special team for checking: we need to evaluate people before they start working in the airport; we use some procedures; we want to know their experience, if they have been working for at least 3 years and then we evaluate them. [If the level if very low we do not hire them. Since our salary is higher than others, there are a lot of people that want to work with us. For these reasons, when we are selecting persons we use a lot of criteria."*

From our analysis in Section 4, the optimal effort level is $a = \frac{\gamma\delta+\rho}{2(1+\delta)}$ from inspection we can see that this is always greater than the effort level without monetary and intrinsic incentives. Furthermore, even if the employee's feeling of burden is very high, the training can still lead to a positive level of effort as long as the level of transferability has a positive value (i.e., $a = \frac{1}{2}r$).

A core conclusion appears to be that specific technical training can develop the employees' motivation and understanding of the rationale behind their tasks, hence mitigate a moral hazard problem, however transferability of value from effort appears to be a important important factor in the employees pay-off function.

We can think of this as 'buy-in' by the agents into a longer term set of goals. From the interviews, it is clear that the perception of each of the agents understands and agrees with the broader goals, but the specificity of the individual micro-tasks is also important in achieving the most efficient risk reduction for a given level of investment in labor and capital.

7. CONCLUSIONS

This study seeks to elucidate the issues surrounding the incentive structure for workers engaged in facilitating risk reduction in an important security setting. In the study, we consider not only a traditional monetary incentives, but also non-monetary utility intrinsic to the agent.

In the our modeling section, we have outlined a pair of models that specifically addresses the optimal contract to align incentives within an airport security setting. Our first model mimics the typical principal agent setting with only pay-off directly (in a certainty equivalent 'monetized' value) relating to effort entering into the optimal solution. Our second model, incorporates trade-offs in welfare that contain feelings of well being not strictly associated with financial pay-offs.

We have then identified a set of potential trade-offs in terms of effort on behalf of the individual security agent versus a remuneration contract that combines static wages and incremental contributions. For ease of elucidation of the specific P–A effect for which we are specifically interested, our risk generating mechanism assumes a 'non-strategic' exogenous attacker. Whilst we have not quantitatively analyzed the extension to a strategic attacker it would be anticipated that targeted attacks that target security lapses from agency costs would in general be more successful, magnifying the costs of the effects we have identified. Therefore, our findings indicate that the critical trade-offs for the agents explicitly addressed in the model and lent weight by the case study are critical issues for society at large and field experiments that attempt to target training and deepen the intrinsic capital security staff should be carefully looked at.

Prior studies of *ex-post* failings in complex socio technical systems (in relation to both security events and accidents) have often outlined causal failures in correctly identifying where incentives are not aligned correctly, see Suzuki (2014), page 1251 for a relevant, but unfortunately after-the-fact summary of agency costs, moral hazard and information asymmetry in nuclear safety. Other numerous recent examples from financial services indicate, for instance, that not incorporating the agency costs of debt in loan contracts can lead to unforeseen and potentially extremely costly events. A study, such as this one, seeks to identify P–A issues *a-priori* to help reduce the likelihood of catastrophic security failures by illustrating to the a policy maker the type risk structure that they are faced with.

ACKNOWLEDGEMENTS

This work has been partly funded by the European Union's 7th Framework Programme under grant agreement no 285223 - SECONOMICS (www.seconomics.org).

REFERENCES

Anderson, R., T. Moore, S. Nagaraja, and A. Ozment (2007). Incentives and information security. In N. Nisan, T. Roughgarden, E. Tardos, and V. Vazirani (Eds.), *Algorithmic Game Theory*, pp. 633–650. Cambridge University Press.

Atallah, M., M. Blanton, V. Deshpande, K. Frikken, J. Li, and L. Schwarz (2006). Secure collaborative planning, forecasting, and replenishment (scpfr). In *Multi-Echelon/Public Applications of Supply Chain Management Conference*.

Bakshi, N. and N. Gans (2010). Securing the containerized supply chain: Analysis of government incentives for private investment. *Management Science 56*(2), 219–233.

Bakshi, N. and P. Kleindorfer (2009). Co-opetition and investment for supply-chain resilience. *Production and Operations Management 18*(6), 583–603.

Belzer, M. H. and P. F. Swan (2011). Supply chain security: Agency theory and port drayage drivers. *The Economic and Labour Relations Review 22*(1), 41–63.

Bernabou, R. and J. Tirole (2003). Intrinsic and extrinsic motivation. *The Review of Economic Studies 70*(3), 489–520.

Bloom, N. and J. Van Reenen (2010). New approaches to surveying organizations. *American Economic Review 100*(2), 105–109.

Brenner, M. E. (2006). Interviewing in educational research. In J. Green, G. Camilli, and P. Elomere (Eds.), *Handbook of Complementary Methods in Education Research (3rd Ed.)*, pp. 357–370. American Educational Research Association, WA.

Brown, G. G. and L. A. Cox, Jr. (2011). How probabilistic risk assessment can mislead terrorism risk analysts. *Risk Analysis 31*(2), 196–204.

Canton, E. (2005). Power of incentives in public organizations when employees are intrinsically motivated. *Journal of Institutional and Theoretical Economics JITE 161*(4), 664–680.

Casadesus-Masanell, R. (2004). Trust in agency. *Journal of Economics & Management Strategy 13*(3), 375–404.

Cramer, J. K. and A. T. Thrall (2009). Understanding threat inflation. In A. T. Thrall and J. K. Cramer (Eds.), *American foreign policy and the politics of fear: Threat inflation since 9/11*. Routledge.

Eisenhardt, K. M. (1989). Agency theory: An assessment and review. *Academy of Management Review 14*(1), 57–74.

Ellig, J. and P. A. McLaughlin (2012). The quality and use of regulatory analysis in 2008. *Risk Analysis 32*(5), 855–880.

Enders, W. and T. Sandler (2011). *The Political Economy of Terrorism*. Cambridge University Press.

Fitoussi, D. and V. Gurbaxani (2012). It outsourcing contracts and performance measurement. *Information Systems Research 23*(1), 129–143.

Garber, R. and E. Paté-Cornell (2012). Shortcuts in complex engineering systems: A principal-agent approach to risk management. *Risk Analysis 32*(5), 836–854.

Hausken, K. (2002). Probabilistic risk analysis and game theory. *Risk Analysis 22*(1), 17–27.

Holmstrom, B. and P. Milgrom (1987). Aggregation and linearity in the provision of intertemporal incentives. *Econometrica 55*(2), pp. 303–328.

Huselid, M. A., S. E. Jackson, and R. S. Schuler (1997). Technical and strategic human resources management effectiveness as determinants of firm performance. *Academy of Management Journal 40*(1), 171–188.

Kantor, P. and E. Boros (2010). Deceptive detection methods for effective security with inadequate budgets: The testing power index. *Risk analysis 30*(4), 663–673.

Kaufmann, C. (2004). Threat inflation and the failure of the marketplace of ideas: The selling of the iraq war. *International Security 29*(1), 5–48.

Knoke, D. and A. L. Kalleberg (1994). Job training in us organizations. *American Sociological Review*, 537–546.

Kunreuther, H. and G. Heal (2003). Interdependent security. *Journal of Risk and Uncertainty 26*(2), 231–249.

Lin, Y. H. and Y. H. Chang (2008). Significant factors of aviation insurance and risk management strategy: An empirical study of taiwanese airline carriers. *Risk analysis 28*(2), 453–461.

Maxwell, J. A. (2009). Designing a qualitative study. In L. Bockman and D. Rog (Eds.), *The SAGE Handbook of Applied Social Research Methods (2nd Ed.)*, pp. 69–100. SAGE Publication Inc., CA.

Merton, R., M. Fiske, and P. Kendall (1990). *The focused interviews: A manual of problems and procedures* (Second ed.). New Your, NY: Free Press.

Milgrom, P. R. and J. Roberts (1992). *Economics, organization and management*, Volume 7. Prentice-hall Englewood Cliffs, NJ.

Murdock, K. (2002). Intrinsic motivation and optimal incentive contracts. *The RAND Journal of Economics 33*(4), pp. 650–671.

Patacconi, A. (2009). Coordination and delay in hierarchies. *RAND Journal of Economics 40*(1), 190–208.

Paté-Cornell, E. and L. A. Cox (2014). Improving risk management: From lame excuses to principled practice. *Risk Analysis 34*(7), 1228–1239.

Peters, R. G., V. T. Covello, and D. B. McCallum (1997). The determinants of trust and credibility in environmental risk communication: An empirical study. *Risk Analysis 17*(1), 43–54.

Pym, D., C. Ioannidis, and J. Williams (2013). Sustainability in information security. In *Workshop on the Economics of Information Security 2011*.

Quinlan, M. and L. Wright (2008). Safe payments: Addressing the underlying causes of unsafe practices in the road transport industry. *Melbourne. National Transport Commission. October*, 78.

Schmidt, S. W. (2007). The relationship between satisfaction with workplace training and overall job satisfaction. *Human Resource Development Quarterly 18*(4), 481–498.

Suzuki, A. (2014). Managing the fukushima challenge. *Risk Analysis 34*(7), 1240–1256.

United States. General Accounting Office (2000). Aviation security: Long-standing problems impair airport screeners' performance: Report to congressional requesters. United States General Accounting Office.

Van Herpen, M., M. Van Praag, and K. Cools (2005). The effects of performance measurement and compensation on motivation: An empirical study. *De Economist 153*(3), 303–329.

Viklund, M. (2003). Trust and risk perception in western europe: A cross national study. *Risk Analysis 4*(23), 727–738.

Yin, R. K. (2010). *Qualitative research from start to finish*. New Your, NY: Guilford Press.

Yin, R. K. (2011). *Applications of Case Study Research*. SAGE Publications, Inc.

APPENDIX

Proofs of Propositions

This section provides the proofs of the propositions and the claims presented in §(4). It is intended to be an electronic supplement.

As a preliminary result we derive the certainty equivalent equations (2) and (3). Suppose that the risk averse employee has an exponential utility function $u_a = -e^{-rw}$, where $w = s(x) - a^2$ and $w \sim \mathcal{N}(\mu, \sigma^2)$. The corresponding density function for w is given as

$$f(w) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left(-\frac{(w-\mu)^2}{2\sigma^2}\right)}.$$

Therefore, the expected utility can be defined as

$$\begin{aligned}
\mathbb{E}[u_a] &= -\mathbb{E}[-e^{-rw}] \\
&= -\int_{-\infty}^{\infty} e^{-rw} \frac{1}{\sigma\sqrt{2\pi}} e^{\left(-\frac{(w-\mu)^2}{2\sigma^2}\right)} dw \\
&= -\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\left(-rw-\frac{(w-\mu)^2}{2\sigma^2}\right)} dw.
\end{aligned}$$

Noting that

$$\begin{aligned}
&-rw-\frac{(w-\mu)^2}{2\sigma^2}\\
&=-rw-\frac{(w-\mu)^2}{2\sigma^2}+r\mu-r\mu+\frac{r^2\sigma^2}{2}-\frac{r^2\sigma^2}{2}\\
&=-\tfrac{1}{2}\left(2r(w-\mu)+\frac{(w-\mu)^2}{\sigma^2}+r^2\sigma^2\right)-r\mu+\frac{r^2\sigma^2}{2}\\
&=-\frac{1}{2\sigma^2}((w-\mu)+r\sigma^2)^2-r\mu+\frac{r^2\sigma^2}{2}.
\end{aligned}$$

From this, we can see that

$$\mathbb{E}[u_a] \quad = \quad -\frac{1}{\sigma\sqrt{2\pi}}e^{\left(-r\mu+\frac{r^2\sigma^2}{2}\right)}\int_{-\infty}^{\infty}e^{\left(-\frac{1}{2\sigma^2}((w-\mu)+r\sigma^2)^2\right)}dw$$

by setting

$$y \quad = \quad \frac{((w-\mu)+r\sigma^2)}{\sigma^2}$$

and under the normality assumption of y we obtain

$$\begin{aligned}
\mathbb{E}[u_a] \quad &= \quad -\frac{1}{\sigma\sqrt{2\pi}}e^{\left(-r\mu+\frac{r^2\sigma^2}{2}\right)}\int_{-\infty}^{\infty}e^{\left(-\frac{y^2}{2}\right)}dy\\
&=-\frac{1}{\sigma\sqrt{2\pi}}e^{\left(-r\mu+\frac{r^2\sigma^2}{2}\right)}\sigma\sqrt{2\pi}\\
&=-e^{-r\left(\mu+\frac{r\sigma^2}{2}\right)}.
\end{aligned}$$

From the certainty equivalent theorem, $u(\pi_a)=E[u_a]$. We therefore get

$$\pi_a=\mu-\frac{r\sigma^2}{2}=\mathbb{E}(w)-\frac{rVar(w)}{2}.$$

Since $w=s(x)-a^2=\alpha(a+\epsilon)+\beta-a^2$, the agent's certainty equivalent is given as

$$\pi_a=\alpha a+\beta-a^2-\tfrac{1}{2}r\alpha^2\sigma^2.$$

The government is risk neutral and has net benefit $u_p=(1-\alpha)(a+\epsilon)-\beta$. Therefore, the government's expected net benefit can be defined as

$$\mathbb{E}[u_p] \quad = \quad a-\alpha a-\beta.$$

Since $u(\pi_p)=\mathbb{E}[u_p]$, the government's certainty equivalent is

$$\pi_p \quad = \quad a-\alpha a-\beta.$$

Proof. [Proposition 1] If the employee's action is observable without costs, the government does not need to take an incentive compatibility constraint into account, and only needs to pay the employee for his action that can guarantee his participation. Hence, the employee's participation constraint holds with equality, and we set the employee's reservation utility equals to zero (i.e., $\pi_a=0$). The government's problem is then to solve the following maximization problem.

$$\max_a \pi_a+\pi_p=\max_a a-a^2-\tfrac{1}{2}r\alpha^2\sigma^2. \tag{A.1}$$

It entails the employee to make the level of action $a^\dagger=1/2$. Inserting this value into the joint surplus and maximizing it with respect to α yields $\alpha^\dagger=0$. Using $a^\dagger$ and $\alpha^\dagger$ in the participation constraint, we get $\beta^\dagger=1/4$ which equals to the cost of his action. Consequently, the government gets net benefits of 1/4.

Proof. [Proposition 2] In order to identify optimal α and β, we first need to explore the employee's problem. Since his problem is to identify an optimal effort level that can maximize π_a for given α and β, it can be denoted as $\max_a \pi_a$ and gives the first-order condition $a^\ddagger=\alpha/2$. Therefore, if incentive wage is not provided (i.e., $\alpha=0$), the employee will not carry out any action (i.e., $a=0$). This condition shows that an optimal action level is only determined by an incentive rate α. Moreover, the condition also means that the employee's marginal benefits of action (i.e., marginal expected reward) are equal to his marginal costs of action.

By inserting optimal effort level $a^\ddagger=\alpha/2$ into (5), we can drop the incentive compatibility constraint and rewrite it as:

$$\max_\alpha \frac{\alpha}{2}-\left(\frac{\alpha}{2}\right)^2-\tfrac{1}{2}r\alpha^2\sigma^2. \tag{A.2}$$

This problem has the first-order condition $1/2-\alpha/2-r\alpha\sigma^2=0$. Rearranging this equation with respect to α gives

$$\alpha^\ddagger=\frac{1}{(1+2r\sigma^2)}.$$

Inserting this value into $a^\ddagger=\alpha/2$ and (5) clearly yields

$$a^\ddagger=\frac{1}{2(1+2r\sigma^2)}$$

$$\pi_p^\ddagger+\pi_a^\ddagger=\frac{1}{4(1+2r\sigma^2)}.$$

Furthermore, inserting $\alpha^\ddagger$ and $a^\ddagger$ into $\pi_a^\ddagger$ and setting this to 0 yields equation (8) below.

$$\beta^\ddagger=\tfrac{1}{4}\alpha^2\left(-1+2r\sigma^2\right)=\frac{2r\sigma^2-1}{4(1+2r\sigma^2)^2}.$$

Proof. [Proposition 3] When the intrinsic incentives are taken into account, in the first stage, the employee chooses his action a for the given satisfaction ρ, burden δ and returns from the burden γ. Therefore, his problem is to decide an effort level a, such that π_a^* is maximized for given α, β, ρ, δ and γ: $\max_a \pi_a^*$. The optimal effort therefore is

$$a^*=\frac{\alpha+\gamma\delta+\rho}{2(1+\delta)}. \tag{A.3}$$

This implies that the employee who has developed a positive level of ρ is willing to exert a strictly positive amount of effort even if there is no monetary incentive, α. A positive level of γ will also increase the employee's effort level, if he bears some psychological burden (i.e., $\delta>0$).

The government's certainty equivalent is identical with (2). Inserting (A.3) into the joint surplus and writing it as a maximization problem with respect to α yields

$$\begin{aligned}
\max_\alpha \;&\frac{-(-2+\alpha)\alpha+(\gamma\delta+\rho)(2+\gamma\delta+\rho)}{4(1+\delta)}\\
&\qquad-\frac{2r(1+\delta)(\alpha+\rho)^2\sigma^2}{4(1+\delta)}\;.
\end{aligned}$$

The first order condition therefore is

$$\frac{1-\alpha}{2+2\delta} - r(\alpha+\rho)\sigma^2 = 0,$$

and rewriting this gives the optimal incentive rate as a function of the employee's burden and satisfaction:

$$\alpha^* = \frac{1-2r\rho\sigma^2(1+\delta)}{1+2r\sigma^2(1+\delta)}. \tag{A.4}$$

Inserting this into (A.3) yields the agent's optimal effort with

$$a^*(\rho,\delta,\gamma) = \frac{\gamma\delta + \frac{1+\rho}{1+2r(1+\delta)\sigma^2}}{2+2\delta}.$$

By setting (12) to zero and substituting a with a^*, the fixed wage β^* as a function of α can be given as:

$$\beta^* = \tfrac{1}{2}r(\alpha+\rho)^2\sigma^2 - \frac{(\alpha+\gamma\delta+\rho)^2}{4(1+\delta)}.$$

Replacing α with α^*, the optimal fixed wage $\beta^*(\rho,\delta,\gamma)$ can be calculated.

The total surplus from taking into account intrinsic incentives can be written as:

$$\pi_p^* + \pi_a^* = \tag{A.5}$$

$$\frac{(1+\gamma\delta+\rho)^2 + 2r\gamma\delta(1+\delta)(2+\gamma\delta+2\rho)\sigma^2}{4\left(1+\delta+2r(1+\delta)^2\sigma^2\right)}.$$

Sensitivity Analysis

The following statements highlight the results of sensitivity analysis for the optimal values listed in §(4.2). For simplicity of exposition, we denote

$$k = r\sigma^2.$$

Claim 1: An increase in k results in $\partial\alpha^*/\partial k < 0$, $\partial a^*/\partial k < 0$ and $\partial(\pi_a^*+\pi_p^*)/\partial k < 0$.

The proof is divided for each derivative of the optimal values. We suppose $k \geq 0$, $\rho \geq 0$, $\gamma \geq 0$ and $\delta \geq 0$.

(i) The derivative of α^* with respect to k is smaller than zero because
$$\frac{\partial\alpha^*}{\partial k} = -\frac{2(1+\delta)(1+\rho)}{(1+2k(1+\delta))^2}.$$

(ii) The derivative of a^* with respect to k is smaller than zero since
$$\frac{\partial a^*}{\partial k} = -\frac{1+\rho}{(1+2k(1+\delta))^2}.$$

(iii) The derivative of $\pi_a^*+\pi_p^*$ with respect to k is smaller than zero since
$$\frac{\partial(\pi_a^*+\pi_p^*)}{\partial k} = -\frac{(1+\rho)^2}{2(1+2k(1+\delta))^2}.$$

Claim 2: The change in ρ results in $\partial\alpha^*/\partial\rho < 0$, $\partial a^*/\partial\rho > 0$ and $\partial(\pi_a^*+\pi_p^*)/\partial\rho > 0$.

Proof. The proof is divided for each derivative of the optimal values. We suppose $k \geq 0$, $\rho \geq 0$, $\gamma \geq 0$ and $\delta \geq 0$.

(i) The derivative of α^* with respect to ρ is smaller than zero because
$$\frac{\partial\alpha^*}{\partial\rho} = -\frac{2k(1+\delta)}{1+2k(1+\delta)}.$$

(ii) The derivative of a^* with respect to ρ is bigger than zero since
$$\frac{\partial a^*}{\partial\rho} = \frac{1}{(2+2\delta)(1+2k(1+\delta))}.$$

(iii) The derivative of $\pi_a^*+\pi_p^*$ with respect to ρ is bigger than zero since
$$\frac{\partial(\pi_a^*+\pi_p^*)}{\partial\rho} = \frac{1+\gamma\delta(1+2k(1+\delta))+\rho}{2\left(1+\delta+2k(1+\delta)^2\right)}.$$

Claim 3: The change in γ results in $\partial\alpha^*/\partial\gamma = 0$, $\partial a^*/\partial\gamma > 0$ and $\partial(\pi_a^*+\pi_p^*)/\partial\gamma > 0$.

The proof is divided for each derivative of the optimal values. We suppose $k \geq 0$, $\rho \geq 0$, $\gamma \geq 0$ and $\delta \geq 0$.

(i) The derivative of α^* with respect to γ is zero because α^* does not depend on γ.

(ii) The derivative of a^* with respect to γ is bigger than zero since
$$\frac{\partial a^*}{\partial\gamma} = \frac{\delta}{2+2\delta}.$$

(iii) The derivative of $\pi_a^*+\pi_p^*$ with respect to γ is bigger than zero since
$$\frac{\partial(\pi_a^*+\pi_p^*)}{\partial\gamma} = \frac{\delta(1+\gamma\delta+\rho)}{2(1+\delta)}.$$

Claim 4: The change in δ results in $\partial\alpha^*/\partial\delta < 0$.

Suppose $k \geq 0$, $\rho \geq 0$, $\gamma \geq 0$ and $\delta \geq 0$. The derivative of α^* with respect to δ is smaller than zero because

$$\frac{\partial\alpha^*}{\partial\delta} = -\frac{2k(1+\rho)}{(1+2k(1+\delta))^2}.$$

The partial derivative signs cannot be determined unambiguously for a^* and $(\pi_a^*+\pi_p^*)$ since it depends on the sizes of α, γ and ρ. If the transferability γ is higher than the sum of incentives $\alpha+\rho$, this would induce the employee to exert higher effort level, thereby resulting in the increase in total surplus. In contrast, if γ is smaller than $\alpha+\rho$, both a^* and $\pi_a^*+\pi_p^*$ will be reduced.

Optimal α and a for different scenarios

This section compares the optimal α and a for different models with various assumptions for the parameters. A series of claims are developed.

Claim 5: If the incentive rate equals zero ($\alpha = 0$), the optimal effort level for the model with motivations and training might be higher than that in the benchmark model.

From $a^\ddagger = \frac{\alpha}{2}$ in (7) and $a^* = \frac{\alpha+\gamma\delta+\rho}{2(1+\delta)}$ in (13), we can compare $a^\ddagger$ and a^* for various scenarios. Since $\alpha = 0$, we have

(i) $\rho > 0, \delta = 0, \gamma = 0$: $a^* = \frac{\rho}{2} > 0$.

(ii) $\rho > 0, \delta > 0, \gamma = 0$: $a^* = \frac{\rho}{2(1+\delta)} > 0$.

(iii) $\rho > 0, \delta > 0, \gamma > 0$: $a^* = \frac{\gamma\delta+\rho}{2(1+\delta)} > 0$.

(iv) $\rho = 0, \delta > 0, \gamma = 0$: $a^* = 0$.

(v) $\rho = 0, \delta > 0, \gamma > 0$: $a^* = \frac{\gamma\delta}{2(1+\delta)} > 0$.

The case where $\delta = 0$ and $\gamma > 0$ is omitted since it is unrealistic.

As can be seen, as long as ρ or γ is bigger than zero, a positive effort can be exerted. However, if both ρ and γ are zero (i.e., (iii)), a^* becomes zero.

Claim 6: If α has a positive value ($\alpha > 0$), the optimal effort level for the model with motivations and trainings is higher than that in the benchmark model when a level of burden δ equals zero.

This is the case where $\alpha > 0$, $\rho > 0$, $\delta = 0$ and $\gamma = 0$. Since $a^* = \frac{\alpha+\rho}{2}$, it is clear that $a^* > a^\ddagger$.

Claim 7: If α and δ have positive values (i.e., Cases 2 to 5), the optimal effort level for the model with motivations and training can only be higher than that in the benchmark model when ρ or γ is sufficiently high.

Since the denominator of a^* in each case is bigger than that of $a^\ddagger$ (i.e., $2(1+\delta)(1+2k(1+\delta)) > 2(1+2k)$), the numerator of a^* should be sufficiently higher than that of $a^\ddagger$ to make $a^* > a^\ddagger$.

This claim implies that a moral hazard problem can be mitigated if the employee's level of emotional motivation or forward transferability on his costly effort is sufficiently high. Therefore, in our training example, even if a training program results in a high burden on the employee, it can be very effective in making the employee exert his due care as long as the employee's effort has higher forward transferability on his costly effort.

Interview Questions

Tables 7 and 7 provide the pro-forma for the questions for the two days of interviews conducted with the stakeholders from Table IV .

Table I . Interview Questions for Round 1

REGULATION

(1) Which are the important security regulations that rule the airport domain?

  (a) Are these regulations applied to every airport, irrespective of its size?
  (b) Which is the authority in charge to design these regulations?

(2) What do you think is the rationale for those security measures? Setting goals, addressing incidents, mandating technology, ecc
(3) When the regulator mandates security investments, does he mandate specific measures OR just generic measures? GENERAL REQUEST ↔ SPECIFIC REQUEST

Specific: you must have at least 3 body scanner
Generic: spend to have less than 3 successful intrusions to the tower

(4) If the regulation is violated, fines are applied? Can you give some examples? Amount? Motivation?
(5) Do authorities prefer to charge security costs on the airport overall budget OR on the passengers flight ticket? COSTS TO BUDGET ↔ COSTS TO PASSENGERS
(6) The national regulation you applied at Anadolu airport envisages a minimum OR a mandatory set of security measures? MINIMUM ↔ MANDATORY

Minimum: you have to do A or more depending on your decision
Mandatory: you have to do exactly A.

(7) How does your airport address the regulation?

  (a) Do you need (or want) to do something beyond the mandatory rules? Why?
  (b) What about other airports?

AIRPORT MANAGEMENT

(1) If you had some money to invest in security, which measure would be your first choice? And your second? Can you motivate this choice?
(2) Think about a technological recent innovation the regulator asked you to introduce: was it in line with the needs of your airport? Did it really improve the overall security?
(3) Do you think other security measures should be requested and mandated by the regulator?
(4) If the regulator increased the minimum mandatory level, would you prefer to invest more in training OR in technological devices?
(5) If you had additional money to invest for the security of your airport, would you prefer to employ a new (or updated) technological device(s) OR to introduce further training programs? TECHNOLOGY ↔ TRAINING
(6) If you had additional money to invest for the security of your airport, would you prefer to hire additional staff OR to introduce further training programs? MORE STAFF ↔ TRAINING
(7) To prevent an attack, would you prefer to improve technological countermeasures OR to (better) develop a manual contingency procedure? TECHNOLOGY ↔ MANUAL.

  (a) Would you do the same for a cyber-attack?

Note: Question sheet for semi-structured interviews. The interviews took place over the course of 14th and 15th November 2013 at the premises of the Anadolu Airport. Interviews were conducted on a one-to-one basis while a English translator attended in some cases. The interviews were recorded and transcribed. All interviewees were asked to briefly introduce themselves and specify their roles.

Table II . Interview Questions for Round 2

AIRPORT MANAGER - TRAINING	AIRPORT MANAGER - SECURITY
(1) Who is responsible for training in your airport? (2) By whom is training provided in you airport? Is a general or a specific training program? Who pays for it? (3) Do you have the chance to decide to whom commit the delivery of training? (a) If outsourcing: Why do you prefer this solution? Criteria? (Cost efficient, qualified expert personnel, better control, ...) Do you have a preferred provider? (b) If insourcing: Why do you prefer this solution? Criteria? If you could outsource, would you do that? Why? (4) The contractual relationship: (a) How can you evaluate the quality of the outsourced/insourced provided training? Monitoring? (Formal and direct monitoring// informal and infrequent? Why?) (b) Is it a long term or short term contract? (c) Do you share sensitive information with the outsourced company? (5) Have you ever experienced conflicts with the outsourced company? Explain? (6) Do you think that the training provided is enough? If you had more money, would you improve training?	(1) Who is responsible for security in your airport? (2) Can you describe the organizational structure of the security staff in your airport? Which actors are involved? Roles/duties? % decided by whom? (3) Do you have the chance to decide to whom commit the delivery of security services? (a) If outsourcing: Why do you prefer this solution? Criteria? (Cost efficient, qualified expert personnel, better control, ...) Do you have a preferred provider? (b) If insourcing: Why do you prefer this solution? Criteria? If you could outsource, would you do that? Why? (4) The contractual relationship: (a) How can you evaluate the quality of the outsourced/insourced provided training? Monitoring? (Formal and direct monitoring// informal and infrequent? Why?) (b) Is it a long term or short term contract? (c) Do you share sensitive information with the outsourced company? (5) Have you ever experienced conflicts with the outsourced company? Explain? (6) Do you have an evaluation system for police staff as well? (7) Who pays for security in your airport? (state/charges on passengers ticket/airport budget)
AIRPORT MANAGER	PRIVATE SECURITY MANAGER(S)
(1) Do you think that the current regulation related to airport security appropriately fits your airport needs? Do you think that the regulation about security measures is enough? (2) Customized vs. uniform regulation: which is more appropriate in your opinion? Why? Explain? (3) When the regulator mandates security investments, does he mandates specific measures or generic measures? (you must have 3 X-ray scanners or just you must have .. scanners?) (4) Do you need to add additional security measures beyond the mandatory rules?	(1) Which security role does your private security company cover in the airport? Duties? Activities? (Mention at least 2) (2) Do you share your everyday work activities with other security agents? Do you have different roles/duties? (How is the interplay with the other security agent managed?) (3) Do you have a specific training in aviation security? (Different training programs for different security staff? How many hours? Provided by whom?) (4) Is your performance regularly monitored? Are security agents in charge with different roles differently evaluated? How? (Are they monitored on measurable outcomes? (ex: security guards and X-ray inspector should have different performance measures)) (5) About the contractual relationship: (a) Is it a long term or short term contract? (b) Does the airport share sensitive information with you? (6) Have you ever experienced conflicts with the airport on the management of the security services? Explain.

Note: The interviews took place over the course of 27th and 28th of February 2014 at the premises of the Anadolu Airport. Interviews were conducted on a one-to-one basis while a English translator attended in some cases. The interviews were recorded and transcribed. All interviewees were asked to briefly introduce themselves and specify their roles. The first row of questions aims at collecting data about the decision of outsourcing/insourcing some services like training and security.

Public Policy And The Security of Critical Infrastructure: Discretionary or Audit Based Regulation?

Matthew Collinson and Julian Williams

Monopoly infrastructure providers, usually private companies owning and managing critical services assets, are subject to security threats that require investments in security controls to ensure an appropriately level of mitigation. Public bodies that have the duty to act as rate setters, regulating the profits of the monopolist, often have specific provisions to ensure that the monopolist makes appropriate security investments. Taking our lead from classic models of regulated industries such as those found in Laffont and Tirole (1993), we model the interaction of a social planner and a regulated monopoly when a representative attacker, generating security risks, acts strategically, in a sub-game. We show that none of the current regulatory regimes found in advanced economies dominates. However, multiple solutions do exist, including several where attackers with advanced characteristics do not engage under certain regulatory regimes, but do engage is inappropriate regulatory mechanisms are enacted.

KEY WORDS: Risk versus rules, regulation of critical infrastructure security

1. INTRODUCTION

Natural monopolies are common in areas such as public utilities where a single service provider is either geographically constrained or the provision of a market based solution is excessively costly. One facet of monopoly public services, such as bulk electricity transmission, is their criticality in respect to the proper functioning of the rest of society. Indeed, within the public policy remit, most public utilities are termed as being Critical Services (CS) and are potentially subject to malicious attack from a variety of antagonists with a broad variety of motivations.

In 2001 the American Society For Civil Engineers identified the vulnerability of Critical Services and their underlying infrastructure, often referred to as Critical National Infrastructure (CNI) to attack from terrorists and other criminals. Indeed, moving on over a decade, Presidential Policy Directive 21 (PPD-21), amongst several others, outlines a policy mandate for federal agencies to engage in pro-active monitoring and defence of designated CNI assets.

The provision of security by a natural monopolist firm provides an interesting problem for the public policy maker. When the firm is a public company owned by well diversified shareholders the risk management controls will view investment in mitigation of forward looking risks as being a risk neutral decision subject to the firms discount rate. Corporate discount rates are, on average, far higher than social discount rates; furthermore, the damage of an adverse security event to the firm, maybe far less than the damage to society overall. As such the monopolist provider of CS may substantively underinvest in mitigating forward looking security risks relative to the degree of investment that

[1]University of Aberdeen
[2]University of Durham

maybe deemed appropriate by the public policy maker acting on behalf of wider society. This paper addresses directly the mechanisms that are available to a benevolent, utilitarian social planner seeking to maximise welfare across society. We will show that the impact of strategic behaviour by attackers results in subtle but important changes to many of the classic results found in the literature on regulated industries, see Laffont and Tirole (1993) for a broad overview. Specifically, we find that rational attackers only engage in effort when their subjective cost benefit analysis satisfies at worst a zero gain. Therefore, CS providers and public policy makers will be expected to have a relative absence of attacking data with which to identify the properties of and categorize their adversaries. To this problem is added the standard problem of the policy maker needing to extract assurance from the monopolist that appropriate investment is being made and that the correct degree of public subsidy is provided to ensure that the firm has the resources to accomplish this task.

Regulatory mechanisms usually settle on a line between purely *tort* based punishments, that allocate costs of events after the fact, usually via a civil legal mechanism[3] or a *rules* or *audit* based approach that addresses compliance with a set of conditions that the social planner sets as a requirement to allow the firm to act as a monopolist. Our main analysis will look at how these regimes operate when differing attack and defence technologies are in place.

For our practical focus we will look at bulk electricity transmission as our example CS and compare the regulatory approaches taken in the UK versus the US as examples of different policy regimes. Bulk electricity transmission is the transmission side of the societally critical provision of electricity and lies between the generation and distribution component of this provision. Usually, bulk electricity transmission refers to the high-voltage lines that run across a country or between countries (called an inter-connector). This is opposed to the generation side; power stations and associated infrastructure and the distribution side; normally lower voltage connections to homes and most industrial locations.

In the US there are a large number of local bulk electricity transmission operators, as of 2014, 390 are registered with the North American Electricity Reliability Corporation (NERC) which provides regulatory oversight for the US and Canada. US regulation of bulk electricity transmission is under the auspices of the Federal Energy Regulatory Commission (FERC) and its Office of Infrastructure Security. NERC is a not-for-profit corporation that evolved from a combination of voluntary standards setters after 2001. All North American bulk electricity transmission providers are members of NERC and contribute to the standards, which are in the form of a detailed set of rules mandating specific technical aspects of transmission, reliability requirements and security provisions.[4] The transmission operators themselves are a wide variety of organisational types, from local cooperatives to large multinational corporations with distributed

By contrast in the UK a much less structured regulatory system is in operation. A single public company National Grid is licensed to operate the bulk electricity in England and Wales and owns the infrastructure, in Scotland, National Grid is the transmission operator, but not the infrastructure owner. The terms of the license specify the technical requirements for the transmission of electricity. However, the licence also specifies general requirements for security and the cost recovery mechanism.[5] The general terms require that the licensee maintain the security of the electricity infrastructure.

In Section 2 we outline the assumptions behind our main game-theoretic model, which features a regulator, an attacker and a firm; we also give specific, natural, functional forms that meet these specifications. In Section 3 we treat the model in such a way that it reduces to a two player interaction between regulator and firm, so that attacks are generated by some exogenous process; this is used to contrast with the full model, showing that it leads to potentially different conclusions and recommendations. In Section 4 we give a fuller treatment of the model in which attacks are then generated by the attacker, which can select more than one level of attacking effort.

[3] In the case of gross negligence a criminal mechanism for corporate officers, however, this is out of scope and we will look at the firm as a shareholder present value maximiser.

[4] The NERC Standards maybe found at `http://www.nerc.com/pa/Stand/Pages/default.aspx`, the security provision section is denoted ‘Critical-Infrastructure-Protection’.

[5] The National Grid Operating Licence, is found at `https://www.ofgem.gov.uk/ofgem-publications/53954/nget-rollover-special-conditions.pdf`, the security requirements section is in “Special Condition D7 parts 1 through 6”.

2. THE MODEL

We will focus on the specific interactions between a regulated monopolist, the public policy maker and a representative strategic attacker. The regulated monopolist versus a public policy maker is the subject of extensive quantitative modelling, see the classic work Laffont and Tirole (1993) for extensive insight in this area. Indeed, our approach is in the lineage of Laffont and Tirole (1986, 1993); however, we will show that when constructing a regulatory agenda for determining appropriate investments in security, the strategic nature of the attacker makes subtle, but important changes to the standard results that permeate the literature on regulation of public utilities and other natural monopolies.

2.1 A Monopolist Firm

We start with a firm whose corporate officers make decisions on a risk-neutral basis. For our purposes we will assume that the dimension of cost is based specifically on investments in security. Spillovers between security and performance investments are not considered and indeed there is little evidence to suggest that additional investment in security benefits the performance of the productive assets of the utility. We assume that the policy maker allows the firm to extract a transfer denoted $t \geq 0$; this is either in the form of a direct payment from government or by allowing the firm to extract a rent using its monopoly power. The firm then incurs two costs $\mathscr{P}(\cdot) \geq 0$ and $\mathscr{A}(\cdot) \geq 0$, which are respectively a tort based punishment action, decided ex post and an contractually pre-agreed punishment system based around deviations from a pre-agreed investment schedule monitored ex-post by audit.

Audit based compliance is presumed to be based around an audit schedule that contains various listed items which are either in compliance or default.[6] Each item is a security control that should be applied.

The bundle of investments in mitigation of security threats consists of pairs $x = (x_1, x_2)$; where x_1 are investments in security towards items on the audit schedule and x_2 are investments in security surplus to the audit schedule. We presume that the policy maker has two damage multipliers at their disposal, written as a pair $w = (w_1, w_2)$. A damage multiplier is a mechanism chosen by the policy maker and designed to set punitive costs on individuals and firms for actions that lead to injurious outcomes to individuals, other firms or society as a whole. As L. Kaplow and S. Shavell (1996) note, the need for punitive damages may be important if the expected benefit accrued to the injurer from causing injurious actions is very high and vastly outweighs the direct cost of tort renumeration.

For our purposes, the entity causing the injurious occurrences is not the monopolist utility itself, but the attacker disrupting the timely and economic supply of the utilities' good. However, current legislative interpretation accrues costs of security breaches, in part, to the firm being attacked if the firm has not taken appropriate preventative security measures to mitigate the actions of attackers. For instance, in bulk electricity transmission, we see that for an accidental loss of power from weather related incidents, transmission service operators (TSOs) have been subject to punitive damages by courts with state legislatures acting as plaintiffs. Whilst we are yet to see a tort action on the basis of a security event for public utilities, several retail and medical organisations have been subject to punitive damages for loss of data caused by security incidents where attackers have gained access to records Westby (2004) [7] Epstein and Brown (2008) [8].

We assume that the efficiency of security investment is described by a family of hyperbolic functions with constant degree of absolute risk reduction to investment; therefore the firms investments have the following properties

$$-\theta_1 = \frac{\mathscr{P}''_{x_1}}{\mathscr{P}'_{x_1}}, \quad -\theta_2 = \frac{\mathscr{P}''_{x_2}}{\mathscr{P}'_{x_2}}$$

where $\mathscr{P}'_{x\in\{x_1,x_2\}}$ and $\mathscr{P}''_{x\in\{x_1,x_2\}}$ are, respectively, the first and second derivative with respect to $x \in \{x_1, x_2\}$, we will also assume that investments in the audited component of investments has the same properties therefore

$$-k = \frac{\mathscr{A}''_{x_1}}{\mathscr{A}'_{x_1}},$$

[6] In the NERC audit schedule each top level item on the audit list is denoted as a requirement. In our model, as in reality, the compliance with these audit requirements is not presumed to be total, but a function of the investment choice of the firm.

[7] Chapter 1, Section C, page 26

[8] Specifically noted is the Class Action Complaint, In re TJX Companies Retail Security Breach Litigation, No 07-10162 3 (D Mass filed Apr 25,2007)

where we can think of k as the firms efficiency in complying with the audit based compliance, we assume that the limit $\lim_{k\to\infty} \mathscr{A} = 0$, i.e. that firms that are highly efficient at complying with audit based regulatory systems suffer a negligible compliance penalty.

We can think of θ_1 and θ_2 as representing the efficiency of the firm in converting fixed investment x_1 and x_2, respectively, into risk reduction, on the presumption that $\mathscr{P}^{'}_{x_i} < 0$ for all $x_i > 0$ and $i \in \{1,2\}$, and that marginal decreases in the probability of attack to investment are diminishing, $\mathscr{P}^{''}_{x_i} > 0$ for all $x_i > 0$ and $i \in \{1,2\}$. We assume that $\theta_1 < \theta_2$, so that on aggregate, the firm knows better how to reduce risk more efficiently than the policy-maker; moreover, if we do not assume this then rules based regulation becomes a dominant strategy for the policy-maker.

We will break our analysis down into two broad aspects, first when attacks are exogenous, and second when attackers are endogenous and play in sub-games with a target CNI firms. We can then illustrate the importance of attackers as endogenous strategic players in changing many of the results for regulating a firm of the type appearing in Laffont and Tirole (1993). The decision making pattern for attackers is less well understood, however, for tractability we will assume that the attacker has a broadly similar efficiency set-up to the firm, whereby

$$-\beta = \frac{\mathscr{P}^{''}_{z}}{\mathscr{P}^{'}_{z}} \quad \text{and} \quad \mathscr{P}^{'}_{z} > 0.$$

where z is a decision variable representing the level of effort variable for the attacker. The numeraire for z is assumed to be a separate from that of x, therefore attackers and defenders are specifically assumed to have different units of account. This captures the notion of non-monetary rewards for attackers.

2.2 The Public Subsidy

The firm is assumed to be a regulated monopoly that receives a transfer t from the social planner, either directly or indirectly in the form of a rent charged to society for the production of the monopoly good. In keeping with the standard literature, $t \in \mathbb{R}$ is unbounded and as such the firm can receive a negative subsidy. For a risk neutral firm we can write the expected pay-off in a single period as follows:

$$\mathscr{U}_F = t - \mathscr{P} - \mathscr{A} - x_1 - x_2$$

for tractability of exposition we assume that $\mathscr{P}$ and $\mathscr{A}$ are composites of a probability of a single successful event and the cost to the firm of this event. We further assume that the probabilistic component is a product of the attacker actions and the firm actions. Therefore we set $\mathscr{P} = \mathscr{D}_F \times \mathscr{P}_A \times \mathscr{P}_F$. The damage to the firm from a successful attack $\mathscr{D}_F \geq 0$ is the sum of two components, D the actual damage to the firm and w_1, the policy makers penalty against the firm for allowing a successful attack; $\mathscr{P}_A \in [0,1]$ is the probability of a successful attack when a firm chooses to not defend itself; we assume therefore that $\partial \mathscr{P}_A / \partial x_{i\in\{1,2\}} = 0$.

The result of the firms choices are encapsulated in the probability $\mathscr{P}_F \in [0,1]$ is the proportional reduction in risk, for a given $\mathscr{P}_A$ for investment, $x \in \{x_1, x_2\}$, by the target in defending itself from attackers, as the change in likelihood for a given attacking effort is assumed to be completely captured by $\mathscr{P}_A$, we assume that $\partial \mathscr{P}_F / \partial z = 0$. The third term in the pay-off is the audit penalty for non-compliance with the social planners audit based regulatory mechanism $\mathscr{A} = w_2 \times \mathscr{A}_F$, total non-compliance, $x_1 = 0$, with the planners regulatory mechanism results in a fine w_2, and complete compliance $\mathscr{A}_F = 0$, results in no compliance penalty to be levied against the firm.

The social planner's pay-off partially mirror's the firms, except that damage to society from a successful attack is assumed to be different from the damage to the firm, therefore $\mathscr{D}_P \neq \mathscr{D}_F$. To ensure an appropriate conversion to the policy makers unit of account from the firm, we have a collection of multipliers $\lambda = \{\lambda_0, \lambda_1, \lambda_2, \lambda_3, \lambda_4\}$. Therefore the social planners pay-off is therefore described as follows:

$$\mathscr{U}_P = -\lambda_0 t - \mathscr{H} - \mathscr{C} + \lambda_1 x_2 + \lambda_2 x_2 + \lambda_3 \mathscr{P} + \lambda_4 \mathscr{A}$$

where $\mathscr{H}$ is the expected social damage from attacks and $\mathscr{C}$ is the benefit a policy maker gains from the regulated monopolist firm's compliance with the social planner's audit based regulatory mechanism. We assume that $\mathscr{H}$ is composed of the probability of a single successful attack and the social damage multiplier μ, as such $\mathscr{H} = \mu \times \mathscr{P}_A \times \mathscr{P}_F$. Furthermore, we assume that $\mathscr{C} = \nu \times \mathscr{A}_P$, and the policy makers assurance is hyperbolic in x_2, such that $\mathscr{C}^{''}{}_{x_1} / \mathscr{C}^{'}{}_{x_1} = -K$.

2.3 The Attacker's Objective

For our main specification, we will assume that attackers act strategically, with a well described set of preferences. This is, of course, very difficult to verify as the literature of the psychology and economic incentives of attackers is extremely limited. However, as an initial treatment we can assume that attackers will have finite, non-zero costs for their activities and that once we have an understanding of the numeraire for rewards versus the cost of attack, we can write out an objective function. Let $\mathscr{R}$ be the expected reward from attacks. We will assume that this is given by the product $r \times \mathscr{P}_A \times \mathscr{P}_F$, therefore in our one period set-up the risk-neutral attacker's pay-off function is:

$$\mathscr{U}_A = -z + \mathscr{R},$$

where $z \geq 0$ is the attacker's level of effort.

For the initial part of our analysis we will fix the attacker's effort (i.e. exogenous attacking effort) as a benchmark model and treat the social planner and monopoly firm interaction as Stackelberg game This can be done formally by restricting the set of available attacker actions to a singleton set. However, before we begin our treatment of this (exogenous attack) set-up we will quantify the model using functional forms that explicitly capture our preceding assumption of an endogenous attacker with a non-trivial choice of effort level.

2.4 Quantifying The Model

At this juncture the analysis is served by placing more explicit functional forms on $\mathscr{P}_{i \in A,F}$, $A_{i \in F,P}$ and r. A natural functional form that satisfied the assumptions implicit in the preceding set-up is to use exponential cumulative distributions to describe the contributions to the likelihood of success and the degree of compliance with the policy makers audit based regulatory mechanism. Therefore, let $\mathscr{P}_F = \exp(-\theta_1 x_1 - \theta_2 x_2)$, $\mathscr{P}_A = 1 - \exp(-\beta z)$, $\mathscr{A}_F = \exp(-k x_1)$ and $\mathscr{A}_P = \exp(-\kappa x_1)$.

To simplify the analysis, we consider only cases with $\lambda_0 = \lambda_1 = \lambda_2 = 1$. Our treatment will thus not capture all of the subtle informational effects in other transfer and regulation models Laffont and Tirole (1986, 1993).

We also suppose that transfers back to the policy-maker from the firm as the result of penalties are essentially negligible from its viewpoint: thus we take $\lambda_3 = \lambda_4 = 0$.

The payoffs therfore take the following forms:

$$\begin{aligned}
\mathscr{U}_A &= -z + r(1 - e^{-\beta z})e^{-(\theta_1 x_1 + \theta_2 x_2)} \\
\mathscr{U}_P &= -t - \mu(1 - e^{-\beta z})e^{-(\theta_1 x_1 + \theta_2 x_2)} - \nu e^{-\kappa x_1} \\
\mathscr{U}_F &= t - x_1 - x_2 - w_1 e^{-k x_1} \\
&\quad -(w_2 + D)(1 - e^{-\beta z})e^{-(\theta_1 x_1 + \theta_2 x_2)}.
\end{aligned}$$

We further assume that $\beta r > 1$. With the quantification above, this turns out to b necessary in order to have the attacker attack even a firm that makes no investment in security, that is, with $x_1 + x_2 = 0$.

3. A TREATMENT WITH EXOGENOUS ATTACKS

3.1 A Stackelberg Game

The principal goal of this model is to show the effects of different policy regimes upon the behaviour of the firm, and to be contrasted with the main model of Section 4. We reduce the model to a Stackelberg game, with the firm following the policy-maker.

As noted above, we restrict the set of actions of the attacker to a singleton $\{\zeta\}$, giving a fixed attacker action $z = \zeta$. Let $m = 1 - e^{-\beta\zeta}$ and $m = \mu M$. The payoffs fo the policy-maker and firm reduce to the forms:

$$\begin{aligned}
\mathscr{U}_P &= -t - M e^{-\theta_1 x_1 - \theta_2 x_2} - \nu e^{-K x_1} \\
\mathscr{U}_F &= t - (w_2 + D) m e^{-\theta_1 x_1 - \theta_2 x_2} - w_1 e^{-k x_1} \\
&\quad -x_1 - x_2.
\end{aligned}$$

The firm has control of x_1 and x_2 only. The policy-maker controls t, w_1, w_2 and k.

The function $\mathscr{U}_F$ is twice differentiable and concave $\frac{\partial^2 \mathscr{U}_F}{\partial {x_1}^2} < 0$, $\frac{\partial^2 \mathscr{U}_F}{\partial {x_2}^2} < 0$, $\frac{\partial^2 \mathscr{U}_F}{\partial x_1 \partial x_2} < 0$. Moreover $\frac{\partial \mathscr{U}_F}{\partial x_i} \longrightarrow -1$ as $x_i \longrightarrow +\infty$.

3.2 Analysis

We assume initially that the firm has complete information about itself and the policy-maker, and perfect information about the game state after the policy-maker has acted.

The Stackelberg game gives rise (as usual) to a strategic form game in which the firm's strategies are functions from the set of states reached through the policy-maker's choice to the set of the firm's actions. We confine our analysis to finding the function given by the firm's optimal choice of action given each of the policy-maker's actions.

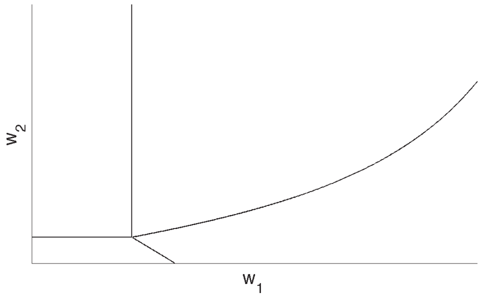


Fig. 1: $m\theta_2 D < 1$

Other strategies, for example constant choices of action that do not depend upon the policy-maker's choice, may also be of interest, especially when repeated play is considered.

The firm's choices of action x_1 and x_2 both have a lower bound (of zero). The concavity of $\mathscr{U}_F$ then means that in any given game state in which the firm can act, it always has an optimal action, but this may involve a choice at one or both of the lower bounds.

Although k is an aspect of the punishment that is under control of the policy-maker, it is convenient (for now) to force it to take a fixed constant value. The residual controls of the policy-maker w_1, w_2, t can then be viewed as points (w_1, w_2) in a space, together with a constraint on feasible pairs (w_1, w_2) that are feasible under the given subsidy t

The optimal reactions of the firm to a given pairs of punishment regimes (w_1, w_2) at a a given transfer, partition the space of into simply-connected regions according to whether the firm's actions lie at their lower bounds. This partiton can take two different (but related) forms, depending upon whether $m\theta_2 D$ is less, great or equal to 1. Ignoring, the special case of equality, the two main forms are shown in Figures 1 and 2.

Figure 1 consists of one horizontal line, one vertical line, one other straight line, and one curve. All four of these lines meet at a single point. The horizontal straight line, $L^{(m)}$ lies at $w_2 = w_2^{(m)} \equiv -D + \frac{1}{m\theta_2}$. The vertical straight line, $L^{(S)}$ lies at $w_1 = w_1^{(S)} \equiv \frac{1}{k}\left(1 - \frac{\theta_1}{\theta_2}\right)$. Let $L^{(D)}$ be the diagonal straight line: it it determined by the equation

$$kw_1 + m\theta_1(w_2 + D) = 1.$$

Let C be the curve: it is defined by $w_2 = -D + \frac{1}{m\theta_2}\left(\frac{w_1}{w_1^{(S)}}\right)^{\frac{\theta_1}{k}}$.

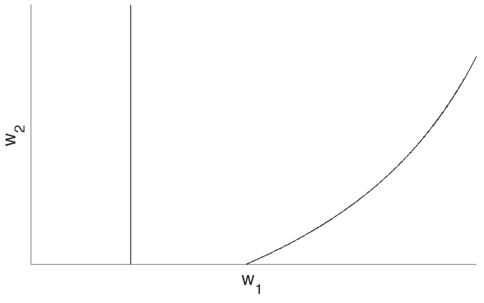


Fig. 2: $m\theta_2 D > 1$

Let the optimal pair of action choices for the firm be (x_1^*, x_2^*) given the policy-maker's choice (w_1, w_2). In both figures, any point on or to the left of $L^{(S)}$ has $x_1^* = 0$, any strictly point above C has $x_1^* > 0$ and $x_2^* > 0$, and any point strictly to the right of C or $L^{(D)}$ has $x_1^* > 0$ and $x_2^* = 0$. in Figure 1 those points below the lines $L^{(m)}$ have $x_2^* = 0$, as do those points below $L^{(D)}$ and with $w_1 > w_1^{(S)}$. Thus in both figures there are regions of pure rules compliance, of pure tort mitigation, and of mixed rules compliance and tort mitigation. In Figure 1 there is an additional region in which there is neither rules compliance nor tort mitigation. Figure 2, is essentially Figure 1 but with $w_2^{(m)} < 0$.

The difference implied by these two figures is that, under a pure rules-based regulatory regime, a firm will engage in a mixture of rules compliance and tort mitigation (i.e. additional 'risk-based' security activity) only if it is subject to sufficiently severe damage arising from attacks.

A second point to note is that the curve C has no asymptote. Thus, ignoring any feasibility constraint imposed by the level of subsidy, under an arbitrarily strict rules punishment,w_1, it is possible for the policy-maker to induce non-zero tort mitigation investment, $x_2^{**} > 0$, by imposing a sufficiently high tort punishment, w_2. This stands in contrast to our main model below.

4. A TREATMENT WITH AN ENDOGENOUS ATTACKER

We return now to the full three-player situation with the payoffs for the players introduced in Section 2.4. Similar to Section 3 we again find regions defined by the reactions to policy set.

Let $w_2^{(S)} = \frac{\beta r}{\beta r - 1}\frac{1}{\theta_2} - D$ and $w_1^{(D)} = \frac{1}{k}(\beta r)^{\frac{k}{\theta_1}}$ and

$w_1^{(A)} = \frac{1}{k}(1 - \frac{\theta_1}{\theta_2})(\beta r)^{k/\theta_1}$. Note that $w_1^{(S)} < w_1^{(A)} < w_1^{(D)}$. Let $x_1^{**}, x_2^{**}, z^{**}$ be the Nash Equilibrium values for x_1, x_2, z respectively in the subgame played with given parameters and policy.

Proposition 1: If $w_1 \leq w_1^{(S)}$ and $w_2 \leq w_2^{(S)}$, or if $w_1 \geq w_1^{(S)}$ and $w_1 k + (w_2 + D)\theta_1 \frac{\beta r - 1}{\beta r} \leq 1$, then $x_1^{**} = x_2^{**} = 0$ and $z^{**} = z^{(1)} \equiv \frac{1}{\beta}\log(\beta r) > 0$.

Thus there is a region in which the firm invests nothing in security in response to policy, and the attacker invests an amount that is determined by its own attacking efficiency β and the reward r.

Proposition 2: If $w_1 \leq w_1^{(S)}$ and $w_2 > w_2^{(S)}$ then $x_1^{**} = 0$ and $x_2^{**} = x_2^{(2)} \equiv \frac{1}{\theta_2}\log\left(\frac{(w_2+D)\theta_2\beta r}{\beta r + (w_2+D)\theta_2}\right) > 0$ and $z^{**} = z^{(2)} \equiv \frac{1}{\beta}\log\left(\frac{\beta r + (w_2+D)\theta_2}{(w_2+D)\theta_2}\right) > 0$.

Thus there is a region in which there is only tort mitigation investment, and no rules compliance investment. The attacker's effort reduces, but does not diminish completely to zero as the tort punishment becomes more severe. In effect, the firm and attacker reach an equilibrium which is acceptable to both in terms of effort and punishment or reward, and which is only at the lower boundary for the rules compliance investment.

Proposition 3: If $w_1^{(S)} < w_1 < w_1^{(A)}$ and $w_2 \geq -D + \frac{1}{\theta_2\left(\left(\frac{w_1^{(S)}}{w_1}\right)^{\theta_1/k} - \frac{1}{\beta r}\right)}$, then $x_1^{**} = \frac{1}{k}(\log(\frac{w_1}{w_1^{(S)}})$ and $x_2^{**} = \frac{1}{\theta_2}\log\left(\frac{\beta r\theta_2(w_2+D)}{\beta r + \theta_2(w_2+D)}\left(\frac{w_1^{(S)}}{w_1}\right)^{\theta_1/k}\right)$ and $z^{**} = \frac{1}{\beta}\log\left(1 + \frac{\beta r}{(w_2+D)\theta_2}\right)$.

Thus there is a region which the firm invests in both rules compliance and tort mitigation. The level of attacking effort, z^{**}, and the mitigation of attacks, captured by $\exp(-\theta_1 x_1^{**} - \theta_2 x_2^{**})$, are dependent only upon the severity of the tort punishment, w_2, and not the rules punishment w_1: an increase in w_1 is matched by a compensatory decrease in w_2. In essence, the firm chooses to invest less in tort-based mitigation, because rules compliance is already providing some mitigation. Thus, within this zone, the only advantage to the policy-maker in increasing w_1 comes from the increased assurance that rules compliance provides — such an increase also cannot be achieved at a lower transfer.

Proposition 4: If $w_1^{(A)} < w_1 < w_1^{(D)}$, or if $w_1^{(S)} < w_1 < w_1^{(A)}$ and $w_2 \leq -D + \frac{1}{\theta_2\left(\left(\frac{w_1^{(S)}}{w_1}\right)^{\theta_1/k} - \frac{1}{\beta r}\right)}$ and $w_1 k + (w_2 + D)\theta_1\frac{\beta r - 1}{\beta r} > 1$, then $x_2^{**} = 0$, and $x_1^{**} > 0$ is the unique solution to the equation

$$w_1 k e^{-kx_1} + (w_2 + D)\theta_1(e^{-\theta_1 x_1} - \frac{1}{\beta r}) = 1 \qquad (1)$$

and $z^{**} > 0$.

Thus there is a region in which the firm invests only in rules compliance and not in tort mitigation, and on which the attacker invests a non-zero attacking effort. Analytically, this case is a little more complicated than the other cases, with the compliance investment being determined by an implict function of w_1 and w_2, and the attacker's effort being determined from that.

Proposition 5: If $w_1 \geq w_1^{(D)}$ then $x_1^{**} = x_1^{(6)} \equiv \frac{1}{k}\log(w_1 k) > 0$ and $x_2^{**} = 0 = z^{**}$.

Thus there are conditions in which the attacker's effort drops to zero. This happens when the rules-based punishment regime is so strict that the firm always invests heavily in rules compliance and this is sufficiently effective at mitigating attacks. This feature was not present in the model with exogenous attacks above. However, as further discussed below, although this may be an ideal situation from a security perspective, this may require a large subsidy in order to make such policies feasible for the firm.

The above results can be summarised geometrically in the space of points (w_1, w_2). Let C be the curve defined by

$$w_2 = -D + \frac{1}{\theta_2\left(\left(\frac{w_1^{(S)}}{w_1}\right)^{\theta_1/k} - \frac{1}{\beta r}\right)}. \qquad (2)$$

We are only interested in C for $w_1 \geq w_1^{(S)}$ and to the left of its asymptote at $w_1^{(A)}$. Let L be the line defined by

$$w_1 k + (w_2 + D)\theta_1\frac{\beta r - 1}{\beta r} = 1. \qquad (3)$$

The curve C meets the line L where $w_1 = w_1^{(S)}$ and $w_2 = w_2^{(S)}$.

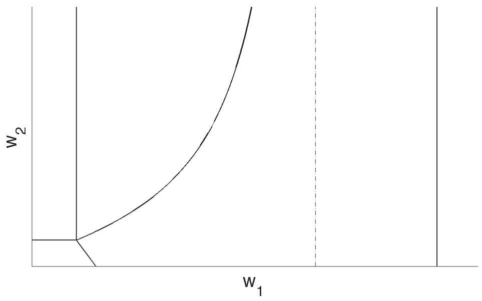


Fig. 3: $D\theta_2(\beta r - 1) < \beta r$

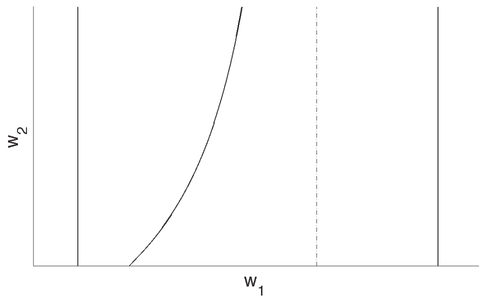


Fig. 4: $D\theta_2(\beta r - 1) > \beta r$

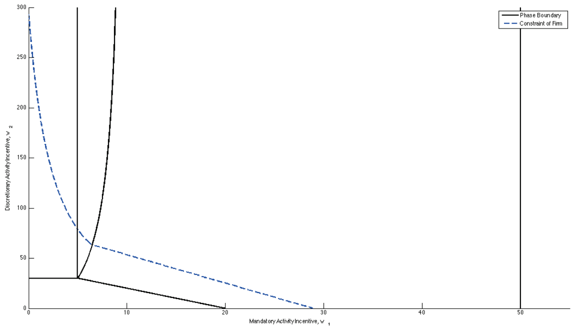


Fig. 5: Reaction Regions with Firm's Constraint

The summary then takes the form of the pair of diagrams given in Figure **??**. There are two diagrams depending upon whether $w_2^{(S)}$ is greater than or less than zero. There are now five regions sepearated by the solid lines shown. The asymptote to C is also indicated. The region in which the attacker's effort drops to zero is the rightmost.

Much of the structure of Figures 1 and 2 is embedded in the Figure 3 and 4, respectively. The parameter m that determines both $w_2^{(m)}$ and $L^{(m)}$ is replaced by $(\beta r - 1)/(\beta r)$ in determining $w_2^{(S)}$ and L.

However, there is now the additional possibility that attacks can be driven completely to zero by choosing a sufficiently severe rules punishment, w_1: this may or may not be feasible at the given level of transfer.

The form of the curved boundary, C, between the mixed investment (by the firm) and pure compliance regions is also different. With endogenous attacker, it has an asymptote. Thus, for a sufficiently high severe rules punishment, $w_1 \geq w_1^{(A)}$, there can be no tort punishment, w_2, that induces a mixed investment response by the firm. In this case the policy maker has forced all of the firm's attention into rules compliance.

The firm cannot be expected to participate in a contract which stipulates punishments that are not sustainable at the given level of transfer. For concreteness we suppose that the firm must achieve a payoff of at least 0. The condition $\mathscr{U}_F \geq 0$ is the firm's budget/ participation/individual rationality constraint. At a given transfer t, this condition defines a subspace of the set of policies (w_1, w_2). An example, is the set of point to the right of the curve with negative slope in Figure 5.

5. VISUALIZATION

It is useful to visualize the shapes defined by the various quantities output by the model. The equilibiria values x_1^*, x_2^*, z^* with given constants allow for the payoffs to be plotted at those equilibria. This allows us to visualize the reaction to policy (w_1, w_2, t), in particular by plotting surfaces over a plane consisting of points (w_1, w_2). The qualitative reaction regions (phases) are seen to capture essential information The following plots have been generated using MATLAB MATLAB (2014) using the values $k = 0.04, D = 10, r = 20, \beta = 0.1, \kappa = 0.02, \mu = 100.0, \theta_1 = 0.04, \theta_2 = 0.05, t = 1.57 * 1/\theta_2$. This happens to be a case in which there is a region with total security investment $x_1^* + x_2^* = 0$. The particular vertical scales in the following plots are not significant.

Figure 5 shows the reaction regions with the firm's constraint plotted in blue. In particular, it is not feasible to force the attacker to reduce its effort to zero at the present level of subsidy.

Figure 6 shows the typical form of the attacker's attacking effort. Figure 7 shows the typical form of the attacker's payoff. The attacker's effort is highest

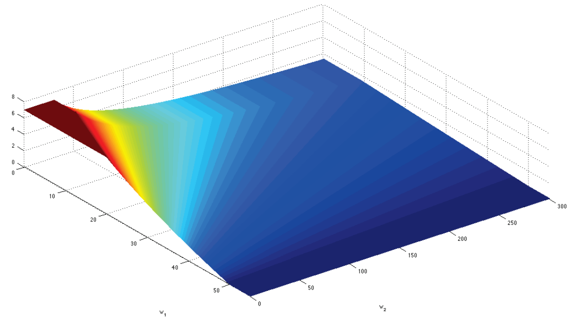


Fig. 6: Attacker Effort, z^*

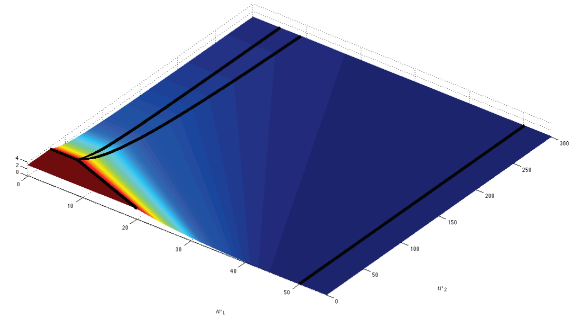


Fig. 7: Attacker Payoff, $\mathscr{U}_A$

when the firm does not defend itself. As w_1 increases to a sufficiently high value, the attacker ceases to attack entirely and its payoff in minimized.

Figure 8 shows the typical form of the firm's rules compliance investment. Figure 9 shows the typical form of the firm's tort mitigation investment. It should be noted that where rules investment is high, tort investment is low, illiustrating the substitution effect between these components. The rules investment is zero for $w_1 < w_1^{(S)}$ and the tort mitigation investment is driven to zero to the right of the curved mixed investment boundary.

Figure 10 shows the typical form of the firm's payoff. This illustrates that security investment resulting from policy is costly.

Figure 11 shows the typical form of the policy-maker's payoff with no assurance from rules. Figure 12 shows the typical form of the policy-maker's payoff with assurance from rules compliance. The firm's constraint has been projected onto the surface in Figure 11; *in this particular case* we can see immediately by inspection of the intercept of the constraint and the axes that the optimal, feasible,

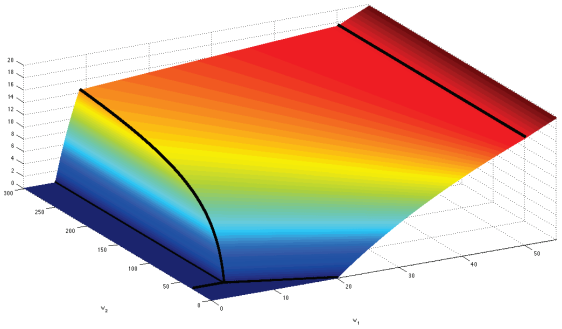


Fig. 8: Rules Compliance Investment, x_1^*

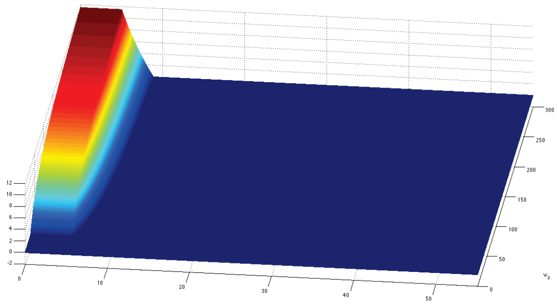


Fig. 9: Tort Mitigation Investment, x_2^*

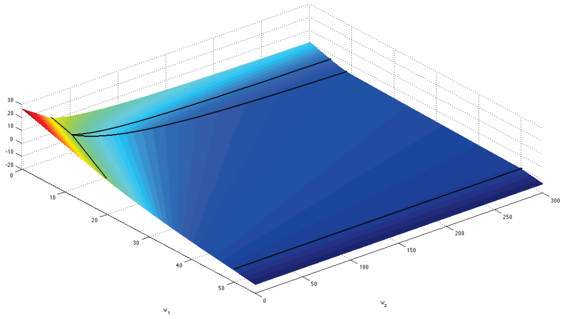


Fig. 10: Firm's Payoff, $\mathscr{U}_F$

pure tort policy, $w_1 = 0$, yields a higher payoff than the optimal, feasible, pure rules policy, $w_2 = 0$.

6. SUMMARY

In recent years and in several countries, cybersecurity has been incorporated into the responsibility of private sector bulk electricity transmission operators. This has sometimes been done by amending legislation or structuring contracts so that a require-

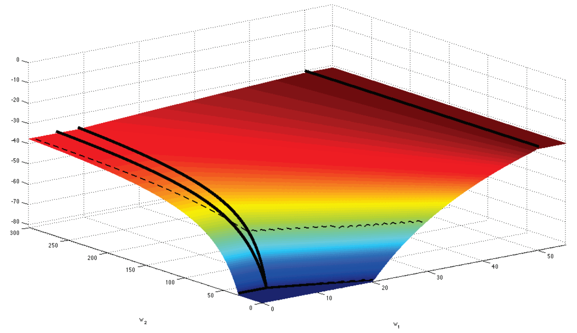


Fig. 11: No Assurance From Compliance, $\nu = 0.0$

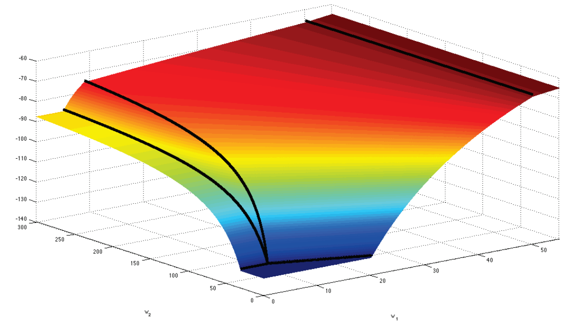


Fig. 12: With Assurance From Compliance, $\nu = 50.0$

ment to provide a reliable service, which is of course essential for this critical service, includes the need to put in place adequate security controls and processes. In most traditional engineering, reliability is such that adverse events are generated stochastically. This view can be integrated simply into economic models featuring two players, regulator and firm, for example Laffont and Tirole (1993). However, security engineering is different. Adverse security events are generated by attackers with agency that actively adjust their attacking effort in anticipation of and reaction to defensive effort. This paper has demonstrated that economic models with exogenous attacks will lead to different policy recommendations than models with adversarial attackers.

Moreover, which is better, rules-based regulation or tort-based regulation, is a complex question that depends on several environmental factors and the preferences of the the policy-maker.

Derivatives of the Firm's Payoff

Suppose a choice (x_1, x_2) by F. The propotion of vulnerabilites that are not mitigated is $e^{-(\theta_1 x_1 + \theta_2 x_2)}$. The proportion of all of the vulnerabilites that are unmitigated and exploited is therefore

$$\text{Exploited}\,(x_1, x_2, z) = (1 - e^{-\beta z})e^{-(\theta_1 x_1 + \theta_2 x_2)}. \tag{A.1}$$

First and second derivatives:

$$\frac{\partial \mathscr{U}_F}{\partial x_1} = -1 + w_1 k e^{-k x_1} \tag{A.2}$$

$$+ (w_2 + D)\theta_1 \,\text{Exploited}\,(x_1, x_2, z) \tag{A.3}$$

$$\frac{\partial^2 \mathscr{U}_F}{\partial {x_1}^2} = -w_1 k^2 e^{-k x_1} \tag{A.4}$$

$$- (w_2 + D)\theta_1^2 \,\text{Exploited}\,(x_1, x_2, z) \tag{A.5}$$

$$\frac{\partial \mathscr{U}_F}{\partial x_2} = -1 + (w_2 + D)\theta_2 \,\text{Exploited}\,(x_1, x_2, z) \tag{A.6}$$

$$\frac{\partial^2 \mathscr{U}_F}{\partial {x_2}^2} = -(w_2 + D)\theta_2^2 \,\text{Exploited}\,(x_1, x_2, z) \tag{A.7}$$

$$\frac{\partial^2 \mathscr{U}_F}{\partial x_1 \partial x_2} = -(w_2 + D)\theta_1\theta_2 \,\text{Exploited}\,(x_1, x_2, z). \tag{A.8}$$

Note that the firm's maximization problem is not convex if $z = 0$.

Attacker Reaction

It is easiest to begin the analysis by calculating the attacker's optimal reaction, given the policy choice $p = (w_1, w_2)$ and the firms's choice (x_1, x_2). Let the attacker's reaction be z^*.

The attacker has a constrained maximization problem over $z \geq 0$ There are two cases: boundary reaction $z^* = 0$, or internal reaction $z^* > 0$.

Now

$$\frac{\partial \mathscr{U}_A}{\partial z} = -1 + \beta r e^{-\beta z} e^{-(\theta_1 x_1 + \theta_2 x_2)} \tag{A.9}$$

$$\frac{\partial^2 \mathscr{U}_A}{\partial z^2} = -\beta^2 r e^{-\beta z} e^{-(\theta_1 x_1 + \theta_2 x_2)}. \tag{A.10}$$

Evidently $\frac{\partial^2 \mathscr{U}_A}{\partial z^2} < 0$ for all x_1, x_2, z, so the attacker's maximization problem is always convex. Moreover, $\frac{\partial \mathscr{U}_A}{\partial z} \longrightarrow 0$ as $z \longrightarrow +\infty$.

Thus, there is a boundary reaction iff

$$\beta r \leq e^{\theta_1 x_1 + \theta_2 x_2}. \tag{A.11}$$

If the reaction is internal then it occurs at

$$z^* = \frac{1}{\beta}(\log(\beta r) - \theta_1 x_1 - \theta_2 x_2). \tag{A.12}$$

In summary,

$$z^* = \max\{0, \frac{1}{\beta}(\log(\beta r) - \theta_1 x_1 - \theta_2 x_2)\}. \tag{A.13}$$

The more that the firm invests in defence, the less the attacker attacks. Note that the firm can cause attack to cease entirely by investing sufficiently. However, the firm's behaviour is conditioned by its interaction with both the attacker and the policy-maker.

A Menagerie of Important Values

Important values of (w_1, w_2).

The following values of w_1 are of significance:

$$\begin{aligned}
w_1^{(S)} &= \frac{1}{k}(1 - \frac{\theta_1}{\theta_2}) \\
w_1^{(1)} &= \frac{1}{k}\frac{\beta r}{(\beta r - 1)} \\
w_1^{(3)} &= \frac{1}{k}(1 - D\theta_1) \\
w_1^{(4)} &= \frac{1}{k}\frac{\beta r}{\beta r - 1}(1 - \frac{\theta_1}{\theta_2}) \\
w_1^{(5)} &= \frac{1}{k}\frac{\beta r}{\beta r - 1} \\
w_1^{(6)} &= \frac{1}{k}(1 - D\theta_1\frac{\beta r - 1}{\beta r}) \\
w_1^{(7)} &= \frac{1}{k}(\beta r)^{k/\theta_1} \\
w_1^{(8)} &= \frac{1}{k}(1 - \frac{\theta_1}{\theta_2})(\beta r)^{k/\theta_1} \\
w_1^{(9)} &= w_1^{(S)}\left(\frac{e^{\theta_2 t - 1 - \theta_2 w_1^{(S)}}}{\beta r}\right)^{(\theta_1+\theta_2)/k} \\
w_1^{(10)} &= w_1^{(S)}\exp\left(\frac{(\theta_2 t - 1 - \theta_2 w_1^{(S)})k}{2\theta_1 + \theta_2}\right) \\
w_1^{(11)} &= t - \frac{1}{\theta_2} - \frac{1}{\theta_2}\log(\beta r) \\
w_1^{(12)} &= \frac{1}{k}\exp(kT - 1) \\
w_1^{(13)} &= w_1^{(S)}\exp(k\alpha)
\end{aligned}$$

Note that

$$0 < w_1^{(S)} < w_1^{(3)} < \frac{1}{k} < w_1^{(5)}$$

and

$$w_1^{(S)} < w_1^{(4)} < w_1^{(5)}$$

and

$$w_1^{(8)} < w_1^{(7)}.$$

The following values of w_2 are significant:

$$\begin{aligned}
w_2^{\circ} &= \tfrac{1}{(1-e^{-\beta z})}\tfrac{1}{\theta_2} - D \\
w_2^{(1)} &= \tfrac{\beta r}{(\beta r-1)}\tfrac{1}{\theta_1} - D. \\
w_2^{(2)} &= \tfrac{\beta r}{(\beta r-1)}\tfrac{1}{\theta_2} - D. \\
w_2^{(3)} &= \tfrac{1}{\theta_2} - D \\
w_2^{(4)} &= \beta r\tfrac{1}{\theta_2} - D \\
w_2^{(5)} &= \tfrac{1}{\theta_2} - D \\
w_2^{(6)} &= \tfrac{1}{\theta_1} - D \\
w_2^{(7)} &= \tfrac{\mu}{1-\theta_1} - D \\
w_2^{(9)} &= -D + \tfrac{\beta r}{\theta_2\left(\beta r\exp(-\theta_2 t+1+\theta_2 w_1^{(S)})-1\right)}
\end{aligned}$$

Geometry of the Interaction

Let L_0 be the vertical line in the (w_1, w_2)-plane at $w_1 = w_1^{(S)}$.

Let L_4 be the vertical line in the (w_1, w_2)-plane at $w_1 = w_1^{(8)}$.

The line L_3 defined by

$$w_1 k + (w_2 + D)\theta_1\frac{\beta r - 1}{\beta r} = 1. \tag{A.14}$$

It intercepts the w_1-axis at $w_1^{(6)}$ and the w_2-axis at $w_2^{(1)}$. It meets the line L_0 where $w_2 = w_2^{(2)}$.

We have L_3 in the upper-right quadrant if

$$\frac{\beta r}{\beta r - 1}\frac{1}{\theta_2} < D < \frac{1}{\theta_1}(1 - \frac{1}{\beta r}). \tag{A.15}$$

Let C_5 be the curve defined by

$$\frac{\beta r\theta_2(w_2 + D)}{\beta r + \theta_2(w_2 + D)} = \left(\frac{w_1}{w_1^{(S)}}\right)^{\theta_1/k} \tag{A.16}$$

This can be rewritten in the form:

$$w_2 = -D + \frac{1}{\theta_2\left(\left(\frac{w_1^{(S)}}{w_1}\right)^{\theta_1/k} - \frac{1}{\beta r}\right)} \tag{A.17}$$

This curve meets the line L_0 at the point $(w_1^{(S)}, w_2^{(2)})$. So C_5 intersects L_0 in the upper-right quadrant iff $w_2^{(2)} \geq 0$.

C_5 has an asymptotes at $w_1 = w_1^{(8)}$: on the curve

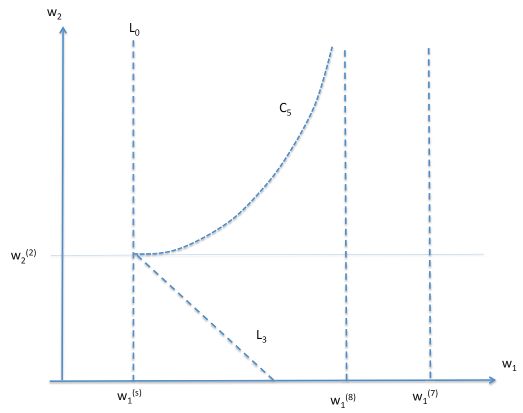


Fig. 1: Geometry of the Interaction

$w_2 \longrightarrow +\infty$ as w_1 tends to $w_1^{(8)}$ from below; $w_2 \longrightarrow -\infty$ as w_1 tends to $w_1^{(8)}$ from above.

Note that C_5 lies entirely below the w_2 axis for $w_1 \geq w_1^{(8)}$.

C_5 crosses the w_1 axis at $w_1^{(S)}\left(\frac{\theta_2 D\beta r}{\beta r+\theta_2 D}\right)^{k/\theta_1}$. It is above the w_1 axis for precisely the values to the right of this, but to the left of the asymptote

Important Values and Functions of x_1, x_2, z.

The following equation is important:

$$w_1 k e^{-kx_1} + (w_2+D)\theta_1(e^{-\theta_1 x_1} - \frac{1}{\beta r}) = 1 \quad \text{(A.18)}$$

This will have a unique solution $x_1 = x_1^{(4)}$, where (w_1, w_2) lie above L_3.

The following values of x_1 are important (for given w_1, w_2):

$$x_1^{(4)} = \text{unique solution of Equation A.18} \quad \text{(A.19)}$$

$$x_1^{(5)} = \frac{1}{k}(\log(w_1 k) - \log(1-\frac{\theta_1}{\theta_2})). \quad \text{(A.20)}$$

$$x_1^{(6)} = \frac{1}{k}\log(w_1 k). \quad \text{(A.21)}$$

$$\text{(A.22)}$$

Note that $x_1^{(4)}$ is a smooth function (by the Implict Function Theorem) and monotonically increasing in w_1 and w_2. Let H be an alternative name for this function, that is $x_1^{(4)} = H(w_1, w_2)$. The derivatives $\frac{\partial H}{\partial w_i}$ can be found by differentiating both sides of Equation A.18.

$$\frac{\partial H}{\partial w_1} = \frac{ke^{-kH(w_1,w_2)}}{w_1 k^2 e^{-kH(w_1,w_2)} + (w_2+D)\theta_1^2 e^{-\theta_1 H(w_1,w_2)}}$$

$$\frac{\partial H}{\partial w_2} = \frac{-\theta_1(\frac{1}{\beta r} - e^{-\theta_1 H(w_1,w_2)})}{w_1 k^2 e^{-kH(w_1,w_2)} + (w_2+D)\theta_1^2 e^{-\theta_1 H(w_1,w_2)}} \quad \text{(A.23)}$$

Both of these are strictly greater than 0.

The following values of x_2 are important:

$$x_2^{(2)} = \frac{1}{\theta_2}\log\left(\frac{(w_2+D)\theta_2\beta r}{\beta r + (w_2+D)\theta_2}\right) \quad \text{(A.24)}$$

$$x_2^{(5)} = \frac{1}{\theta_2}\log\left(\frac{\beta r\theta_2(w_2+D)}{\beta r + \theta_2(w_2+D)}\left(\frac{w_1^{(S)}}{w_1}\right)^{\theta_1/k}\right) \quad \text{(A.25)}$$

The following values of z are important:

$$z^{(1)} = \frac{1}{\beta}\log(\beta r) \quad \text{(A.26)}$$

$$z^{(2)} = \frac{1}{\beta}\log\left(\frac{\beta r + (w_2+D)\theta_2}{(w_2+D)\theta_2}\right) \quad \text{(A.27)}$$

$$z^{(5)} = \frac{1}{\beta}\log\left(1 + \frac{\beta r}{(w_2+D)\theta_2}\right). \quad \text{(A.28)}$$

Cases to Analyse in The Equilibrium Between Firm and Attacker

The calculations supporting these results are contained in the appendix.

The maximization problems for the firm and the attacker are solved simultaneously. Let the equilibrium be $(x_1^{**}, x_2^{**}, z^{**})$. We continue to suppose that $\beta r > 1$. Suppose that $D > 0$, although perhaps arbitrarily small — this simplifies the analysis, so that $w_2 + D > 0$. Figure 2 provides a map of cases that we analyze below.

Case 1. $w_1 \leq w_1^{(S)} = \frac{1}{k}(1 - \frac{\theta_1}{\theta_2})$ and $w_2 \leq w_2^{(2)} = \frac{\beta r}{\beta r - 1}\frac{1}{\theta_2} - D$.

In this case $x_1^{**} = x_2^{**} = 0$. Moreover $z^{**} = z^{(1)} > 0$.

Case 2. $w_1 \leq w_1^{(S)}$ and $w_2 > w_2^{(2)}$.

In this case $x_1^{**} = 0$ and $x_2^{**} = x_2^{(2)} > 0$. Moreover $z = z^{(2)} > 0$.

In cases 1 and 2, there is insufficient incentive to encourage compliance by the firm. In case 1, there is also insuffcient incentive for the firm to make discretionary security spend. spend, but in case 2, there is sufficient incentive.

As $\beta r \longrightarrow 1$ from above, the value $w_2^{(2)} \longrightarrow \infty$. That is, if the reward to the attacker is sufficiently

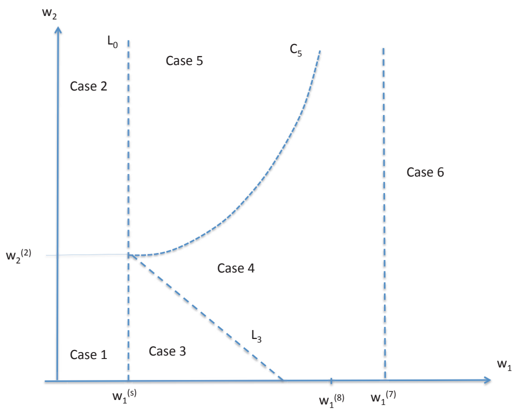


Fig. 2: Case Map for Analysis of Reactions in Subgames

small, then case 2 (effectively) vanishes — it is simply not worth spending on security as the attacker does not attack and there are limited punishments for non-compliance. When βr becomes sufficiently large, $w_2^{(2)}$ becomes zero (and then negative). In this situation, case 1 vanishes: even if there are low punishments for non-compliance, attacks mean that discretionary investment is still worthwhile. Note also that varying the value of D, changes the value $w_2^{(2)}$. If the firm is the final resting place (residual claimant) for a high-level of damage D arising from attacks, then $w_2^{(2)}$ will be lower than it would be if D were low. Thus, a high-level D will cause the firm to commence discretionary security spend at a relatively low-level of risk-based punishment w_2, when w_1 is also low ($w_1 < w_1^{(S)}$).

As $\theta_1 \longrightarrow \theta_2$ from below, $w_1^{(S)} \longrightarrow 0$: thus, as compliance investment becomes as efficient at mitigation as discretionary investment cases 1 and 2 become confined to a pure risk-based policy. However, case 1 is really only separated from case 3 for the purposes of analysis — they really have the same equilibrium between the firm and the attacker.

For θ_1 to become close to θ_2, this can likely only be achieved by an expensive (for the policy-maker) process of tailoring precise, current rules to the firm in its environment. A similar effect can be produced by taking k to be large, that is by using a risk-based punishment in which increasing levels of non-compliance lead to more rapidly increasing levels of punishment. However, this has further effects on the rest of the phase space, as witnessed by cases 4 and 5 below.

Case 3. $w_1 > w_1^{(S)}$ and (w_1, w_2) lies below (or on) L_3.
In this case $x_1^{**} = x_2^{**} = 0$. Moreover $z^{**} = z^{(1)} = \frac{1}{\beta}\log(\beta r) > 0$.

In case 3, there remains insufficient incentive for the firm to spend on compliance, regardless of any other factors. There is also insufficient incentive for the firm to make discretionary security spend, despite a level of residual attack $z^{(1)}$. This level of residual attack will be low if βr is close to 1, particularly if the returns to the attacker do not diminish slowly (β is not very small).

Case 4. $w_1 > w_1^{(S)}$ and (w_1, w_2) lies strictly below and right of C_5, but strictly above and right of L_3, and strictly to the left of $w_1^{(7)}$.
In this case $x_1^{**} = x_1^{(4)} > 0$ and $x_2^{**} = 0$. Moreover $z^{**} > 0$.

Case 5. $w_1^{(S)} \leq w_1 < w_1^{(8)}$ and (w_1, w_2) lies above C_5.
In this case $x_1^{**} = x_1^{(5)} > 0$ and $x_2^{**} = x_2^{(5)} > 0$. In this equilibrium $z^{**} = z^{(5)} > 0$.

Case 6. $w_1 \geq w_1^{(7)}$ and (w_1, w_2) lies strictly above L_3.
$x_1^{**} = x_1^{(6)} > 0$ and $x_2^{**} = 0 = z^{**}$

In this case, the firm expects that the attacker will not attack, whilst the attacker expects the firm will nevertheless invest heavily in compliance (and therefore makes attacks insufficiently beneficial to the attacker), because of a high-level of punishment for non-compliance.

Subgame Equilibria Calculations

The existences of cases 1,2,3 depends upon the geometry suggested in the 'geometry of the interaction' section. The lines and curves L_0, L_3, C_5 and the value $w_2^{(2)}$ move around as the model parameters change. This makes some cases disappear for some parameter choices.

Cases 1, 2.
Since $\frac{\partial \mathscr{U}_F}{\partial x_2} \mid_{\substack{x_1 = x_1^{**} \\ x_2 = x_2^{**} \\ z = z^{**}}} \leq 0$, we also have $\frac{\partial \mathscr{U}_F}{\partial x_1} \mid_{\substack{x_1 = x_1^{**} \\ x_2 = x_2^{**} \\ z = z^{**}}} \leq w_1 k - (1 - \frac{\theta_1}{\theta_2})$ (in case either $z^{**} = 0$ or $z^{**} > 0$. We must therefore have $x_1^{**} = 0$ because $w_1 < w_1^{(S)}$.

Suppose $x_2^{**} > 0$. Then $z^{**} > 0$, and $-1 + \beta r e^{-\beta z^{**}} e^{-\theta 1 x_1^{**} - \theta_2 x_2^{**}} = 0$. But then it must be the case that $(w_2 + D)\theta_2(e^{-\theta_2 x_2^{**}} - \frac{1}{\beta r}) > 0$.

In **Case 1** this is not possible, since $w_2 < w_2^{(2)}$, so $x_2^{**} = 0$, and then $z^{**} = \frac{1}{\beta}\log(\beta r)$.

In **Case 2** we find that $x_2^{**} = 0$ leads to a contradiction, and therefore that $x_2^{**} = x_2^{(2)}$ and $z^{**} = z^{(2)}$.

Case 3.

As in case 4 below we get $x_2^{**} = 0$. However, here A.18 has no solution, so $x_1^{**} = 0$. We then have $z^{**} = \frac{1}{\beta}\log(\beta r)$.

Case 4.

Similar to 5, except that we have that $x_2^{(5)} < 0$.

Firstly, it follows that $x_2^{**} = 0$. If $x_2^{**} > 0$, then $\frac{\partial \mathscr{U}_F}{\partial x_2} \mid_{\substack{x_1 = x_1^{**} \\ x_2 = x_2^{**} \\ z = z^{**}}} = 0$, so that $\frac{\partial \mathscr{U}_F}{\partial x_1} \mid_{\substack{x_1 = x_1^{**} \\ x_2 = x_2^{**} \\ z = z^{**}}} = w_1 k e^{-k x_1^{**}} + (1 - \frac{\theta_1}{\theta_2})$, and hence $x_1^{**} > 0$. In this case, the equilibrium lies at the stationary point, but then $x_2^{**} < 0$ since (w_1, w_2) lies below C_1, yielding a contradicton.

Since (w_1, w_2) lies above lies above L_3, equation A.18 has a solution $x_1^{(4)}$ and so we have $x_1^{**} = x_1^{(4)} > 0$.

Case 5.

Note that there can be no equlibrium here with $z^{**} = 0$ (for the first-order condition on $\mathscr{U}_A$ would require $w_1 > w_1^{(7)}$). Therefore $z^{**} > 0$.

In this case , $e^{\theta_1 x_1^{**} + \theta_2 x_2^{**}} \leq \beta r$ and

$$-1 + \beta r e^{-\beta z^{**}} e^{-\theta 1 x_1^{**} - \theta_2 x_2^{**}} = 0. \tag{A.29}$$

Under the given conditions, a stationary-point with $\frac{\partial \mathscr{U}_F}{\partial x_1} |_{z=z^{**}} = 0$ and $\frac{\partial \mathscr{U}_F}{\partial x_2} |_{z=z^{**}} = 0$ exists and lies at

$$x_1^{(S)} = x_1^{(5)} = \frac{1}{k} \log\left(\frac{k w_1}{1 - \frac{\theta_1}{\theta_2}}\right) = \frac{1}{k} \log\left(\frac{w_1}{w_1^{(S)}}\right) \tag{A.30}$$

and

$$x_2^{(S)} = x_2^{(5)} = \frac{1}{\theta_2} \log\left(\frac{\beta r \theta_2 (w_2 + D)}{\beta r + \theta_2 (w_2 + D)} \left(\frac{w_1^{(S)}}{w_1}\right)^{\theta_1 / k}\right) \tag{A.31}$$

Now $w_1 \geq w_1^{(S)}$ by assumption, so $x_1^{(S)} > 0$. Moreover (w_1, w_2) lies above the curve C_5, so $x_2^{(S)} \geq 0$. (The curve C_5 is chosen to make this so.)

We therefore have $x_1^{**} = x_1^{(S)}$ and $x_2^{**} = x_2^{(S)}$.

Moreover, we find that

$$z^{**} = z^{(5)} = \frac{1}{\beta} \log\left(1 + \frac{\beta r}{(w_2 + D)\theta_2}\right). \tag{A.32}$$

Note that $(w_2 + D)\theta_2 > \beta r$, under the given assumptions, so $z^{**} > 0$, and the equilibrium at $(x_1^{**}, x_2^{**}, z^{**})$ is consistent.

Case 6.

Suppose that $z^{**} = 0$. In this case $x_2^{**} = 0$. Now $\beta r > 1$ and $\beta r \leq e^{\theta^1 x_1^{**} + \theta_2 x_2^{**}}$, so it must be the case that $x_1^{**} > 0$. Using $\frac{\partial \mathscr{U}_F}{\partial x_1} \mid_{\substack{x_2 = x_2^{**} = 0 \\ z = z^{**}}} = 0$, we find the stationary point of at

$$x_1^{(S)} = x_1^{(6)} = \frac{1}{k} \log(w_1 k). \tag{A.33}$$

If $w_1 > w_1^{(7)}$, then $\beta r \leq e^{\theta^1 x_1^{**} + \theta_2 x_2^{**}}$. $\beta r \leq e^{\theta^1 x_1^{**} + \theta_2 x_2^{**}}$ If $w_1 > w_1^{(7)}$ then $x_1^{**} = x_1^{(6)}$.

If $w_1 < w_1^{(7)}$ then there is no NE with $z^{**} = 0$. If $w_1 \geq w_1^{(7)}$ then there is no NE with $z^{**} > 0$.

Payoffs at Equilibria in Subgames

Equilibria Within Cases (Firm Reaction Phases)

Case 1 In this case, the reactions by the attacker and firm reduce the payoff functions to the functions of w_1 and w_2 only (not x_1, x_2, z):

$$\mathscr{U}_A = r(1 - \frac{1}{\beta r}) - \frac{\log(\beta r)}{\beta} \tag{A.34}$$

$$\mathscr{U}_F = t - w_1 - (D + w_2)\left(1 - \frac{1}{\beta r}\right) \tag{A.35}$$

$$\mathscr{U}_P = -\left(1 - \frac{1}{\beta r}\right)\mu - t - \nu \tag{A.36}$$

The attacker is insensitive to policy and the choices of the firm. The firm loses more as compliance and risk penalties are increased. *Within this case*, the policy-maker P does not itself care about w_1 and w_2 for a fixed transfer making the case feasible.

The firm's individual rationality constraint (IRC) demands that for a response of this type to be feasible, the policy-maker must allocate any transfer $t \geq D\left(1 - \frac{1}{\beta r}\right)$. Since the firm does not respond in this case (by setting $x_1 > 0$ or $x_2 > 0$), the optimal choice for both the basic and alternate assurance policy-maker is to set $t = D\left(1 - \frac{1}{\beta r}\right)$ and $w_1 = 0$ and $w_2 = 0$.

This is essentially a world in which security of very little concern. Importantly, note that observance of the firm's rationality constraint forces a de facto indemnification of the firm by the policy-maker for the otherwise unindemnified damage.

Case 2

We find on the region defined in case 2:

$$\begin{array}{rcl}\mathscr{U}_A & = & \frac{r}{\theta_2(D+w_2)} - \frac{1}{\beta}\log\left(\frac{\theta_2(D+w_2)+\beta r}{\theta_2(D+w_2)}\right)\\ \mathscr{U}_F & = & t - w_1 - \frac{1}{\theta_2} - \frac{1}{\theta_2}\log\left(\frac{(D+w_2)\theta_2\beta r}{(D+w_2)\theta_2+\beta r}\right)\\ \mathscr{U}_P & = & -t-\nu-\frac{\mu}{\theta_2(D+w_2)}\end{array} \tag{A.37}$$

$\mathscr{U}_P$ monotonically increases as w_2 increases, but is insensitive to w_1. Thus, within this case, the policy-maker wishes simply to make w_2 as large as possible, before taking constraints into consideration.

We find for the firm:

$$\mathscr{U}_F = \;=\; t - w_1 - \frac{1}{\theta_2} - \frac{1}{\theta_2}\log\left(\frac{1}{\frac{1}{(D+w_2)\theta_2}+\frac{1}{\beta r}}\right) \tag{A.38}$$

Note that $\mathscr{U}_F$ is decreasing in both w_1 and w_2.

In this case the IRC, $\mathscr{U}_F \geq 0$, then becomes

$$w_1+\frac{1}{\theta_2}\log\left(\frac{\beta\theta_2 r(D+w_2)}{\theta_2(D+w_2)+\beta r}\right) \leq t-\frac{1}{\theta_2}. \tag{A.39}$$

Note that this can have a vertical asymptote that lies within case 2, or to the right of its right-hand boundary. In this case, for those w_1 in case 2 to the left of the asymptote, all w_2 satisfy the IRC.

Case 2 is feasible (provided it exists becasue the parameters allow it) when the point $(0, w_2^{(2)})$ in the (w_1, w_2)-plane satisfies the IRC. This happens when

$$t \geq \frac{1}{\theta_2}. \tag{A.40}$$

Note that this is the minimum transfer that allows the firm to operate under a purely risk-based policy-regime: it is simply inversely proposritonal to the 'efficiency of mitigation of discretionary spend' θ_2.

Issues of asymmetry of information play a major role at this point. The parameter θ_2 may not be completely transparent to the policy-maker. There is a residue of adverse selection here: a firm that is inefficient should receive a larger transfer! So there is a perverese incentive for the firm to declare that it is less efficient than it really is in order to recieve a larger transfer, and make a gain by not spending it all on security. (This may be transferred into gold-plated (security) assets rather than profits so as not to be caught out). It declares $\hat{\theta_2}$ when really it has efficiency parameter $\theta_2 > \hat{\theta_2}$. It then gets a transfer $1/\hat{\theta_2}$ and then improves its utility to

$$\begin{array}{rcl}\hat{\mathscr{U}_F} & = & \frac{1}{\hat{\theta_2}} - w_1 - \frac{1}{\theta_2} - \frac{1}{\theta_2}\log\left(\frac{(D+w_2)\theta_2\beta r}{(D+w_2)\theta_2+\beta r}\right)\\ & > & \frac{1}{\theta_2} - w_1 - \frac{1}{\theta_2} - \frac{1}{\theta_2}\log\left(\frac{(D+w_2)\theta_2\beta r}{(D+w_2)\theta_2+\beta r}\right)\end{array}$$

whilst the policy-maker (society) loses out commensurately. Note that the actual efficiency does not change, and is not used by the policy-maker in the calculation of punishments (only the observed level of exploitation), so only the transfer term changes.

Moreover, the firm loses if it presents an overestimate of its efficiency θ_2. It has a strong incentive to not overestimate.

Moreover, if, realistically, its ability to hide such gains is as a percentage of the transfer (rather than the actual amount), then the firm has an interest that the transfer should be as large as possible. This may be why the regulator feels that it has to be extremely careful about over-financing security in a risk-based system.

Note also that if the regulator can observe changes in efficiency at some 'large' threshold, then the firm will have an incentive only to increase its efficiency parameter θ_2 to just below that threshold. There is an incentive to become a little more efficient, but only up to a certain point. The Laffont-Tirole style analysis of contractual incentives for efficiency gain is probably appropriate in such a circumstance.

There is an additional complication in the security situation: it is difficult for the firm to know its own θ_2, since measurement of this depends upon the attacker as well as itself. In the purely risk-based environment, the firm might not know βr in the non-zero rate of exploitation $e^{-\theta_2 x_2} - \frac{1}{\beta r}$, and this is how it measures θ_2 using a known level of x_2. Moreover, the attacker may not be able to estimate θ_2 correctly. There can be two-way information asymmetry. The incentive not to underestimate may be amplified by this.

The above considerations are for when the policy-maker is to choose at the minimal feasible value of w_2. Now let us consider the optimal, feasible w_2.

Suppose that case 2 is feasible.

The individual rationality constraint for the firm, $\mathscr{U}_F \geq 0$, can be rewritten in the form

$$\exp(-\theta_2 t+\theta_2 w_1+1)-\frac{1}{\beta r} \leq \frac{1}{\theta_2(D+w_2)}. \tag{A.41}$$

Let $L = \exp(-\theta_2 t+\theta_2 w_1+1) - \frac{1}{\beta r}$. It may be that L is positive, zero or negative.

A policy (w_1, w_2) in case 2 satisfies the firm's IRC if $w_1 \leq w_1^{(11)}$, or else

$$w_2 \leq -D + \frac{\beta r}{\theta_2(-1 + \beta r \exp(1 + w_1\theta_2 - t\theta_2)))}; \tag{A.42}$$

The policy-maker would like to set as large a w_2 as possible. However, P would also like to minimize t.

Recall that $w_1^{(11)}$ is a monotonically increasing function of t.

If $w_1 \leq w_1^{(11)}$ then the policy-maker can never drive the firm down to $\mathscr{U}_F = 0$ by increasing w_2, because the transfer is too large and the environment is too safe to generate sufficient risk to absorb that transfer.

Suppose that the policy-maker wishes to achieve a policy (w_1, w_2). If $w_1 < w_1^{(11)}$, then too large a transfer has been allocated. Indeed, the optimal transfer is given by finding t by setting the constraint of Equation A.42 to be binding.

Now since $\mathscr{U}_P$ is not sensitive to w_1 (since $x_1 = 0$ within this case), it is never worse for the policy-maker to reduce w_1. On the other-hand, the IRC for the firm is decreasing in w_1 (on this case) where it is defined. It may be that there many equally good optima to the left of the asymptote. However, none is better than the purely risk-based policy $(0, \min(\overline{w_2}, -D + \frac{\beta r}{\theta_2(-1+\beta r \exp(1-t\theta_2)))})$, where $\overline{w_2}$ is a cap on the level of punishment under consideration.

This tells us how to set t.

Firstly, it is clear that the policy-maker should set

$$\frac{1}{\theta_2} \leq t \leq \frac{1}{\theta_2}(1 + \log(\beta r)) \tag{A.43}$$

since if $t \geq \frac{1}{\theta_2}(1 + \log(\beta r))$ then $w_1^{(11)} \geq 0$, and the optimal w_2 is at the upper bound $\overline{w_2}$, so that the increase of transfer will only reduce the policy-maker's payoff. Note again that the maximum transfer increases as θ_2 decreases, so this would tend to induce the firm to understate θ_2.

Now suppose that $\frac{1}{\theta_2} \leq t \leq \frac{1}{\theta_2}(1+\log(\beta r))$. In this case we find $\mathscr{U}_P = \frac{mu}{\beta r} - \nu - \alpha(t)$ where $\alpha(t) = t + \mu \exp(1 - t\theta_2)$.

If $\theta_2\mu \leq 1$ then $\mathscr{U}_P$ attains its maximum at $t = \frac{1}{\theta_2}$. If $\theta_2\mu \geq 1$ and $\mu\theta_2 \leq \beta r$ then $\mathscr{U}_P$ attains its maximum at $t = \frac{1}{\theta_2}(1 + \log(\mu\theta_2))$. If $\theta_2\mu \geq 1$ and $\mu\theta_2 \geq \beta r$, then $\mathscr{U}_P$ attains its maximum at $t = \frac{1}{\theta_2}(1 + \log(\beta r))$.

Note that the policy-maker's optimum transfer depends on the aggressiveness of the environment, but also upon a combination of the social damage multiplier μ and the efficiency parameter θ_2.

Note that (within this case) it is in the firm's interests for P to believe that θ_2 is small, whilst $\mu\theta_2$ is large. Thus the firm would wish to play up the possibility of social damage, whilst exaggerating the costs of mitigation.

Case 3 This case produces the same equilibrium $x_1^{**}, x_2^{**}, z^{**}$ as case 1, so the same $\mathscr{U}_A, \mathscr{U}_F, \mathscr{U}_P$.

Case 4

The analysis in this case is less straightforward than the others, because the compliance reaction x_1 is determined by an implict function of w_1 and w_2.

We find in this case:

$$\begin{aligned}
\mathscr{U}_A &= r\left(e^{-\theta_1 x_1} - \tfrac{1}{\beta r}\right) - \tfrac{1}{\beta}(\log(\beta r) - \theta_1 x_1) \\
\mathscr{U}_F &= t - (D + w_2)\left(e^{-\theta_1 x_1} - \tfrac{1}{\beta r}\right) - w_1 e^{-k x_1} \\
&\quad -x_1 \\
\mathscr{U}_P &= -t - \mu(e^{-\theta_1 x_1} - \tfrac{1}{\beta r}) - \nu e^{-\kappa x_1}
\end{aligned} \tag{A.44}$$

where $x_1 = x_1^{(4)} = H(w_1, w_2)$ is given by the function H from Equation A.18 and $z = \frac{1}{\beta}(\log(\beta r) - \theta_1 x_1)$.

Note that the policy-maker is not indifferent to w_1 and w_2 within this case. Recall that the function H is monotonically increasing in w_1 and w_2. Within this case, the policy-maker wishes $H(w_1, w_2)$ to be as large as possible. We find $\partial\mathscr{U}_P/\partial w_i > 0$ for $i = 1, 2$. Thus it wishes to maximize both w_1 and w_2. It may seem perverse that that there can be an advantage in increasing the risk-based penalty when this results in no increased in risk-based behaviour by the firm. However, this is because an increase in w_2 (remaining within case 4) causes an increase in compliance x_1, and this results in benefit to the policy-maker.

There are constraints on how the policy-maker can increase w_1 and w_2. There is the feasability of the policy (wrt L_3 and C_5) and there is the firm's IRC and investment upper-bounds. Let us consider the IRC first.

The firm's utility is a complicated function of w_1 and w_2. It can be re-writen as

$$\mathscr{U}_F = (t - \tfrac{1}{\theta_1}) + w_1(\tfrac{k}{\theta_1} - 1)e^{-k x_1} - x_1 \tag{A.45}$$

The firm's IRC occurs where $\mathscr{U}_F = 0$.

Now

$$\frac{\partial \mathscr{U}_F}{\partial w_2} \quad = \quad -(e^{-\theta_1 H} - \tfrac{1}{\beta r}) - (w_2 + D)\tfrac{\theta_1}{\beta r}\tfrac{\partial H}{\partial w_2} < 0. \qquad \text{(A.46)}$$

In contrast, the sign of $\frac{\partial \mathscr{U}_F}{\partial w_1}$ can be negative, positive or zero on the region defined by case 4, depending upon the parameters.

Case 5 We find:

$$\text{Exploited}\,(x_1^{**}, x_2^{**}, z^{**}) = \frac{1}{\theta_2(w_2 + D)} \qquad \text{(A.47)}$$

and

$$\mathscr{U}_A \quad = \quad \tfrac{r}{\theta_2(D+w_2)} - \tfrac{1}{\beta}\log(1 + \left(\tfrac{\beta r}{\theta_2(D+w_2))}\right))$$

$$\begin{aligned}\mathscr{U}_F \quad &= \quad t - \tfrac{1}{\theta_2} - w_1^{(S)} - w_1^{(S)} \log\left(\tfrac{w_1}{w_1^{(S)}}\right) \\ &\quad -\tfrac{1}{\theta_2}\log\left(\tfrac{\beta r\theta_2(w_2+D)}{\beta r+\theta_2(w_2+D)}\right)\end{aligned}$$

$$\mathscr{U}_P \quad = \quad -t - \tfrac{\mu}{\theta_2(D+w_2)} - \nu\left(\frac{w_1^{(S)}}{w_1}\right)^{(\frac{\kappa}{k})} \qquad \text{(A.48)}$$

where, recall $w_1^{(S)} = \frac{1}{k}(1 - \frac{\theta_1}{\theta_2})$.

The policy-maker wishes to increase both w_1 and w_2 as much as possible. The firm's IRC will lead to a trade-off between w_1 and w_2. Note that $\mathscr{U}_F$ is decreasing in both w_1 and w_2. The firm can thus participate in case 5, precisely when

$$\mathscr{U}_F(w_1^{(S)}, w_2^{(2)}) = t - \frac{1}{\theta_2} - w_1^{(S)} \geq 0.$$

Therefore a case 5 response requires at least a minimum transfer

$$t = w_1^{(S)} + \frac{1}{\theta_2} = \frac{1}{k} + \frac{1-\theta_1}{\theta_2}. \qquad \text{(A.49)}$$

Getting a mixed response from the firm thus requires an additional $w_1^{(S)} = \frac{1}{k}(1 - \frac{\theta_1}{\theta_2})$ transfer from the firm, compared to the pure risk-based approach. Note that the same incentive effects involving revelation by the firm of the true value of θ_2 apply as in case 2 (because $\theta_1 < 1$).

The firm's IRC $\mathscr{U}_F \geq 0$ gives that the policy (w_1, w_2) must not lie above the curve C_2 given by

$$w_2 = -D+ \frac{\beta r}{\theta_2\left(\beta r e^{(-\theta_2 t+1+\theta_2 w_1^{(S)})}\left(\frac{w_1}{w_1^{(S)}}\right)^{\frac{1}{k}(\theta_2-\theta_1)} -1\right)} \qquad \text{(A.50)}$$

Below this curve (in case 5), $\mathscr{U}_F > 0$, and above it $\mathscr{U}_F < 0$.

The curve has an asymptote at

$$w_1 = w_1^{(9)} = w_1^{(S)}\left(\frac{e^{\theta_2 t-1-\theta_2 w_1^{(S)}}}{\beta r}\right)^{k/(\theta_2-\theta_1)} \qquad \text{(A.51)}$$

Note that the curve is not meaningful for $w_1 \leq w_1^{(9)}$. To the left of this asymptote, $\mathscr{U}_F > 0$. To the right of $w_1^{(9)}$, the curve C_2 has negative slope. When $w_1^{(9)}$ is to the right of $w_1^{(S)}$, the asymptote for case 2 (at $w_1^{(11)}$) is also to the right of $w_1^{(S)}$, and so all of case 2 is feasible.

Above $w_2^{(2)}$, the IRC curve for case 2 meets L_0 if and only if the IRC fr case 5 meets L_0, and in which case they meet at the same point.

Now consider any feasible policy (w_1, w_2) in case 5. Any $w_1' < w_1$ gives (w_1', w_2) a feasible policy. For a basic policy-maker P with $\nu = 0$, there is never any advantage to using a mixed policy (w_1, w_2) in case 5 over the projected pure risk one $(0, w_2)$. The minimum transfer required to achieve (w_1, w_2) be on the IRC, giving $\frac{1}{\theta_2} + w_1^{(S)} + w_1^{(S)}\log(\frac{w_1}{w_1^{(S)}}) + \frac{1}{\theta_2}\log(\frac{\beta r\theta_2(w_2+D)}{\beta r+(w_2+D)t\theta_2})$, whereas for $(0, w_2)$ it is $\frac{1}{\theta_2} + \frac{1}{\theta_2}\log(\frac{\beta r\theta_2(w_2+D)}{\beta r+(w_2+D)t\theta_2})$. P can always achieve a better payoff at $(0, w_2)$ than it can at (w_1, w_2).

Things are more complex for the policy-maker with assurance, $\nu > 0$. For a given transfer, the optimum for $\mathscr{U}_P$ on case 5 will occur on the IRC, or at $\overline{w_2}$ where the firm's IRC intersects C_5 above $\overline{w_2}$ or $w_1^{(9)} > w_1^{(8)}$ (the asymptote for the IRC is to the right of the asymptote for the right-hand case boundary.) Note that $\mathscr{U}_P^{(5)}(w_1, w_2) > \mathscr{U}_P^{(2)}(w_1', w_2) = \mathscr{U}_P^{(2)}(0, w_2)$. That is, the payoff in case 5 at (w_1, w_2) is always higher than the payoff for any (w_1', w_2) in case 2 with the same w_2.

However, the transfer required to make $(0, w_2)$ feasible is less that the transfer required to make (w_1, w_2) feasible, as noted above. There is a tension here between the transfer allocated and the component of the payoff corresponding to assurance.

The height of the firm's IRC (above the w_1-axis) is monotonically increasing in t as is the w_1 position of its aymptote. Hence, there is never any point considering a t that is larger than that which makes the IRC meet C_5 where $w_2 = \overline{w_2}$. That is, the maximum transfer that a policy-maker should ever choose for this case is

$$t = \frac{1}{\theta_2} + w_1^{(S)} + \frac{1}{\theta_1}\log\left(\frac{\beta r\theta_2(\overline{w_2}+D)}{\beta r + \theta_2(\overline{w_2}+D)}\right). \tag{A.52}$$

Note that the final summand here is greater than 0, provided we have chosen $\overline{w_2} > w_2^{(2)}$ Thus the range of sensible transfers for Case 5 is where both

$$\frac{1}{\theta_2} + w_1^{(S)} \leq t$$

and

$$t \leq \frac{1}{\theta_2} + w_1^{(S)} + \frac{1}{\theta_1}\log\left(\frac{\beta r\theta_2(\overline{w_2}+D)}{\beta r + \theta_2(\overline{w_2}+D)}\right).$$

Note that the upper bound on t above may be greater than a fixed upper bound on the transfer that the policy-maker wishes to consider, $\bar{t}$, in which case the range of w_2 on case 5 is further truncated (by finding the value of $\overline{w_2}$ entailed by putting $\bar{t}$ in equation A.52). So we may suppose that the optimum occurs on the firm's IRC. On this curve, the polic-maker's payoff can be written (for fixed t) in terms of w_1 alone.

Conclusion for Case 5. If a policy-maker wants an effective partly rules-based, partly risk-based system, then several things are true.

Firstly, the punishment for non-compliance must be sufficiently biting ($w_1 > w_1^{(S)}$ in the model) and this may depend upon the efficiency of risk-based mitigation (θ_2) as well as the efficiency of compliance (θ_1) and the rate at which punishments increase (k).

The punishment for inadequate risk-based mitigation must be sufficiently biting ($w_2 > w_2^{(2)}$). This may depend upon the efficiency of risk-based mitigation and the aggressiveness of the attacker, but may be reduced by any damage (D) that the firm may itself sustain. The upper-limit on risk-based punishment must be greater than the absolute w_2-threshold for a mixed reaction ($\overline{w_2} > w_2^{(2)}$).

More generally, as the punishment for non-compliance increases from the minimum threshold for a compliance reaction ($w_1^{(S)}$), this may require an increase in the level of risk-based punishment (w_2) in order to continue eliciting the same level of (non-zero) risk-based response (e.g., the curve C_5 slopes upwards, but so do the contours of x_2 on this region).

The policy-maker should be aware that too heavy a punishment for non-compliance can kill off any discretionary risk-based response (crowding out when $w_1 \geq w_1^{(8)}$, the asymptote for curved boundary C_5 of case 5).

There must be a genuine desire for compliance on behalf of the policy-maker (there must be a significant term in its payoff function, as in $u_{P_{ass}}$); otherwise the pure risk-based system achieves the same results at less cost (less transfer required for the basic P).

If the policy-maker cares about assurance and there is a mixed policy (eliciting a mixed response) then there is then a trade-off here between the effects of w_1 and w_2. Which combination is most effective depends upon many aspects of the system (combinations of parameters).

Case 6 The level of exploitation is mitigated in this case down to zero.

The payoffs reduce in this case to:

$$\begin{aligned}\mathscr{U}_A &= 0\\ \mathscr{U}_F &= t - \tfrac{1}{k} - \tfrac{\log(kw_1)}{k}\\ \mathscr{U}_P &= -t - \nu\left(\tfrac{1}{kw_1}\right)^{\frac{\kappa}{k}}.\end{aligned} \tag{A.53}$$

Within this case it does not matter to the basic policy-maker with $\nu = 0$, how w_1 and w_2 are chosen. Note that the payoff to $\mathscr{U}_P$ at a given transfer is better in this case than anywhere else in the space of (w_1, w_2) points. The policy maker with assurance prefers w_1 to be as large as possible. The firm's IRC here gives that the largest feasible w_1 is

$$\frac{1}{k}e^{kt-1} \tag{A.54}$$

and w_2 is irrelevant (since there are no attacks and $\mathscr{U}_F$ does not depend upon w_2).

Thus, even for the policy-maker with $\nu > 0$ there is a bound on the level of w_1. For a given t enabling this case, the maximum level of $\mathscr{U}_P$ is

$$-t - \nu\exp(\kappa(\frac{1}{k} - t). \tag{A.55}$$

In order for this case to be feasible, it must be what $\frac{1}{k}e^{kt-1} \geq w_1^{(7)}$. That is,

$$t \geq \frac{1}{k} + \frac{1}{\theta_1}\log(\beta r). \tag{A.56}$$

For the basic policy-maker with $\nu = 0$ the optimal transfer for this case is the minimum one $t \geq \frac{1}{k} + \frac{1}{\theta_1}\log(\beta r)$. At this point, all attackers drop out and no more needs to be spent. hence the optimal payoff for P in such cases (taken over all t) is $-(\frac{1}{k}+\frac{1}{\theta_1}\log(\beta r))$.

For the policy-maker with assurance, $\nu > 0$, the function of t given in the expression A.55 must be maximized. An unconstrained maximum for this occurs at $t^* = \frac{1}{k} + \frac{1}{\kappa}\log(\kappa\nu)$. If $(\kappa\nu)^{\frac{1}{\kappa}} \leq (\beta r)^{\frac{1}{\theta_1}}$ then the optimal t is the minimum one for this case. If $\frac{1}{k}+\frac{1}{\kappa}\log(\kappa\nu) \geq \bar{t}$ then the optimal transfer is the global maximum $\bar{t}$. Otherwise, the optimal transfer is t^*.

Conclusion for Case 6. A sufficiently large transfer and a pure compliance approach can completely neutralize attackers. The size of the transfer depends on the aggressiveness, βr, of the attacker.

Assume that the policy-maker targets drop-out of all attacks as in this case. Note that in the setting of the transfer, the firm may wish to suggest that the efficiency of mitigation, θ_1, is low, but the aggressiveness of attackers, βr, is high. The firm will also seek a regime with a low rate at which compliance punishment increases, k.

For the policy-maker without assurance, $\nu = 0$, at any given level of transfer, the payoff in this case (assuming it is feasible) is better than in any other case. (However, note that this does not say that this case is optimal when the transfer may be varied.)

Note that no maximum level of w_1 is required for this model (no companion to $\overline{w_2}$) ,as the firm's IRC always bounds the response in the w_1 dimension.

For a policy-maker who does not care about assurance, the optimal transfer (for this case) is the minimum one for this case.

For a policy-maker who does care about assurance, the optimal transfer (for this case) depends upon its preference for assurance, the agressiveness of the environment and the maximum transfer available.

ACKNOWLEDGEMENTS

This work has been partly funded by the European Union's 7th Framework Programme under grant agreement no 285223 - SECONOMICS (www.seconomics.org).

REFERENCES

Epstein, R. A. and T. P. Brown (2008). Cybersecurity in the payment card industry. *The University of Chicago Law Review 75*(1), 203–223.

L. Kaplow and S. Shavell (1996). Accuracy in the assessment of damages. *Journal of Law and Economics 39*(1), 191–210.

Laffont, J.-J. and J. Tirole (1986). Using cost observation to regulate firms. *Journal of Political Economy 94*(3), 614–641.

Laffont, J.-J. and J. Tirole (1993). *A Theory of Incentives in Procurement and Regulation.* The MIT Press.

MATLAB (2014). *version 8.3.0.532 (R2014a).* Natick, Massachusetts: The MathWorks Inc.

Westby, J. R. (Ed.) (2004). *International Guide to Cyber Security.* American Bar Association.

Fairness in Airport Security Expenditures: Equilibrium and Optimum

Woohyun Shim, Fabio Massacci, Alessandra Tedeschi, Julian Williams

Following the September 11 attacks, various security regulations that require increased expenditures have been enacted for airport security. While these regulations might increase overall security level, we presently do not have in-depth discussion on whether they are fair for airports with different nature. Particularly, by applying an earlier analysis of interdependent security risks, we investigate whether mandated security expenditures are fair for small, medium and large airports. We first develop a game-theoretical model with interdependent security risks, in order to evaluate the Nash equilibrium and socially optimal security expenditures. We then conduct a simulation analysis to test fairness of mandated security expenditures, and discuss the implications of the results.

KEY WORDS: Risk versus rules, regulation of critical infrastructure security

1. INTRODUCTION

After September 11, 2001, security regulations in the aviation industry have been tightened intensively. Accordingly, security costs represent up to 35% of overall airport operating costs (European Commission (2009)), and airport operators need to decide the best mechanism for the resource allocation in compliance with regulatory standards. As for a policy-maker, determining the optimal level of security expenditures has become a major task.

It is likely that that policy-makers have long believed that a high level of security investments are essential to address the threats and risks posed by international terrorism and to restore public confidence in the aviation security. Therefore, they have tried to enforce regulatory rules that mandate security expenditures of airport operators more than a certain amount.

As airports' security expenditures are directed by the regulators, however, various questions regarding the fairness of these mandatory expenditures have arisen. For example, some authors have recently pointed out that the optimal security expenditures are likely to vary across airports (e.g., Bier et al. (2008)), and each airport might have different security preferences. In a series of interviews, we also found that the regulators' passion for making a sound security environment by mandatory expenditures does not align well with the interest of airport operators. The airport operators seem to think that mandatory security expenditures that are set uniformly might not align well with the airports' incentives since security activities required by heterogeneous airports are different.

In this situation, while fully tailored security charges and mandatory security expenditures would maximize social surplus, the regulators might not be able to do this since it will be too costly for them to explore every available options. Even if they can identify all the possible options, it would be very difficult for them to determine the correct level of mandatory expenditures due to limited information. As a result, the regulators face to set

[1] DISI - University of Trento, Italy
[2] DISI - University of Trento, Italy
[3] Deep Blue, Italy
[4] University of Durham, UK

the level of security expenditures based on their limited knowledge and it might result in suboptimal outcome. Furthermore, the security interdependence between airports might make this problem more severe.

In this study, we therefore aim at addressing directly to this issue and investigating whether airport security investments mandated by the regulators are determined fairly for airports with different characteristics. In detail, we demonstrate *how mandatory security investments might undermine fairness in the context of airport security.* We identify taxonomies of technological factors and interdependence of airport security, and study how they can cause a divergence between unregulated private actions and those that would maximize the overall social surplus against terrorism.

In order to make our study more constructive, we develop a formal model that captures the interaction between airports, attackers and a regulator. The model contributes to the literature on public policy and economics that employs a game theoretical model to study strategic investment decisions. One of the crucial innovation of our approach is *to take into account the importance of interdependence of security environment and the strategic interaction with attackers in analysing security expenditure in civil aviation.* The model illustrates the circumstances under which the social and private incentives to invest in airport security can be expected to differ. One of the surprising results is that even though mandated security expenditures minimize the overall cost to society from an attack, the distribution of security expenditures for airports with different nature would be unfair. That is, a security regulation tends to shift the burden of security expenditures from large airports with high risk of terrorist attack to small airports with low risk of terrorist attack.

The paper proceeds as follows: Section 2 reviews the related literature. Section 3 outlines our model and show how private and social incentives for security can differ. Section 4 provides numerical illustration for different settings. Finally, Section 5 offers some concluding remarks.

2. LITERATURE REVIEW

This study is grounded in two distinct, but interconnected research domains: Economics of public good provision and fair cost allocation, which can both inform a policy-maker's decision making process. We first review the literature on traditional public good provision and fair cost allocation and then explore more recent works that has evolved as a new paradigm for managing aviation security risk.

It has long been accepted that security has public good characteristics (i.e., nonrivalry and nonexcludability). In public economics, there has been a wide array of studies on social optimum characterizations with interdependence in the provision of a public good including security. A formal approach on how economics look at the government role in public good provision began in the seminal book of Wallace (1972). In this book, under the assumption of no interdependence or interdependence among jurisdictions, Oates provides analysis on the tradeoff between the government's different fiscal systems from economic point of view for the first time. This influential work brought together the idea of an ideal regulatory system that can guide public goods and services that should be offered in localities and determine how to finance them.

In the subsequent works, many authors have also tried to identify most efficient regulatory system in pubic good provision (e.g., Oates (1999); Dur and Roelfsema (2005); Besley and Coate (2003)). The crucial feature of these studies is that the optimal regulatory structure for public good provision brings about a potential trade-off: for example, Besley and Coate (2003) show that the trade-off between different regulatory systems are determined by the heterogeneity of preference and the level of interdependence. They further argue that a decentralized regulatory system can enjoy the benefits from reflecting diverse preferences for public good provision, it would cause the costs since it cannot enjoy the economies of scale and internalize externalites. Similarly, they conclude that, while a centralized regulatory system for public good provision can internalize externalities, it would experience the coordination failure that causes costs for those whose preferences are not taken into account.

The other research area this study builds on is economics of fair cost allocation. When a large-scale network is held together by several stakeholders, the important thing is not only the development of a mechanism that can optimize the outcome of the whole network but also the fair allocation of costs among stakeholders. In the previous literature, various researchers in different fields including telecommunication (e.g., Skorin-Kapov (2001); Skorin-Kapov and Skorin-Kapov (2005)), transportation (e.g., Gelareh and Nickel (2011);

O'Kelly (1998); O'kelly (1987); Jaillet et al. (1996); Thomson, William (2007) and supply chain network (e.g., Bouchery et al. (2014); Vidal and Goetschalckx (2001)) have studied how a network system formed jointly by many agents can achieve a high level of fairness together with efficiency. In these studies, the researchers characterize network cost allocation problems, and argue that, in order to construct and operate the network, the involved agents should share the costs fairly. For example, in the field of civil aviation, O'kelly (1987) introduces the first mathematical formulation for an airline passenger network problem, and presents various alternative solutions for structuring an efficient airline network using a data set of air-passenger interactions from 25 U.S. cities. Thomson, William (2007) analyzes the problem of cost allocation in an airport network. In this study, the author proposes different approaches for exploring the problem of how several airlines should share the costs for using an airstrip in an airport network.

While the models developed in the above-mentioned fields are well suited to explain the problems and solutions for the public good provision and the fair cost allocation in various types of networks, they might not be applied directly to a security study. One of the main reasons is that these studies assume all players in the model work for maximizing social welfare. For example, in the literature on public good provision, the authors assume that objective of an agent who provide public goods is to maximize public goods surplus (Besley and Coate (2003)). In the studies on the network cost allocation, many authors consider a case where agents involved in a network cooperate in order to minimize the overall costs or to maximize the overall payoffs. However, many organizations are profit (or utility) maximizers rather than social welfare maximizers, and might behave based on private incentives rather than social incentives. Particularly, in the aviation industry, as noted in European Commission (2007), many airports are privately-owned or corporatised, and are likely to behave based on their private incentives that are biased by self-interest.

Another reason to be noted is that the previous studies do not consider a network adversary in the system. For example, the previous studies on the above-mentioned areas only consider agents who jointly work for achieving a higher level of efficiency in the system. However, in the security field, it is important to take into account adversarial agents, for example, terrorists and hackers, who cause a negative effect on the performance of the system since they affect the overall costs and payoffs.

As a result, in the field of security economics, the literature has developed a separate spectrum of discussions on the security provision by private agents and the effect of the interactions between attackers and defenders. For example, a series of studies written by Kunreuther and Heal (e.g., Kunreuther and Heal (2003); Heal and Kunreuther (2003)) analyzes the problems of a private agent's incentive to invest in security. On the other hand, Florêncio and Herley (2013), Cremonini and Nizovtsev (2009), Fultz and Grossklags (2009), Pym et al. (2013) and Ioannidis et al. (2013), for example, study the interactions between attackers and defenders in a security setting.

However, there has been relatively few applications that specifically focus on the economic theory of security provision and attacker-defender interaction to the field of aviation security. Heal and Kunreuther (2005) is one well known example that considers how the nature of security interdependence may affect the level of airline investment. Their work focuses on the relationship between airlines and other firms and discusses the negative effect of interdependence (i.e., negative externalities). However, the study does not address the interaction between attackers and defenders and fair cost allocation issues. It is therefore of interest to observe how the interaction and the cost allocation problem can be tackled in a setting of airport security. Adding these perspectives will provide an important insight into economics of airport security where these aspects are essential to be considered.

Particularly, in this study, we consider a situation where more than two agents (airports) operate jointly in the network. In order to maintain and guarantee security in the network, the agents are assumed to make security investments jointly. We particularly formulate a cost allocation problem associated with airports' security expenditures in civil aviation using a game-theoretic approach where a regulator enforces the level of security expenditures. Our model is different from the previous literature in the following two points. The first point is that, while the previous literature does not necessarily consider adversaries in the model, we take into account the role of attacker behaviour in analyzing airports' strategic investment decisions. We believe that, in studying cost allocation related to an attack on the aviation network, capturing the interaction among airports, attackers and a policy-maker is important since decisions made by each agent affect decisions of other agents. The second point is that we explicitly treat security interdependence among airports by associating the traffic volume between airports. We therefore consider a case where increasing the security level of one airport contribute toward the reduction of security risks in other airports.

3. THE AIRPORT MODEL

In this section, we use two theoretical tools, game theory and utility maximizing (or loss minimizing) models, to analyze a cost allocation problem for airport security expenditures. We use a game-theoretic model for formulating a cost allocation problem associated with airport security expenditures. Briefly speaking, game theory is the mathematical framework for exploring how agents' strategic interactions and decisions affect overall outcomes of the system (Morrow (1994)). A model based on game theory can predict how the agents in a game behave, based on their own preferences, in a conflict situation, and can analyze the decisions the agents made to maximize their payoffs. The benefit of using a game-theoretical model over traditional optimization models is that a game-theoretic model can handle problems for strategic decisions with multi-criteria decision-making for multiple players. Therefore, the model developed here will make it possible to understand how the agents in the aviation ecosystem behave in a conflict situation and to analyze the decisions the agents made to maximize their utility.

3.1 System Description

We model a series of strategic interaction among three classes of players: a group of identical attackers, heterogeneous airports and finally a single policy-maker. We particularly assume that airports in the network are at risk of a potential terrorist attack. We also consider that these airports are heterogeneous in nature, particularly, with their sizes. We refer to an airport which is classified into category i as type i airport ($i = 1, ..., n$) and assume that the number of airports in category i is N_i. Therefore, the total number of airports considered is denoted by $N_T = \sum_{i=1}^{n} N_i$. We define security expenditure made by type i airport as x_i, and the vector of security expenditures $(x_1, ..., x_n)$ as X.

As for attackers, we consider that all potential attackers are identical: they have the same characteristics for launching an attack on an airport. We refer N_A as the total number of attackers in the ecosystem and assume that each of the attackers launches only one attack. Furthermore, we use an assumption that total number of attackers N_A is smaller than the total number of airports N_T (i.e., $N_T > N_A$) and that all attacks are uniformly distributed over the N_T airports. As a result, the average number of attacks, denoted as η, on N_T airports is $\eta = N_A/N_T$. It should be noted that the model considers that η is determined endogenously through the strategic interaction with other players. By treating attacker dynamics endogenously, our model can provide a useful insight into the attackers' behaviour.

We define σ_i as the probability that one or more attacks mounted against type i airport are successful. It is assumed that the probability σ_i is conditional on the strategic decisions of the players. Specifically, we let $\sigma_i = \sigma_i(X, \eta)$ implying that the probability σ_i depends on X and η. For this study, σ_i is considered to have the following properties as discussed in Pym et al. (2013); Ioannidis et al. (2013):

Property 1: *$\partial\sigma_i/\partial\eta > 0$ for all i, which implies that an increase in the average number of attacks made against a target rises the probability of a successful attack;*

Property 2: *$\partial\sigma_i/\partial x_i < 0$ for all i, whicn means that an increase in the security expenditure of a target decreases the probability of a successful attack;*

Property 3: *$\partial^2\sigma_i/\partial x_i^2 > 0$ for all i, which implies that the effectiveness of an increase of security expenditure reduces (i.e., decreasing marginal returns to security expenditure).*

Property 4: *$\partial\sigma_i/\partial x_j \leq 0$ for all i and j, which shows a potential benefit of other airport's security expenditure on the target. It represents an ecosystem with positive externalities for security expenditure.*

Property 5: *$\partial\eta/\partial x_i < 0$ for all i, which implies that increased security expenditure of a target decreases the average number of attacks.*

There might be a wide array of functional forms that satisfies these properties (e.g., Gordon and Loeb (2002); Pym et al. (2013); Ioannidis et al. (2013);

Pym et al. (2014)). In this study, we adopted the functional forms used in Pym et al. (2014, 2013):

$$\sigma_i(X,\eta) = A_i e^{(-\alpha_i x_i - \sum_{j=1}^{n} \tau_{ij}\delta_{ij}x_j)}\eta^{\beta} \tag{1}$$

where α_i and β are positive constants, $\tau_{ij} \in [0,1]$, $\delta_{ij} \in [0,1]$ and $A_i \in (0,1]$.

The motivation for Eq. (1) is as follows. A_i is the probability that an attack on type i airport is successful in the absence of additional security expenditure by the airport. That is, it represents type i airport's risk from no additional security investment in a give period. α_i captures type i airport's marginal reduction in risk from additional security investment. Therefore, other things being constant, α_i represents the level of the reduction of the probability σ_i by the factor $1/e$. β measures inefficiency of attack and captures the marginal efficiency of of an additional attacker per target. Therefore, the increase in β reduces the chances of a successful attack if other things remain unchanged.

In order to capture externalities where the security actions of one airport decrease the risks or losses faced by other airports, we employ two parameters: δ_{ij} and τ_{ij}. δ_{ij} denotes interdependence coefficient which shows the degree of interdependence between type i and type j airports. A security level of one airport is the combined outputs of security effort of other airports as well as the airport itself. Interdependence coefficient therefore shows the extent to which the security level of a target airport type depends on the security level of other types of airports. τ_{ij} represents an actual structural characteristic of the relationship between different types of airport in the aviation ecosystem. In this study, we consider that τ_{ij} can be estimated by measuring a fraction of traffic volume between type i and j airports.

Since all potential attackers are assumed to be identical, we let each attacker have the same constant cost, C for mounting an attack. The cost C incurred for launching an attack includes the attacker's opportunity cost of the lost return from pursuing the next best option (Ioannidis et al. (2013); Pym et al. (2014, 2013)). It should be noted that, while all attackers are considered to be identical, their expected reward from a successful attack on different types of airports might differ. This is due to the fact that attackers might be able to achieve far higher rewards by successfully striking a large airport than a small airport. We therefore assume that the expected reward obtained from a successful attack is different by the types of airports and use R_i as the reward per attack against type i airport when one or more of these attacks turns out to be successful.

We consider that attackers wish to maximize their expected profit. The expected profit which an attacker obtains from mounting an attack on type i airport is given by the following expression.

$$\sigma_i(X,\eta)R_i - C. \tag{2}$$

Attackers are likely to be motivated to mount attacks on the population of target airports as long as the cost of mounting an attack is lower than the expected reward from mounting an attack (Ioannidis et al. (2013); Pym et al. (2014, 2013)). This implies that the equilibrium number of attacks per target, η^*, should meet the following condition:

$$\sum_{i=1}^{n} \sigma_i(X,\eta^*)R_i\frac{N_i}{N_A} = C. \tag{3}$$

While the left-hand side of Eq. (3) shows an attacker's expected reward from an attack made against the polulation of target airports, the right-hand side of the equation is an attacker's cost of launching an attack. This equation ensures that, in equilibrium, more attacks will be launched as long as the expected reward from an attack exceeds the cost of the attack. By reorganizing and rewriting Eq. (3), the following equation can be identified:

$$\sum_{i=1}^{n} \sigma_i(X,\eta^*)\rho_i f_i = \eta^*, \tag{4}$$

where f_i is the fraction of type i airports and ρ_i is reward/cost ratio of an attack on type i airports. It should be noted that the equilibrium level of attackers per target, η^*, satisfying Eq. (4) depends in general on the vector of security expenditures by the N_T airports. The dependence of $\eta^*(X)$ on the level of security expenditures chosen by the various airports has important implications for public policy since this implies that appropriate security expenditures can actually deter an attack (Pym et al. (2014, 2013)).

As for airports, they are assumed to be risk neutral. We let L_i represent the expected loss suffered by type i airport when one or more successful attacks on the airport occurs. Without loss of generality, we suppose that the magnitude of L_i does not rely on the total number of successful attacks but only on whether there is a successful attack. For instance, we can think of a case where the factor allowed a successful attack is removed after

Likelihood:	Frequent	Probable	Occational	Remote	Not credible
Skills	No limitations	Engineering knowledge	Specialist knowledge	Expert knowledge	Inside information
Means	No limitations	Publicly available	Available with difficulty	Hard to obtain	Extremely scarce
Opportunity	Always	Frequently	Regularly	Seldom	Never
Profit	Large	Significant	Fair	Little	None
Attention	World-wide media attention	Regional media attention	Fair attention of local media	Little attention of local media	No media attention
Impunity	No chance of punishment	Little chance of punishment	Fair chance of punishment	High chance of punishment	Certainty of punishment
Detection	Not possible to predict or detect	Detection due to 'chance'	Fair chance of detection	High chance of detection	Certainty of detection

Table I .: Examples of Attacker's likelihood of an attack

the successful attack occurs. Therefore, an airports of type i will select its level of security expenditure x_i by minimizing the expected loss

$$\sigma_i(X, \eta^*)L_i + x_i. \tag{5}$$

Eq. (5) contains both the expected damage from a successful attack, $\sigma_i L_i$, and the security cost, x_i, which type i airport should pay whether or not there is a successful attack. It should be mentioned that, similarly with the equilibrium number of attacks against airports, η^*, which depends on the the entire vector of security expenditures X, we also assume that the security expenditures of other airports can potentially affect the expected loss of type i airport through σ_i. Therefore, Eq. (5) takes into account ecosystem externalties. In Table II we present the set of structural parameters underlying our model.

Table II .: Description of Model Parameters and Choices.

Defender's parameters	
x_i	Type i airport's security investment.
α_i	Type i airport's marginal risk reduction.
A_i	Type i airport's zero investment risk.
L_i	Type i airport's assets at risk
Attacker's parameters	
η	Attacker's attack intensity.
β	Elasticity of attacking intensity on an airport.
ρ_i	Reward/cost ratio for attacks on type i airports.
Policy-maker's parameter	
v_i	Social planner's weight for Type i airport.
δ_{ij}	Interdependence coefficient between type i and j airports.
Environmental parameters	
f_i	Fraction of type i airports.
τ_{ij}	Fraction of traffic volume between types i and j airports.

3.2 Modeling Non-cooperative Nash Equilibrium

We first model the problem of a Nash equilibrium of the game between attackers and airports: a strategy of an attacker is a choice whether or not to launch an attack on the target population based on the condition described in Eq. (4), and a strategy of type i airport is a choice of security expenditure, x_i. In this game, there is a strategic interaction between the choices of attackers and airports. The expected payoff for an attacker is affected, in part, by the choices of airports' strategies on security expenditures. Similarly, the expected loss for an airport is determined partly by the choices of attack participation of attackers. As a result, in a Nash equilibrium, the strategies of both parties should be optimal given the expectations about the strategies chosen by other parties, and these expectations have to be correct when all of them behave optimally.

One important assumption we made for a Nash equilibrium is that airports do not consider ecosystem externalities. As indicated in Pym et al. (2013), while the externality effect of security action might be an socially beneficial byproduct, an individual decision-maker is likely to underestimate its value when considering the costs and the benefits of security action. As a result, without social coordination, individual agents might choose the level of security expenditures based only on their private incentives. In our model, this implies that the interdependence coefficient τ_{ij} equals to 0 for all i and j. From Eqs. (1) and (5), therefore, the expected loss from attacks can be given as:

$$V_i = A_i e^{-\alpha_i x_i} \eta^\beta L_i + x_i. \tag{6}$$

Since the objective of type i airport in the presence of exogenous η is to find an optimal level of security expenditures minimizing the expected loss from an attack. This can be given by:

$$x_i^* = \operatorname*{argmin}_{x_i} A_i e^{-\alpha_i x_i} \eta^\beta L_i + x_i. \tag{7}$$

Differentiating Eq. (7) with respect to x_i and setting it equal to zero yields

$$A_i e^{-\alpha_i x_i} \eta^\beta L_i \alpha_i = 1. \tag{8}$$

Therefore, for a given η, x^* has the following analytic solution:

$$x_i^* = \frac{\log\left(A_i \eta^\beta L_i \alpha_i\right)}{\alpha_i}. \tag{9}$$

Eq. 9 indicates that, as A_i, L_i and η increase, the equilibrium level of security expenditure rises.

We now assume that attackers are risk neutral and make rational choices to participate in attacks. From Eqs. (1) and (4), we can find the equilibrum level of attacker per target to be

$$\eta^* = \sum_{i=1}^{n} A_i e^{-\alpha_i x_i} \eta^\beta \rho_i f_i \tag{10}$$

Solving Eqs. (8) and (10) simultaneously, we can obtain the following proposition.

Proposition 1: **(Existence of Nash equilibria)** Under the preceding assumptions, the Nash equilibrium for targets and attackers can be denoted as:

$$x_i^* = -\frac{\log\left(\frac{\left(\sum_{i=1}^{n} \frac{f_i \rho_i}{\alpha_i L_i}\right)^{-\beta}}{\alpha_i A_i L_i}\right)}{\alpha_i} \tag{11}$$

$$\eta^* = \sum_{i=1}^{n} \frac{f_i \rho_i}{\alpha_i L_i} \tag{12}$$

3.3 Modeling Socially Optimal Security Expenditure

In this section, we consider a Stackelberg policy-maker who desires to minimize a weighted average of the expected losses suffered by the population of airports. A policy-maker can be regarded as a regulator or a law maker who can influence the security expenditures of airports. Since it is important for him to consider socially desirable ecosystem conditions, we assume that the policy-maker takes into account an externality effect of security expenditures between airports. We particularly consider positive exteranlities: an increase in the security expenditure of an airport has a positive effect on other airports' security levels.

We now consider that the policy maker sets the vector of the levels of security expenditures for all types of airports to minimize the following weighted average of the targets' expected losses.

$$V = \sum_{i=1}^{n} v_i f_i [A_i e^{(-\alpha_i x_i - \sum_{j=1}^{n} \tau_{ij} \delta_{ij} x_j)} \eta^\beta L_i + x_i] \tag{13}$$

where v_i are positive weights indicating how much importance the policy-maker places on the expected loss of type i airports. As can be seen, τ_{ij} and δ_{ij} are used to take into account interdependence. In the presence of an exogenous η, type i airport minimizes losses with respect to x_i:

$$x_i^* = \operatorname*{argmin}_{x_i} \sum_{i=1}^{n} v_i f_i [A_i e^{(-\alpha_i x_i - \sum_{j=1}^{n} \tau_{ij} \delta_{ij} x_j)} \eta^\beta L_i + x_i]. \tag{14}$$

We suppose that the choice of X which minimizes the objective in Eq. (14) satisfies the usual first-order conditions for an optimum. These first-order condition with respect to x_i can be given as:

$$\begin{aligned} & A_i L_i \alpha_i \eta^\beta \left[e^{(-\alpha_i x_i - \sum_j \tau_{ij} \delta_{ij} x_j)}\right] \qquad (15) \\ + & \sum_j \left[A_j L_j \eta^\beta \tau_{ij} \delta_{ij} e^{(-\alpha_i x_i - \sum_j \tau_{ij} \delta_{ij} x_j)}\right] = 1 \end{aligned}$$

As for the attacker intensity, the following equation can be found.

$$\eta^* = \sum_{i=1}^{n} \left[f_i \rho_i A_i e^{(-\alpha_i x_i - \sum_{j=1}^{n} \tau_{ij} \delta_{ij} x_j)}\right]^{\left(\frac{1}{1-\beta}\right)}. \tag{16}$$

As with many other previous studies(e.g., Baldwin and Krugman (2004); Calzolari and Lambertini (2007)), the characterization of socially optimal security expenditure are not analytically solvable, and this leads to numerical simulations in the next section.

4. NUMERICAL SIMULATIONS

We now investigate quantitatively how the policy-maker's decision on the levels of security expenditures affects the aviation security ecosystem. However, a generalized and comprehensive understanding must take precedence in order to make quantification of the parameters used in the model. Therefore, we first explore the features related to airport security and then present how the model illustrated in the previous section is parameterized to replicate certain features of European airports. Finally, we discuss the methodology used in the numerical simulation followed by data, assumptions and cases from European airports. The main motivation of this subsection is to observe different outcomes of policies for airport security expenditures where externalities of security are present. The model considers Nash equilibrium security expenditures as a status quo and calculates the impact of a government security policy and altering variables in the equilibrium such as the level of interdependence between different types of airports.

4.1 Airport Security and Financing

Since the beginning of modern aviation it was obvious that airports and airplanes have provided unique opportunities for various types of attackers including hijackers and terrorists. Even a very small explosion can paralyze an airport operation or panic passengers in an airborne aircraft. In a crowded airport or airplane, a single attacker can threat and harm a huge number of people with a small firearm. Since the main objective of airport security is to prevent any illegal and dangerous activities including terrorism, there are a number of different methods and procedures in airports associated with security provision. According to a report published by ACI Europe (2003), while responsible body for security service provision differs from country to country, the main measures for airport security can be summarized as shown in Table III .

In the aftermath of September 11, 2001, the general public's concerns on airport security has put greater pressure on airliners, airport operators and public authorities to improve security methods in airports. As a result, in Europe, based on the recognition that addressing aviation security at a national level is ineffective and efficient, building a harmonized structure for civil aviation security has been a main focus in designing airport security regulations. Therefore, a regulation containing standardized security rules and procedures that must be complied by European airports came into effect (i.e., Regulation (EC) No 2320/2002). Furthermore, various entities have tried to develop risk management tools. For instance, as displayed in Table I and IV , Eurocontrol have published a series of reports (e.g., Eurocontrol (2010)) that summarizes the examples an attacker's likelihood of an attack and likelihood of successful attack.

Checks on access of staff to restricted areas	Reliability check on applicants for obtaining badge
Badge regime	Checks on passengers and hand baggage
Baggage reconciliation	Checks on hold baggage
Checks on cargo/airmail	Armed protection land-side
Armed protection airside	Protection on parked aircraft
Video supervision	

Table III .: Main security measures used in airports

Success likelihood	Physical	People	Electronic
High	Physical access possible	Can introduce or engineer staff	Normal function or known vulnerability
Medium	Physical barriers in depth	Access control, staff checking & training	Well isolated & access controlled
Low	Protection + inspection & audit	Include separation polices & audit	Internal barriers, regular assessment

Table IV .: Examples of likelihood of a successful attack

The idea of harmonization in civil aviation security was further advanced by revising and elaborating Regulation (EC) No 2320/2002 (i.e., Regulation (EC) No 300/2008). One of the main objectives of this revised regulation was to make the aviation security regulation more flexible and up-to-date against terrorists' technologies. In addition, the regulation employed the concept of 'one-stop security' which can remove a re-screening procedure for transfer passengers arriving from non-EU countries (Falconer (2008)).

Therefore, it can be said that Europe has a common view on what to be implemented as security rules and measures, but not on who should pay for security (Falconer (2008)). For example, ACI Europe argues that the governments should fund civil aviation security since a terrorist's attack commonly targets states and security provision by individual airports might distorts competition (e.g., ACI Europe (2010, 2009a)). However, according to a report by Irish Aviation Authority & Aviasolutions (2004), there has been inconsistent security funding schemes in various European countries. As shown in Table V , while some European countries levy a security tax in order to fund security costs burne by the government, other countries make an airport pay for security through security charges levied from passengers. The report further categorized countries by the responsibility of security activities. In detail, Table V illustrates that, whereas state security taxes with the centralised model is a commonly used system for security financing, many countries employ diversified and mixed approaches.

Provision of security activities	Centralised model	Decentralised model
	Austria, Finland, Germany, Iceland, Italy, Luxembourg, Norway, Portugal, Spain, Sweden, Switzerland	Belgium, Denmark, France, Greece, Ireland, UK
Countries charging state security taxes	Austria, Germany, Iceland, Italy, Netherlands, Portugal, Spain	Belgium, France
Countries charging airport security charges	Luxembourg, Sweden, Switzerland, Germany, Netherlands	Belgium, France, Greece, Iceland, UK

Table V .: Structure of airport security provision in European countries

The report also indicated that, even if the details of a security financing system differ from country to country, passengers are the main funder of airport security and most countries adopted rather simple flat rate levied on a per-passenger basis (Irish Aviation Authority & Aviasolutions (2004)). No matter what security financing scheme is employed by a country, one of the important question, given that huge sums of monies are being levied on passengers and being invested on airport security, is whether these monies are fairly allocated among airports. In the next subsection, we discuss how our simulation is conducted to illustrate security cost allocation among airports, and show whether the current regulatory system for aviation security financing used in Europe is an effective and efficient solution for security threats.

4.2 Parameter Calibration

In order to set the parameter values for numerical simulation, we estimate some parameters using actual data from various sources including an airport's annual report and an industry report. We also derive some of the parameter values from formal and informal interviews with various stakeholders.

Table VI provides a full list of the numerical values to parametrize the model. When we interviewed various stakeholders, they stated that airports are commonly categorized into small, medium and large airports. There is no universal criterion for classifying airport sizes and no universal definition for small, medium and large airports. For example, while U.S. Department of Transportation uses the total paved runway area to classify airports (Federal Aviation Administration (2003)), U.S. Congress uses passenger enplanements for classification (U.S. Congress (1984)). Classification using these criteria has been somewhat arbitrary.

In this study, we categorize airport size by considering both the hub and spoke network designs and the airports' traffic volumes since these are the widely used mechanisms for measuring airport size. We consider large airports as hubs with a lot of destinations and an uneven number of aircraft routed to them. Medium airports are assumed to be airports serving for large hubs but working also as hubs for small airports. Small airports are considered to be outlying airports with very low traffic volume. Using the data of total 509 European airports derived from Vitali et al. (201q), we sort airports by the total number of outbound flights per day. Since slight changes in classification criteria might shift conclusions of the analysis, from a series of discussions with stakeholders, we carefully classify the airports with the assumption that 3% of the airports are large airports, 10% are medium airports and the rest airports are small airports. Consequently, 15 airports are categorized into large airports, 50 airports are categorized into medium

Table VI .: Values of Model Parameters.

Parameter	Value	Parameter	Value
f_1	15/509	v_1, v_2, v_3	1
f_2	50/509	δ_{ij}	0.001
f_3	444/509	τ_{11}	0.10104
A_1	0.375	τ_{12}	0.3675
A_2	0.169	τ_{13}	0.094792
A_3	0.025	τ_{22}	0.23828
L_1	€140M	τ_{23}	0.17618
L_2	€17.5M	τ_{33}	0.022203
L_3	€7M	ρ_1	1.1765
α_1	0.071	ρ_2	1.1765
α_2	0.766	ρ_3	1.1765
α_3	2.786	β	0.1

airports, and 444 airports are classified into small airports.

Following the notations used in the previous section, we refer $i = 1$ as large airports, $i = 2$ as medium airports, and $i = 3$ as small airports. The fractions of each type of airports, therefore, are $f_1 = 15/509$, $f_2 = 50/509$ and $f_3 = 444/509$. From the data, we also identify that 15 airports classified as large airports have average 400 outbound flights per day, 50 medium airports have average 50 outbound flights per day, and 444 small airports have average 20 outbound flights per day. In addition, we assume that there are 37M passengers per year for a large airport (i.e., Munich airport, Germany), 2.7M passengers per year for a medium airport (i.e., Verona airport, Italy) and 0.54M passengers for a small airport (i.e., Falconara airport, Italy), repectively.

Loss from a successful attack for each type of airports, L_i is calculated based on the numbers of days of potential airport shutdown and flights grounded from a successful attack. According to International Air Transport Association (IATA), Hurricane Sandy caused €50K loss per canceled flight for seven days (International Air Transport Association (2012)). Other studies for Eyjafallajokulls volcanic ash plume in Iceland also indicate the closure of most of Europe's airspace over a period of seven days (Brooker (2010); Mazzocchi et al. (2010)). We therefore use these as our proxy values for losses from a successful attack. For example, losses incurred from a successful attack are 7 days * 400 flights grounded * €50k for large airports (i.e., €140M), 7 days * 50 flights grounded * €50k for medium airports (i.e., €17.5M), and 7 days * 20 flights grounded * €50k for small airports (i.e., €7M). It should be noted that in reality the losses could be much higher than the values used here since a successful terrorist attack can cause huge casualty as well as damage on economy and society.

Fractions of traffic volume between two types of airports, τ_{ij}, are also calculated from the actual traffic data derived from Vitali et al. (201q). We particularly used the following formula to calculate τ_{ij}:

$$\tau_{ij} = \frac{q_{ij} + q_{ji}}{\sum_i \sum_j q_{ij}} \tag{17}$$

where q_{ij} denotes total number of outbound traffic from type i airports to type j airports.

As for the policy-maker's weight for type i airports, v_i, we set all of the weights equal to one since, in a series of conversations with policy-makers, we learned that they treat all airports equally.

4.3 Results of Simulation

While the calibration of the above mentioned parameters are possible using the existing data and evidence, there are also parameters with no available data and information for calibration. Whenever actual data for a parameter value is not available, we try to employ values from previous research or derive values by making reasonable guesses. Once the solution using the parameter values is found, we calculate the change in security expenditures in equilibrium as the one of the parameters changes.

Particularly, there is no available data for zero investment risk for type i airports, A_i, and therefore we decide to estimate A_i based on our reasonable assumption. In detail, we define A_i as

$$P(\text{selected}_i)P(\text{successful}_i|\text{selected}_i),$$

where $P(\text{selected}_i)$ represents the probability that type i airports are selected by an attacker when the airports do not make any additional investment in security, and $P(\text{successful}_i|\text{selected}_i)$ is the probability that type i airports are successfully attacked when selected. As for $P(\text{selected}_i)$, we estimate it by calculating $(1/N_i)/(\sum_1^n 1/N_i)$. For example, as for large airports, $P(\text{selected})$ is $(1/15)/(1/15+1/50+1/444)$. $P(\text{successful}_i|\text{selected}_i)$ is more problematic to estimate. However, in a conversation with stakeholders, we got an impression that larger airports might have a lower value for this probability since they already have various security measures in place. We therefore arbitrary assume that $P(\text{successful}_1|\text{selected}_1) =$

0.5, $P(\text{successful}_2|\text{selected}_2) = 0.75$ and $P(\text{successful}_3|\text{selected}_3) = 1.0$. Consequently, the estimated values are $A_1 = 0.375$, $A_2 = 0.169$ and $A_3 = 0.025$.

Another parameter difficult to calibrate is type i airport's marginal reduction in risk from security expenditure, α_i. As shown in Figure 1, as α_i increases, the probability of a successful attack on type i airports decreases. In order to identify values for α_i, we assume that the global target of the policy-maker is to have one serious incident every d_i days for each target i. Therefore, $d_i I_i \sigma_i = 1$ where I_i is the number of inbound flights per day. From Eq. (1), we have $d_i I_i A_i e^{-\alpha_i x_i} = 1$. It can be rewritten as $\log d_i + \log I_i + \log A_i - \alpha_i x_i = 0$. This gives the value of α_i as:

$$\alpha_i = \frac{\log d_i + \log I_i + \log A_i}{x_i}. \tag{18}$$

Since the interval for security is the same $d_i = d_j = 365 * Y$ where Y is the number of years without incidents we get

$$\alpha_i = \frac{5.9 \log Y + \log I_i + \log A_i}{x_i}. \tag{19}$$

From the review of annual reports of various airports and interviews we have identified that x_i is between €5 to 7, and hence used the average value, €6. While it is not possible to estimate Y directly, it might be rational to consider that Y is reasonably long. We therefore assume that $Y = 10$. From Eq. (18), we therefore get $\alpha_1 = 0.071$, $\alpha_2 = 0.766$ and $\alpha_3 = 2.786$.

As for the parameter of the policy-maker, we assume that interdependence coefficients, δ_{ij}, which ranges from 0 to 1, are all assumed to be relatively small, 0.001. This is due to the fact that one-stop security has not yet been fully implemented in all European countries.

Calibration of parameter values for attackers is also difficult. To obtain a point estimates of reward/cost ratios for attacks on type i airports, ρ_i, are adopted from Pym et al. (2014). According to them, the cost/reward ratio for cyber-attackers is 0.1 which makes the reward-cost ratio 10. In the age of telecommunication and the Internet, a threat caused by an attacker can draw nationwide, or even worldwide attention. This implies that the attacker might gain relatively high rewards. As a result, we believe the reward-cost ratio 10 is reasonable.[5] We further assume the reward/cost ratios for all types of airports are same.

The most difficult parameter value to be estimated for the simulation is attacker efficiency loss, β, as no information exists. In order to understand this parameter, we generate Figure 2 for an illustrative purpose. As shown in the figure, as attack inefficiency increases (i.e., as β increases), the probability of a successful attack decreases. We use 0.1 for the value for β since it is believed that attackers' efficiency is relatively due to the support from big organized support groups.

From the above parameter values, we can investigate our main question on whether security expenditures controlled by the policy-maker is fair for airports with different nature. Our simulation model provides an overview of the intuition of this question. While the simulation might not be appropriate for specific quantification for real aviation security environment, we have tried to stay close to real data and believe that it can provide a useful insight into the current situation.

For our starting numerical example, we calculate point estimate of security expenditure of each type of airports for Nash equilibrium and social optimum as shown in Table VII . The table shows that, when the interdependence between airports is relatively small (i.e., $\delta_{ij} = 0.001$), the government's regulation in order to minimize expected social losses makes medium and large airports invest in security relatively close to Nash equilibrium security expenditures. However, as interdependence increases, the regulation makes medium and large airport underinvest in security, and small airports overinvest in security compared to Nash equilibrium security expenditures. This implies that medium and large airports can get benefits from the rule whereas small airports take great costs. As a result, security interdependence makes small airports carry a security burden of medium and large airports.

As an extension, we further investigate how the changes in the degree of interdependence between specific types of airports affect airports' security expenditure. For example, we can think of an one-stop security check solution whereby passengers and their baggage does not need to be re-screened at a connecting airport (commonly, medium and large airports) if they had gone through the security check adequately at the airport of origin. We assume that the policy-maker wants to distribute security burden fairly to airports by adjusting

[5] Note that the general results do not change even if it is reduced to a small number, say 1.

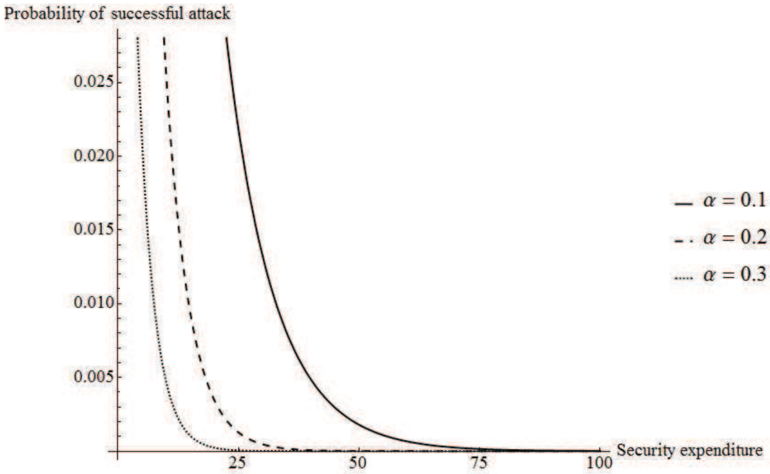


Fig. 1: Probability of a successful attack as a function of security expenditure. Note that, as α increases, the probability of a successful attack decreases.

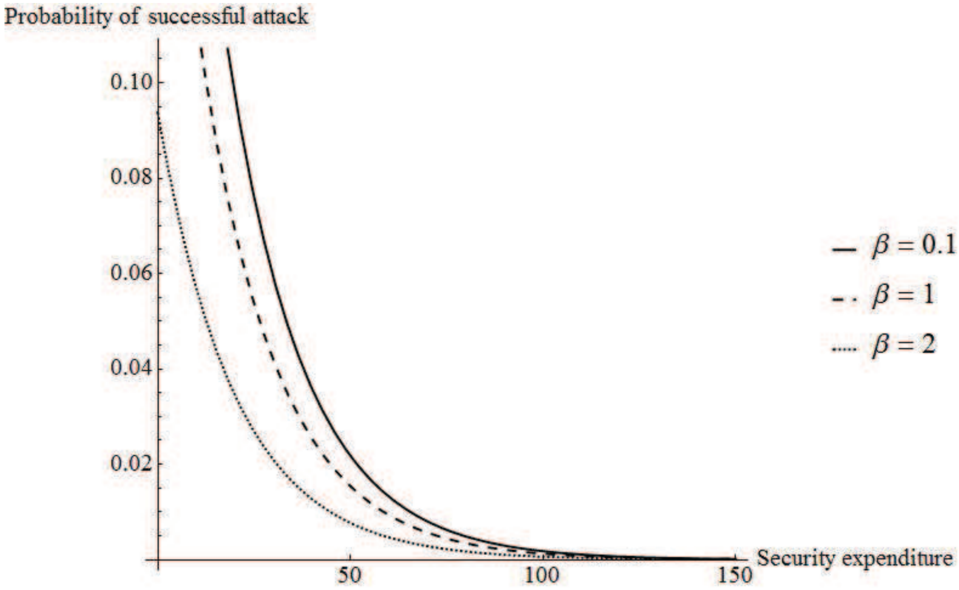


Fig. 2: Probability of a successful attack as a function of security expenditure. An important point to note is that, as an attacker becomes more inefficient, the probability of a successful attack decreases.

security interdependence between two different types of airports.

Figures 3 to 5 present Nash equilibrium and socially optimal security expenditures per passenger when interdependence changes. As displayed in Figure 3, if the policy-maker enacts a regulation that increases interdependence between large and medium airports, the unfairness in security expenditures becomes more severe since mandated security expenditure for medium and large airports gets much less than Nash equilibrium whereas small airports are not affected by the regulation. This implies that the ones taking the benefits from the regulation are medium and large airports.

Figure 4 illustrates the effect of the increase in the interdependence between large and small airports. In this case, as the interdependence increases, mandated security activities for both

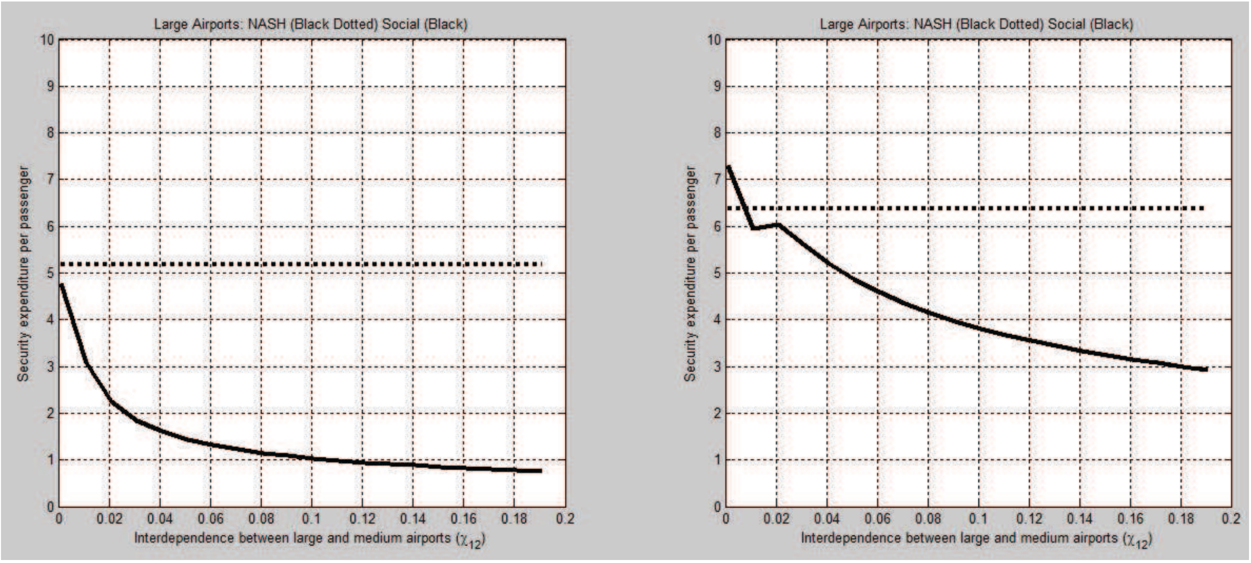


Fig. 3: Nash equilibrium and social optimum security expenditures per passenger for large and medium airports when δ_{12} increases. The dotted line represents the Nash equilibrium security expenditure per passenger and the solid line is the socially optimal security expenditure per passenger. The value of δ_{12} increase from 0.001 to 0.2.

Table VII .: Nash and socially optimal security expenditures with different values for interdependence coefficients.

	Total Security Expenditure		*Security Expenditure per Passenger*	
Type	Nash Equilibrium	Social Optimum	Nash Equilibrium	Social Optimum
$\delta_{ij} = 0.001$				
Large	€189M	€179M	**€5.1**	€4.8
Medium	€17M	€20M	€6.3	**€7.4**
Small	€4M	€5M	€7.6	**€9.9**
$\delta_{ij} = 0.01$				
Large	€189M	€83M	**€5.1**	€2.2
Medium	€17M	€15M	**€6.3**	€5.5
Small	€4M	€5M	€7.6	**€9.2**
$\delta_{ij} = 0.1$				
Large	€189M	€28M	**€5.1**	€0.8
Medium	€17M	€8M	**€6.3**	€3.0
Small	€4M	€4M	€7.6	**€8.1**

large and small airports become lower than Nash equilibrium, but medium airports are not affected (i.e., their mandatory spending is higher than the Nash equilibrium). This means that medium airports need to carry a security burden of small and large airports.

If the policy-maker increases the interdependence between medium and small airports as shown in Figure 5, the policy-maker can reduce the gap between Nash equilibrium and social optimum security expenditures for medium and small airports, while he can let large airports make security effort relatively close to Nash equilibrium. This implies that increase of interdependence between medium

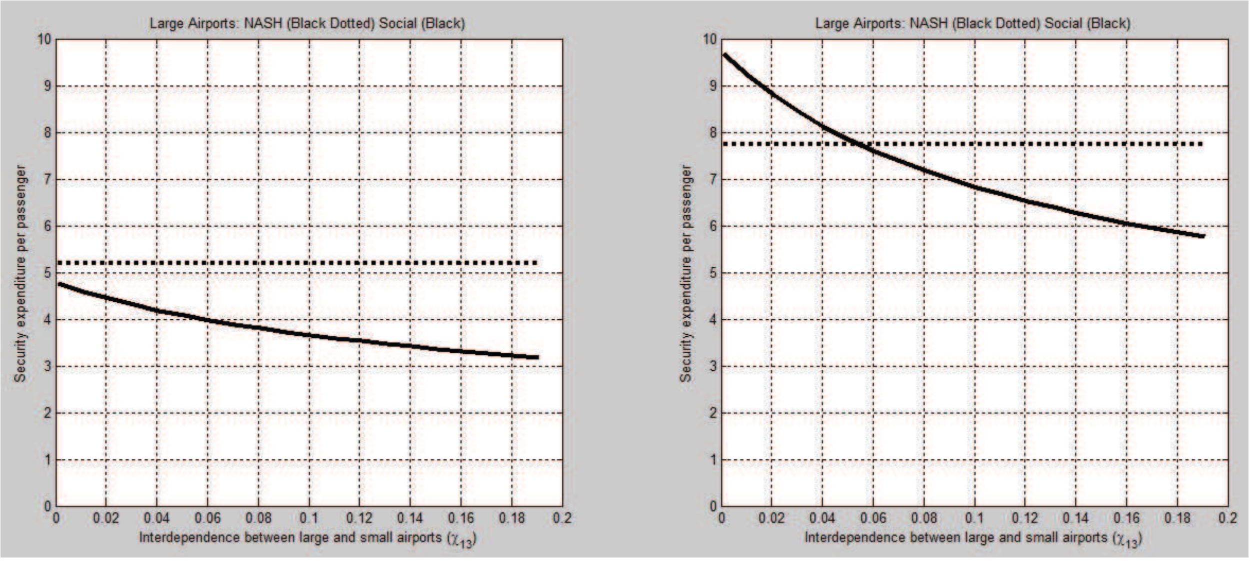


Fig. 4: Nash equilibrium and social optimum security expenditures per passenger for large and small airports when δ_{13} increases. The dotted line represents the Nash equilibrium security expenditure per passenger and the solid line is the socially optimal security expenditure per passenger. The value of δ_{13} increase from 0.001 to 0.2.

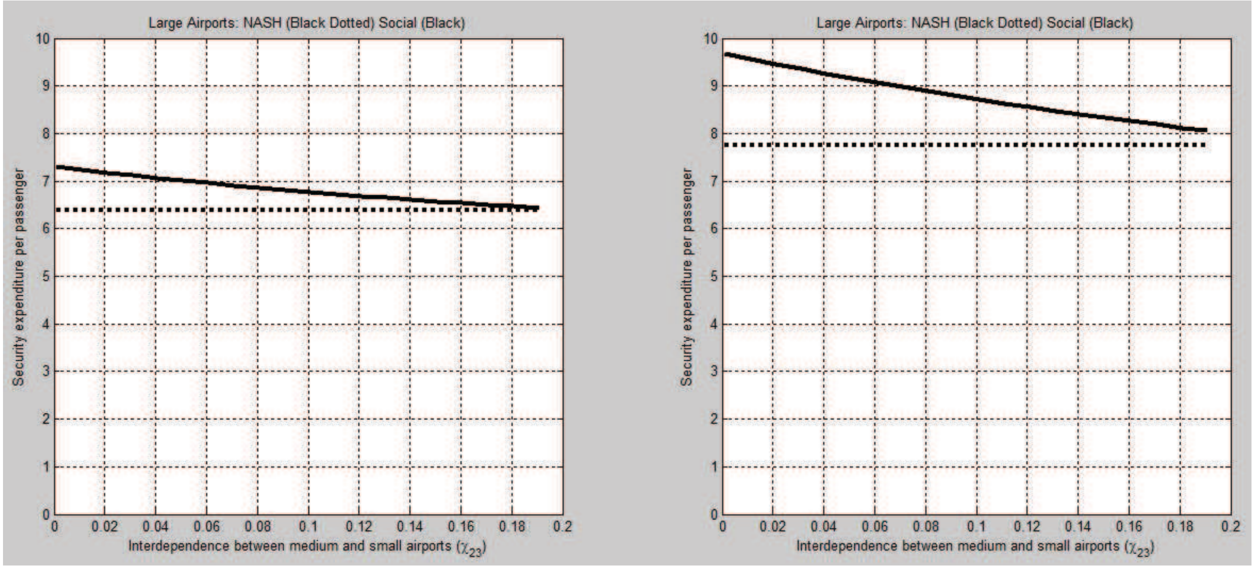


Fig. 5: Nash equilibrium and social optimum security expenditures per passenger for medium and small airports when δ_{23} increases. The dotted line represents the Nash equilibrium security expenditure per passenger and the solid line is the socially optimal security expenditure per passenger. The value of δ_{23} increase from 0.001 to 0.2.

and small airports would be the good way to ensure the fairness of security activities among airports with different.

5. SUMMARY

This study offers a contribution to the ongoing discussion on the fairness of public policy on security expenditures of airports with different nature. The paper uses a game-theoretic model in which there is a strategic interaction between players. In this study, we show that the security expenditures chosen by a social planner might not be fairly allocated within different types of airports, and might require a rational for better adjustment for a social planner.

In detail, we show that, while the divergence between private and social incentives for security expenditures suggests the rationale for regulatory rules for security expenditures, it does not guarantee the fairness of such rules. In the simulation, we illustrate that a social planner might ask small airports spend more on security than medium and large airports. If passenger went through security checking in a small airport, they do not need to go over the security check again after they arrive in a large airport. As a result, if the social planner distribute the same amount of money per passenger to all airports, the large and medium airports might actually make profits, because they are using small airports to conduct security procedures for them.

We further identify that, as interdependence of all types of airports increases, lager airports essentially save more and more money. This implies that the more the airports are interconnected, the more the larger airports can save money compared to small airports. Furthermore, the simulation results show that the increase in the interdependence between medium and small airports would be the most efficient way to reduce unfairness in the regulation on security expenditures for airports.

As a future study, it would be useful to investigate a case where the interdependence coefficients are different. For example, it would be interesting to analyze the situation where small airports are feeders for security.

ACKNOWLEDGEMENTS

This work has been partly funded by the European Union's 7th Framework Programme under grant agreement no 285223 - SECONOMICS (www.seconomics.org).

REFERENCES

ACI Europe (2003). Financing civil aviation security costs in europe. ACI Europe.

ACI Europe (2009a). Aci europe position on the proposal for a directive on security charges (com (2009) 217). ACI Europe.

ACI Europe (2009b). Aci europe position on the proposal for a directive on security charges (com (2009) 217). Airports Council International Europe.

ACI Europe (2010). A level playing field for european airports - the need for revised guidelines on state aid. ACI Europe.

ACI Europe (2012). Revision of the 2005 ec guidelines on financing of airports and start-up aids to airlines departing from regional airports. Airports Council International Europe.

Baldwin, R. E. and P. Krugman (2004). Agglomeration, integration and tax harmonisation. *European Economic Review 48*(1), 1–23.

Barber, R. (2001). Hackers profiled who are they and what are their motivations? *Computer Fraud & Security 2001*(2), 14 – 17.

Besley, T. and S. Coate (2003). Centralized versus decentralized provision of local public goods: a political economy approach. *Journal of public economics 87*(12), 2611–2637.

Biennier, F. and J. Favrel (2005). Collaborative business and data privacy: Toward a cyber-control? *Computers in Industry 56*(4), 361 – 370. ¡ce:title¿The Digital Factory: An Instrument of the Present and the Future¡/ce:title¿.

Bier, V. M., N. Haphuriwat, J. Menoyo, R. Zimmerman, and A. M. Culpen (2008). Optimal resource allocation for defense of targets based on differing measures of attractiveness. *Risk Analysis 28*(3), 763–770.

Botsis, T., A. M. Lai, W. Palmas, J. B. Starren, G. Hartvigsen, and G. Hripcsak (2012). Proof of concept for the role of glycemic control in the early detection of infections in diabetics. *Health Informatics Journal 18*(1), 26–35.

Bouchery, Y., S. Fazi, and J. C. Fransoo (2014). Hinterland transportation in container supply chains. In C.-Y. Lee and Q. Meng (Eds.), *Handbook of Ocean Container Transport Logistics*, Volume 220 of *International Series in Operations Research & Management Science*, pp. 497–520. Springer International Publishing.

Bozorgi, M., L. K. Saul, S. Savage, and G. M. Voelker (2010). Beyond heuristics: learning to classify vulnerabilities and predict exploits. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '10, New York, NY, USA, pp. 105–114. ACM.

Bradbury, D. (2004). Documentation dearth undermines open source security. *Infosecurity Today 1*(5), 6 –.

Bratus, S. (2007). Hacker curriculum : How hackers learn networking. *Distributed Systems Online, IEEE 8*(10), 2–2.

Brooker, P. (2010). Fear in a handful of dust: aviation and the icelandic volcano. *Significance 7*(3), 112–115.

Calzolari, G. and L. Lambertini (2007). Export restraints in a model of trade with capital accumulation. *Journal of Economic Dynamics and Control 31*(12), 3822–3842.

Campbell, K., L. A. Gordon, M. P. Loeb, and L. Zhou (2003). The economic cost of publicly announced information security breaches: empirical evidence from the stock market. *Journal of Computer Security 11*(3), 431–448.

Carney, M. and K. Mew (2003). Airport governance reform: a strategic management perspective. *Journal of Air Transport Management 9*(4), 221 – 232.

Cavusoglu, H., H. Cavusoglu, and J. Zhang (2008). Security patch management: Share the burden or share the damage? *Management Science 54*(4), 657–670.

Chow, J., J. Chiesa, P. Dreyer, M. Eisman, T. W. Karasik, J. Kvitky, S. Lingel, D. Ochmanek, and C. Shirley (2005). Protecting commercial aviation against the shoulder-fired missile threat. Technical report, RAND Corporation.

Cohen, M. A. (1987). Optimal enforcement strategy to prevent oil spills: an application of a principal-agent model with moral hazard. *Journal of Law and Economics 30*(1), 23–51.

Collie, M. P. (1988). The legislature and distributive policy making in formal perspective. *Legislative Studies Quarterly*, 427–458.

Cremonini, M. and D. Nizovtsev (2009). Risks and benefits of signaling information system characteristics to strategic attackers. *Journal of Management Information Systems 26*(3), 241–274.

Dur, R. and H. Roelfsema (2005). Why does centralisation fail to internalise policy externalities? *Public Choice 122*(3-4), 395–416.

Eurocontrol (2010). Atm security risk management toolkit. Eurocontrol.

European Commission (2007). Directive of the european parliament and of the council on airport charges. European Commission.

European Commission (2009). Report from the commission on financing aviation security. European Commission.

Eurostat (2013). Nearly 830 million air passengers in 2012. Eurostat, the statistical office of the European Union.

Evers, J. (2005). Hacking for dollars. *Cnet news.com.*

Falconer, R. (2008). Revised eu regulatory framework for aviation security agreed. *Airport Business*.

Federal Aviation Administration (2003). Buildings for storage and maintenance of airport snow and ice control equipment and materials. U.S. Department of Transportation.

Florêncio, D. and C. Herley (2013). Where do all the attacks go? In *Economics of Information Security and Privacy III*, pp. 13–33. Springer.

Fultz, N. and J. Grossklags (2009). Blue versus red: Towards a model of distributed security attacks. In *Financial Cryptography and Data Security*, pp. 167–183. Springer.

Furnell, S. and M. Warren (1999). Computer hacking and cyber terrorism: the real threats in the new millennium? *Computers & Security 18*(1), 28 – 34.

Gelareh, S. and S. Nickel (2011). Hub location problems in transportation networks. *Transportation Research Part E: Logistics and Transportation Review 47*(6), 1092 – 1111.

Gordon, L. A. and M. P. Loeb (2002, November). The economics of information security investment. *ACM Trans. Inf. Syst. Secur. 5*(4), 438–457.

Gdor, I. and G. Magyar (2005). Cost-optimal topology planning of hierarchical access networks. *Computers & Operations Research 32*(1), 59 – 86.

Hall, R. W. (1989). Configuration of an overnight package air network. *Transportation Research Part A: General 23*(2), 139 – 149.

Haubrich, D. (2003). September 11, anti-terror laws and civil liberties: Britain, france and germany compared1. *Government and Opposition 38*(1), 3–28.

Heal, G. and H. Kunreuther (2003). You only die once: Managing discrete interdependent risks. Technical report, National Bureau of Economic Research.

Heal, G. and H. Kunreuther (2005). Ids models of airline security. *Journal of Conflict Resolution 49*(2), 201–217.

Himanen, P. (2010). *The hacker ethic.* Random House.

Hosmer Jr, D. W., S. Lemeshow, and R. X. Sturdivant (2000). *Applied logistic regression* (Second ed.). Wiley-Interscience Publication.

International Air Transport Association (2012). Iata economic briefing: The impact of hurricane sandy. International Air Transport Association.

Ioannidis, C., D. Pym, and J. Williams (2013). Sustainability in information stewardship: Time preferences, externalities, and social co-ordination. In *The 12th Workshop on the Economics of Inforamtion Security (WEIS 2013).*

Irish Aviation Authority & Aviasolutions (2004). Study on civil aviation security financing. Irish Aviation Authority & Aviasolutions.

Jacobson, S. H., T. Karnani, and J. E. Kobza (2005). Assessing the impact of deterrence on aviation checked baggage screening strategies. *International Journal of Risk Assessment and Management 5*(1), 1–15.

Jacobson, S. H., T. Karnani, J. E. Kobza, and L. Ritchie (2006). A cost-benefit analysis of alternative device configurations for aviation-checked baggage security screening. *Risk Analysis 26*(2), 297–310.

Jaillet, P., G. Song, and G. Yu (1996). Airline network design and hub location problems. *Location Science 4*(3), 195 – 212. Hub Location.

Kunreuther, H. and G. Heal (2003). Interdependent security. *Journal of Risk and Uncertainty 26*(2-3), 231–249.

Loch, K. D., H. H. Carr, and M. E. Warkentin (1992). Threats to information systems: Today's reality, yesterday's understanding. *MIS Quarterly 16*(2), pp. 173–186.

Loper, D. K. (2000). *The criminology of computer hackers: a qualitative and quantitative analysis.* Ph. D. thesis, Michigan State University. College of Social Science.

Martimort, D. and W. Sand-Zantman (2006). Signalling and the design of delegated management contracts for public utilities. *The RAND Journal of Economics 37*(4), 763–782.

Martin Hvidt Thelle, Torben Thor Pedersen, F. H. (2004). Airport competition in europe. Copenhagen Economics.

Mazzocchi, M., F. Hansstein, M. Ragona, et al. (2010). The 2010 volcanic ash cloud and its financial impact on the european airline industry. In *CESifo Forum*, Volume 11, pp. 92–100. Ifo Institute for Economic Research at the University of Munich.

Mell, P. and K. Scarfone (2007). *A Complete Guide to the Common Vulnerability Scoring System Version 2.0.*

Morrow, J. D. (1994). *Game theory for political scientists.* Princeton University Press Princeton, NJ.

Oates, W. E. (1999). An essay on fiscal federalism. *Journal of economic literature 37*, 1120–1149.

O'kelly, M. E. (1987). A quadratic integer program for the location of interacting hub facilities. *European Journal of Operational Research 32*(3), 393 – 404.

O'Kelly, M. E. (1998). A geographer's analysis of hub-and-spoke networks. *Journal of Transport Geography 6*(3), 171 – 186.

Pym, D., J. Swierzbinski, and J. Williams (2013). The need for public policy interventions in information security. *Manuscript at http://homepages. abdn. ac. uk/dj pym/pages/InfoSecPubPol. pdf. Submitted for publication.*

Pym, D., J. Williams, and I. Gheyas (2014). Resilience in information stewardship. In *The 13th Workshop on the Economics of Inforamtion Security (WEIS 2014).*

Rogers, M. K. (2001). *A social learning theory and moral disengagement analysis of criminal computer behavior: An exploratory study.* Ph. D. thesis, University of Manitoba.

Sandler, T. (2000). Arms trade, arms control, and security: Collective action issues. *Defence and peace economics 11*(3), 533–548.

Sandler, T. and W. Enders (2004). An economic perspective on transnational terrorism. *European Journal of Political Economy 20*(2), 301–316.

Sandler, T. and H. Lapan (1988). The calculus of dissent: An analysis of terrorists' choice of targets. *Synthese 76*(2), 245–261.

Seidenstat, P. (2004). Terrorism, airport security, and the private sector. *Review of Policy Research 21*(3), 275–291.

Shanmugapriya, R. (2013). A study of network security using penetration testing. In *Information Communication and Embedded Systems (ICICES), 2013 International Conference on*, pp. 371–374.

SH&E Limited (2006). Capital needs and regulatory oversight arrangements. SH&E Limited.

Skorin-Kapov, D. (2001). On cost allocation in hub-like networks. *Annals of Operations Research 106*(1-4), 63–78.

Skorin-Kapov, D. and J. Skorin-Kapov (2005). Threshold based discounting networks: The cost allocation provided by the nucleolus. *European Journal of Operational Research 166*(1), 154 – 159. Metaheuristics and Worst-Case Guarantee Algorithms: Relations, Provable Properties and Applications.

Stefano Baronci (2007). Aci europe position on airport charges. Airports Council International.

Steube, G. (2004). *A logistic regression model to distinguish white hat and black hat hackers*. Ph. D. thesis. AAI3132749.

Taylor, P. A. (1999). *Hackers: crime in the digital sublime*. Routledge.

Thomson, William (2007). Cost allocation and airport problems. Rochester Center for Economic Research, Working Paper.

Turgeman-Goldschmidt, O. (2005). Hackers' accounts: Hacking as a social entertainment. *Social Science Computer Review 23*(1), 8–23.

U.S. Congress, O. o. T. A. (1984). Airport system development. *OTA-STI-231, Washington, DC*, 109–120.

Vidal, C. J. and M. Goetschalckx (2001). A global supply chain model with transfer pricing and transportation cost allocation. *European Journal of Operational Research 129*(1), 134 – 158.

Vitali, S., M. Cipolla, S. Micciche, R. Mantegna, G. Gurtner, F. Lillo, V. Beato, and S. Pozzi (201q). Statistical regularities in atm: network properties, trajectory deviations and delays. In *SESAR Innovation Days*.

Wallace, O. (1972). *Fiscal federalism*. New York: Harcourt Brace Jovanovich.