The Economics of Cybersecurity: Boomerang Effects from Misaligned Incentives

Jose J. Gonzalez*,** and Konstantin Lenchik*

*Norwegian Information Security Laboratory, NTNU Gjøvik, Norway

**Centre for Integrated Emergency Management, University of Agder, Norway

Abstract

Externalities, like misaligned incentives that charge to third parties the costs for bad information security, are tough barriers to overcome. A number of proposals for regulatory options have been suggested. However, the claim that misaligned incentives have their impact on third parties is not the whole truth. Security systems are complex not only in the sense of being composed of many interdependent parts. The most challenging part of their complexity resides in the propagation of effects, resulting in highly unexpected, counterintuitive dynamic behaviour. In particular, unintended side effects can act as boomerangs that impact hardest on the owner of the security defences who intends to push the costs of bad security to third parties. Using system archetypes and concept models we explain how misaligned incentives in the security of ATM systems acted against banks imposing the burden of proof of fraud claims on their customers. We argue that an analysis of unintended side effects arising from the misalignment of incentives is likely to benefit both agents responsible for information security and third parties.

1 Introduction

Misaligned incentives are responsible for bad cybersecurity to the extent that "security failures is caused at least as often by bad incentives as by bad design" [1, p. 610]. Misaligned incentives occur, e.g., when the organization responsible for the security of system does not bear the full costs of its failure [1, p. 610ff, 2, p. 105ff]. A number of regulatory principles have been proposed to overcome misaligned incentives hindering good cybersecurity [2, p. 107ff]. However, the way to apply those proposals in practice is long and in the meantime security continues to suffer.

This papers suggests an additional path to mitigate the occurrence of misaligned incentives: Those responsible for security, from now on called

"defenders", can be hit, albeit with a time delay, quite severely themselves by when pushing significant costs to third parties. The perceived incentive on the side of the defender is doubly misaligned: 1) because third parties, by design, suffer from the resulting externality; and 2) since ultimately the chosen security strategy hits the defender as a boomerang with a revenge owing to unanticipated side effects of the bad security solution.

This contribution continues with a section on counterintuitive behaviour resulting from feedback and time delays in complex systems (§2). Next, in §3 lessons from the experiences with ATM security designs in European and US banks illustrate that a security solution with misaligned incentives, as were possible in some European countries in the 1980'ies, hit the banks with a time delay. In contrast, in the US where legal regulations forced the banks to pursue well aligned incentives, the ATM security designs proved ultimately more profitable for the banks and the banks' customers. In §4 we use simple causal loop diagrams ("system archetypes") to provide intuitive qualitative models of the ATM security cases in Europe and USA. In §5 we sketch simple concept simulation models for ATM security that can be used to test the behaviour of the two alternative approaches to AMT security, viz. putting the burden of proof on customers or the bank assuming the burden of proof in case of ATM fraud. Section 6 provides concluding remarks.

2 Counterintuitive dynamic behaviour

Information security is an extremely complex field involving technology, organization, legislation and standards with physical, logical and social layers [3, p. 1-12]. Disciplines as different as information and communication technology, information systems, management science, microeconomics, psychology, sociology and law are involved in the design, management and maintaining of information security. Time delays, non-linearity and feedback are rampant in security management. It can safely be concluded that information security management belongs to the class of dynamical complex systems.

There are two types of complexity, combinatorial and dynamic. Combinatorial complexity is the aggregate impact of high number of system components; combinatorial complexity can be efficiently dealt with by decomposing the system in subsystems, each small enough to be easily handled. Dynamic complexity refers to the behaviour over time caused by non-linear relations and

feedback among the system components. Dynamic over time behaviour driven by nonlinear relations tends to be difficult to predict, even when the system is small. In addition, more often than not, interdependencies in real systems propagate from node to node with time delays, so that the distance between interventions affecting the system and the results caused by the interventions can be large both in terms of where and when the results show up [4].

A main consequence of the above is that interventions in dynamically complex systems always have side effects.

In the spirit of reference [5], consider first the outcome intended by the decision maker. The intervention must be applied over some period of time and the outcome will be some time-dependent result that in turn will influence the dosage of the intervention (expressed by the influence arrows in Figure 1). The closed loop of causes and effects describes a pattern of feedback occurring over some time interval. The feedback is shown symbolically by the loop labelled 'Intended consequence feedback loop'.

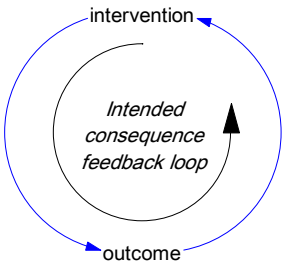

Figure 1 The outcome is achieved gradually during some time interval. The intervention changes over time as it is administered in relation to the evolving outcome.

Owing to the interdependent system components the outcome will cause side effects. Unless the decision maker has done an excellent job of modelling the system so as to anticipate the side effects, the system reaction will be unintended and, more often than not, unexpected. Again, one has feedback acting over some period of time (labelled on Figure 2 as 'Unintended consequences feedback loop'). The line labelled "system boundary" indicates that the unintended consequences are hidden from the view of the decision maker. In dynamically complex real systems the

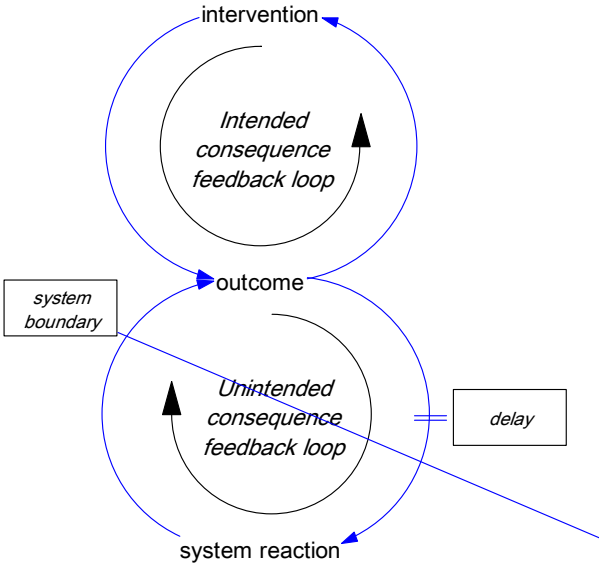

Figure 2 The system reaction arises with significant time delays and is hidden from the view of the decision makers

effects of interventions tend to show up far away from the origin of the intervention. Also, the unintended consequences can appear with significant time delays as side effects, so that the causal connection between the intervention and the system reaction is not apparent. The significant time delay is shown on Figure 2 by the || on the influence arrow going from 'outcome' to 'system reaction'.

Another important aspect is that, quite often, the dynamic complexity of the system makes the unintended consequences highly counterintuitive.

Initially, in dynamic complex systems the intervention mostly achieves the intended outcome, but as the system reaction evolves the unintended consequence often compromises the intended outcome. This phenomenon is known as policy resistance [4].

3 The security of ATMs in Europe and USA

In a survey of fraud against Automatic Telling Machines (ATMs) [6], Anderson found that patterns of fraud depended on whether the bank's customer or the bank itself was liable for them. In some countries, including the USA, if a customer disputed a transaction, the bank had the burden of proof that the customer was mistaken or lying; this gave the banks a motive to protect their systems properly. But in several European countries (including Britain, Norway and the Netherlands), the customer had the burden of proof: the bank was right unless the customer could prove it wrong – an almost impossible task. The "lucky" banks in these countries became complacent and careless. Eventually, avalanches of fraud demolished their complacency. In contrast, the banks in the USA and other countries having the burden of proof suffered much less fraud. Most remarkably they spent less money on security than their European counterparts. Thus, better aligned incentives, whereby the defender suffered most if security was bad, turned out to be the best investment for the banks and for the banks' customers as well [1, p. 611, 6].

4 Qualitative model of ATM security

Consider first the European ATM case. The banks acted by setting up the ATM system so that if the customer disputed the transaction, the burden of proof was on the customer. Thus, the bank's intervention is 'Burden of proof on customers', see Figure 3. The intended outcome was to reduce the number of

fraudulent transactions by the customer (represented by the variable 'Fraudulent transactions') to some acceptable target. Thus, one has as intended consequence a control strategy, expressed by the balancing feedback loop labelled 'B: Customer is liable' on Figure 3. The influence arrow from 'Burden of proof on customers' to 'Fraudulent transactions' has a minus sign – a negative polarity – expressing that the two variables move in opposite direction. That is, if the burden of proof on customers is increased, the outcome – fraudulent transactions – gets reduced (and vice versa).

The unintended consequence of the bank putting the burden of proof on the customer is an increase in the bank's complacency [1, p. 611, 6], – shown on Figure 3 by the influence arrow from 'Burden of proof in customers' to 'Bank's complacency'. Note that this arrow has positive polarity, expressing that the variables move in the same direction. That is, an increase in the burden of proof exerted on customers increases the bank's complacency, whereas if the bank exerted less pressure on making the customer liable, the bank's complacency would decrease.

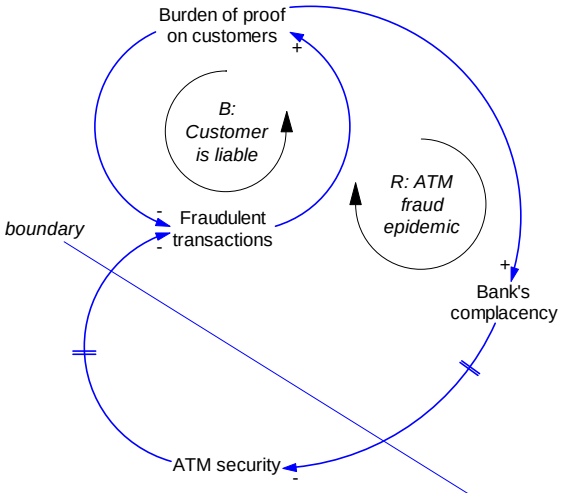

Figure 3 Archetype for the European ATM case

In turn, the variable 'Bank's complacency' influences 'ATM security' with negative polarity: an increase in the bank's carelessness decreases the ATM security over time – with some time delay, indicated by ||, as too little is done to analyse the causes of fraud, discover vulnerabilities and exploits, and remedy them. Over time, again with some delay, 'ATM security' influences 'Fraudulent transactions' with negative polarity – expressing that a decrease in 'ATM security' increases the rate of fraudulent transactions – as more and more crooks discover the poor security in the ATMs along with the bank barking up the wrong tree.

Note that the influence arrow from fraudulent transactions to burden of proof on the customer closes a second feedback loop. Walking along the influence links and considering their polarities it can be recognized that this feedback

look is reinforcing (R): if, e.g., the bank increases the burden of proof on customers, the chain of influences along the feedback loop 'R: ATM fraud epidemic', ultimately forces the bank to a further increase of the burden of proof on the customers.

The archetype on Figure 3 is an out-of-control archetype in the terminology of Wolstenholme [5]. The balancing feedback loop 'B: Customer is liable' expresses the intended consequence of the bank's its strategy, viz. to control fraud. The unintended consequence is expressed by the reinforcing feedback loop 'R: ATM fraud epidemic'. Reinforcing feedback loops can act viciously or virtuously, depending on whether they are triggered to increase or decrease unpleasant effects. In this case, the reinforcing feedback loop is vicious indeed. Owing to the banks' refusal to recognize their prominent part in the bad ATM security [6] – expressed symbolically by the boundary line on Figure 3 – and the time delays in the chain of influences, crooks exploited the numerous vulnerabilities in the ATMs, producing an avalanche of fraud that at long last caused major customer dissatisfaction, loss of reputation and ultimately forced the banks to improve the neglected ATM security – at much higher costs than a well-designed proactive security would have required [1, 6].

Consider now the US ATM case. Here, the bank's strategy built on the opposite principle than in the European case. If the customer disputed an ATM transaction the burden of the proof was on the bank. Thus, the bank's intervention is 'Burden of proof on bank' on Figure 4. The intended consequence is to reduce the number of fraudulent transactions (represented by the variable 'Fraudulent transactions') to some acceptable target. The bank assumed the responsibility and spent resources on ATM security as needed (expressed by 'Security spending')[1, 6], which affected fraudulent transactions with negative polarity. To the extent that fraudulent transactions occurred, the burden of proof on the bank was exerted, closing the loop. The

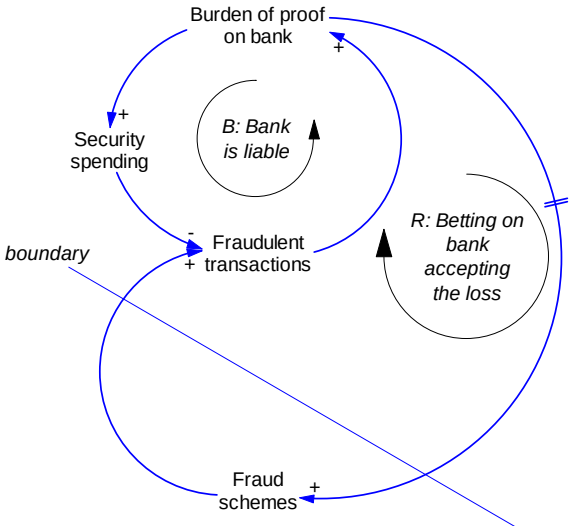

Figure 4 Archetype for the US ATM case

intended consequence was controlling, resulting in a balancing feedback loop, labelled 'B: Bank is liable'.

Customers and non-customers know that it is difficult and costly for the bank to prove who did the fraudulent transaction. They know too that the bank will not act if the fraudulent transactions involve small sums of money. Hence, dishonest customers and professional crooks speculated on that, and (with some time delay) they came up with ingenious 'Fraud schemes' (positive polarity), which increased the number of 'Fraudulent transactions' (positive polarity). The unintended outcome was a reinforcing loop ('R: Betting on the bank to accept the loss').

Also for the US case the problem archetype is an out-of-control one, following the terminology of Wolstenholme [5]. But the impact of the out-of-control archetypes was quite different for European and American banks.

In the European case the banks did not pay enough attention to the ATM security. As the unintended consequence showed up, with significant time delays (Figure 3), the banks were increasingly facing bad publicity and loss of customers, as well as getting involved in costly court disputes. Sometimes the customers won, making the banks losing face. In the end, the banks had no choice but to acknowledge that the original security solution was bad and to make big investments in security. The investments was very costly, since the ATM system was neither designed nor maintained with security in mind, and the solution was less good than if the bank had made security a strong priority in the first place [6].

In the US case, the banks designed and maintained the ATM system with security in mind. Figure 4 shows that ATM security is embedded in the intended outcome feedback loop. Although advances in fraud schemes forced the banks enhance the ATM security, the fact that the banks were security aware and that they were not losing face facilitated a quick reaction and the remedy was less costly than in the European case. This is in accordance with the facts [1, 6].

5 Simulation model of ATM security

Research on information security is hindered scarcity of data, owing to several reasons [7]: attackers conceal as many aspects of their attacks as possible;

organizations gather data on attacks for specific purposes that are not necessarily aligned with scientific data sampling; organizations controlling data assets are very reluctant to share data on those assets out of fear of bad publicity. Our case is not different: Unfortunately, all the information available about the ATM security case in Europe or the US is qualitative and can be summarised in a few statements – as was done on p. 4. All that can be said about the reference behaviour is that the ATMs in some European countries were exposed to an avalanche of fraud while the ATMs in the US were much safer. Further, that US banks ultimately invested less in ATM security while their ATMs nevertheless were more secure than their European counterparts.

The archetype analysis of the previous section is part of a course on security management at the Norwegian University of Science and technology (NTNU), Campus Gjøvik, Norway. After performing the archetype analysis students are given the challenge to build concept models [8]. To this effect a case is described for the ATM security in a typical European and a typical US bank at the time of the introduction of ATMs. The case description includes a qualitative reference behaviour pattern for each of the cases along with some hints. Basically, it is required that the simulation reproduces two key observations about patterns of behaviour (p. 4): 1) that ATMs in some European countries were exposed to an avalanche of fraud while the ATMs in the US were much safer; 2) that the US banks invested less in ATM security while their ATMs nevertheless were more secure than their European counterparts.

In the following we describe concept models for the ATM cases. The model structure and the equations are kept as simple as possible, partly to avoid going too far given the scarcity of empirical data, partly for pedagogical reasons.

The core structure of the concept models is shown on Figure 5. So far this structure applies for both ATM cases (European and US banks).

ATMs have vulnerabilities that can be exploited to commit fraud. Vulnerabilities exist in two states, represented by the stocks 'Vulnerabilities dormant' and 'Vulnerabilities active'. Dormant vulnerabilities have not yet been discovered and, hence, cannot be exploited. By chance or clever schemes vulnerabilities are discovered and become 'active – that is, exploitable. The process rendering dormant to active vulnerabilities is represented by the flow 'vulnerability activation' in Figure 5. Active vulnerabilities can be fixed, represented by the flow 'vulnerability fixing'. The pressure to fix known

('active') vulnerabilities was significantly higher for US banks – who had the burden of proof with regards to fraud claims– than for European banks, who made customers liable and didn't suffer much when fraud was committed.

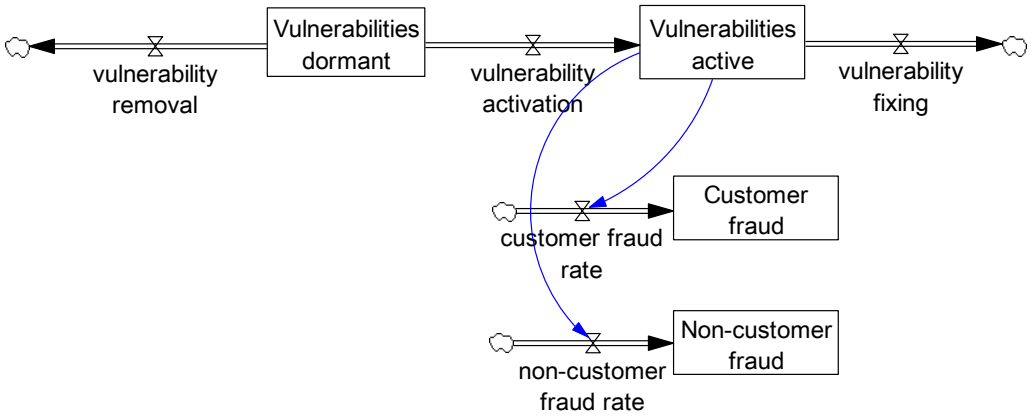

Figure 5 Core structure for a concept model of the ATM case

A proactive posture would in addition imply investment in discovery and removal of as yet unknown dormant vulnerabilities (represented by the flow 'vulnerability removal') in Figure 5. Discovery and removal of dormant vulnerabilities is more demanding and costly than fixing active vulnerabilities – which manifest themselves by the fact that they are exploited and fraud occurs.

We may safely assume that US banks to a much larger extent did assume such proactive posture, whereas European banks mainly acted after the consequences of their neglect of ATM security led to an escalation in angry customer complaints and bad publicity.

Figure 5 shows the two possible mechanisms for ATM frauds, viz customer fraud and non-customer ("crook") fraud. The model has to differentiate between frauds committed by bank customers and by non-customers because the European bank exerts pressure on customers ("burden of proof on customers"). The influence arrows going from the stock 'Vulnerabilities active' to the flows 'customer fraud rate' and 'crook fraud rate' express that the fraud rates depend on the extent to which there are active vulnerabilities in the ATMs.

The full concept models for the European and US ATM case are shown on Figures 6-7 respectively.

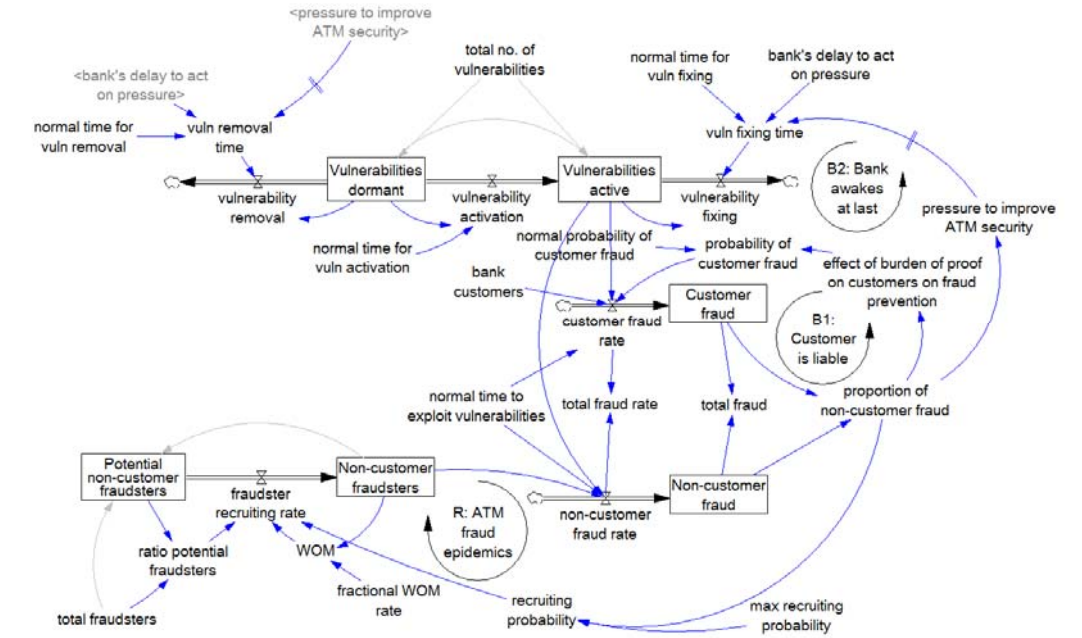

Figure 6 Concept model for the European ATM case

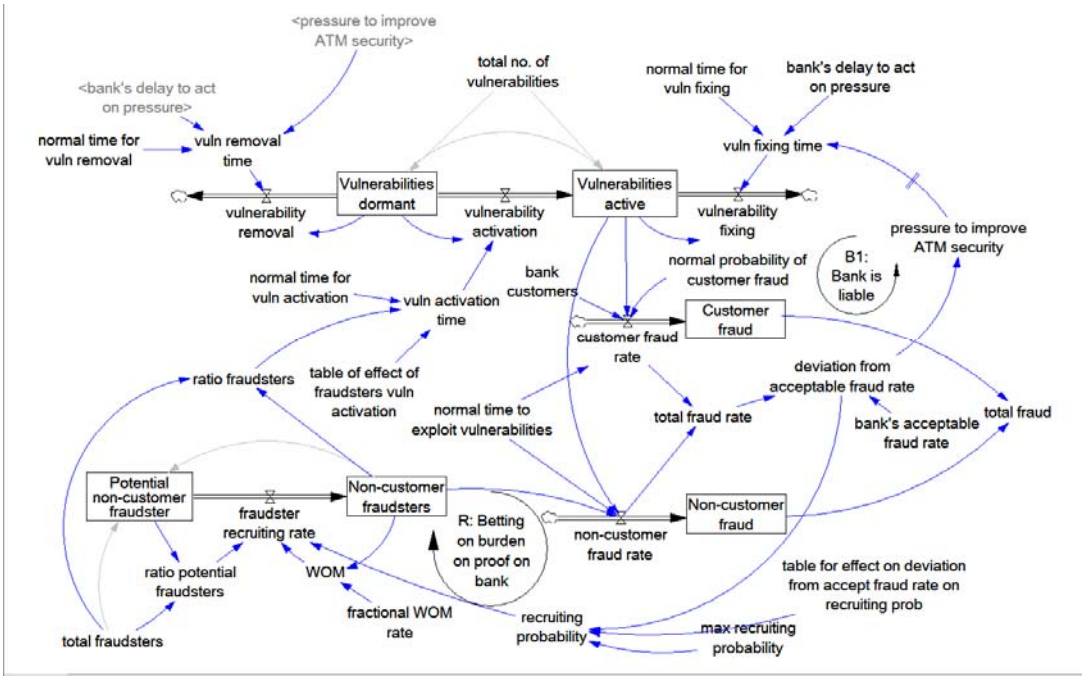

Figure 7 Concept model for the US ATM case

Beyond the core structure shown on Fig. 5, the full concept models for the two ATM cases share some additional features.

Both models assume that the number of bank customers is constant (10 000 people) and that the probability that customers commit fraud is very low. For the US case, since the ATM security is kept high and the bank does take legal

measures unless the fraud committed is high, we assume that this probability stays constant over time at *'normal probability of customer fraud'=0.025 %*.

In the European case, since the bank's policy was to put the burden of proof on customers, people who took money from their own account, but argued that the transaction was fraudulent (committed by others), experienced that their strategy does not work.

Hence, for the European case we assume *'probability of customer fraud'= 'normal probability of customer fraud'*'effect of burden of proof on customers on fraud prevention'*. Figure 8 shows the relevant part of the European ATM model in this respect. The lookup variable '*effect of burden of proof on customers on fraud prevention*', with input '*proportion of non-customer fraud*', is defined as an S-shape curve monotonously declining from unity to zero. Hence, the higher the proportion of non-customer fraud is, the lower the probability of customer fraud becomes.

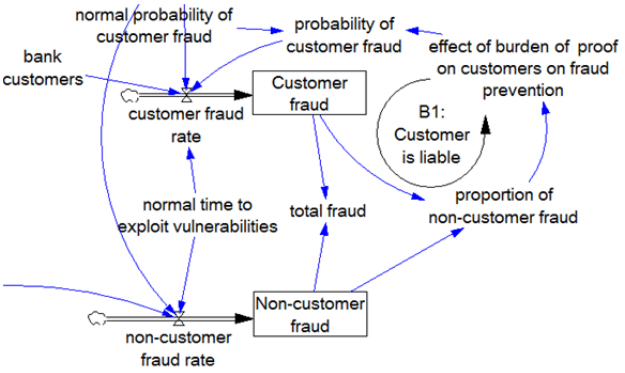

Figure 8 The balancing loop "Customer is liable" expresses that the bank's strategy of burden of proof on customers influences the bank customer's probability of committing fraud.

Both ATM models assume that fraudsters are recruited via a word-of-mouth process (Figure 9). In the European ATM case the recruiting probability is computed as *'max recruiting probability'*'proportion of non-customer fraud'*, where *'max recruiting probability'=1*. For the US ATM case we set *'recruiting probability'='max recruiting probability'*'table for effect on deviation from accept fraud rate on recruiting prob(deviation from acceptable fraud rate)'*, which uses a table function expressing the effect of the US bank's strategy on recruiting of potential fraudsters.

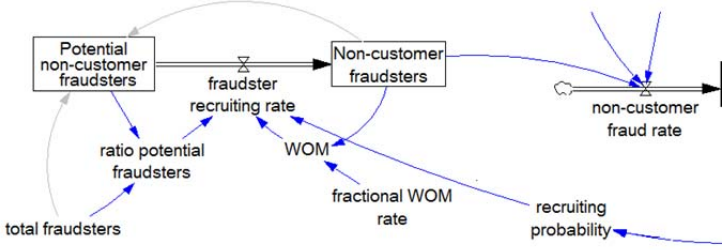

Figure 9 Common model structure for recruiting of fraudsters

The US bank's strategy, burden of proof on the bank, implies that the bank will not care to take legal measures as long as the fraud rate is lower than some acceptable fraud rate (*'bank's acceptable fraud rate'* in the model). The table function '*table for effect on deviation from accept fraud rate on recruiting prob*' expresses that the higher the bank's acceptable fraud rate is, the higher recruiting probability of fraudsters will be (Figures 9 and 10).

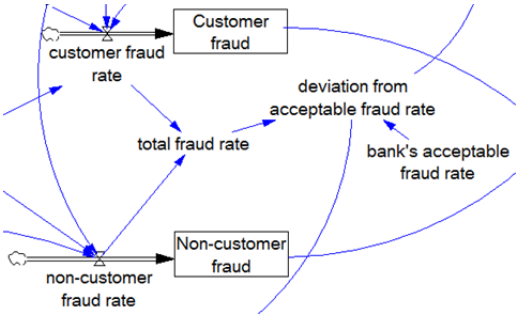

Figure 10 Since the US bank assumes the burden of proof, it doesn't act legally unless the fraud rate surpasses the acceptable fraud rate. The higher the acceptable fraud rate, the higher the probability to recruit fraudsters.

In both ATM cases the bank will feel pressure to improve the ATM security.

In the US case, the higher the value of '*deviation from acceptable fraud rate*', the higher the pressure to improve the ATM security. This translates to shorter times to fix active vulnerabilities and to remove dormant vulnerabilities.

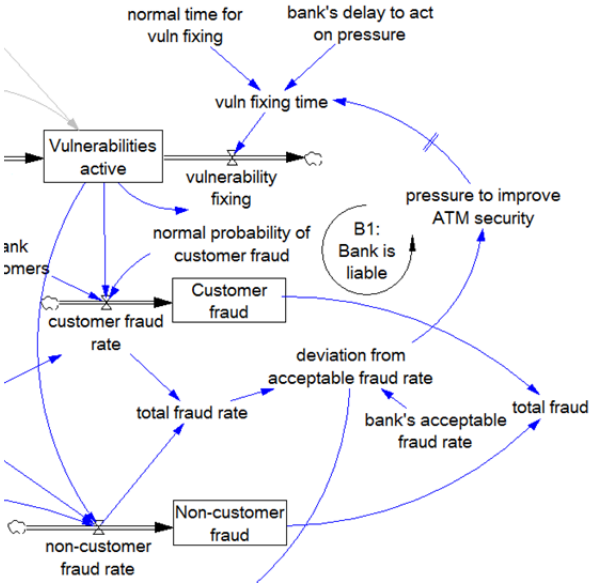

Figure 11 The US bank feels higher pressure to fix active vulnerabilities if the deviation from the acceptable fraud rate increases

Figure 11 shows the balancing loop 'Bank is liable' expressing that the higher the deviation from the acceptable fraud rate is, the shorter the time to fix active vulnerabilities becomes. A similar structure connects '*deviation from acceptable fraud rate*' to the time to remove dormant vulnerabilities (cf. Figure 7).

In the European ATM case the bank did not react until the avalanche of fraud (reinforcing loop "ATM fraud epidemics", Figure 6Figure 6), which resulted in innocent customers losing money, led to massive customer protests and bad publicity. The balancing loop 'Bank awakes at last' expresses that the higher the avalanche of fraud is, the shorter the time to fix active vulnerabilities becomes – albeit with a significant delay. A similar structure connects '*pressure to improve ATM security*' to the time to remove dormant vulnerabilities (cf. Figure 6).

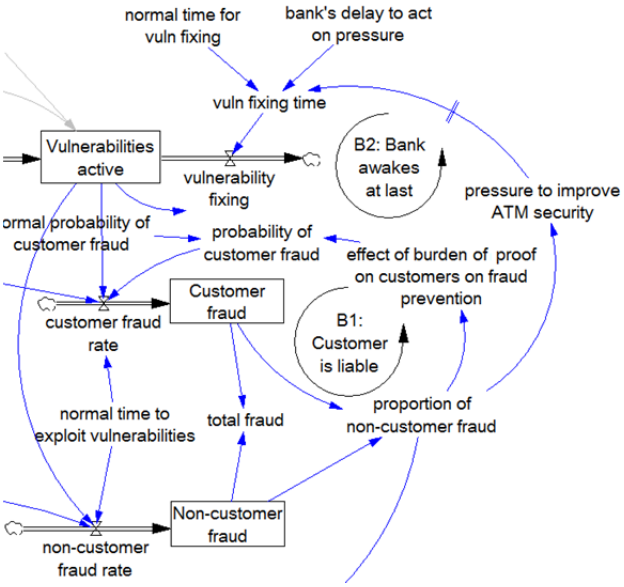

Figure 12 The European bank felt increasing pressure to improve ATM security as the avalanche of fraud triggered massive customer protests and bad publicity (B2: Bank awakes at last).

We mention finally, that fraudsters reacted by devising schemes to activate dormant vulnerabilities as counterstrategy to the to the fact that US banks accepted fraud below some threshold while at the same time taken care to continuously improve ATM security (that is both fixing active vulnerabilities and taking efforts to remove the dormant vulnerabilities). This is represented in Figure 7 by the link connecting '*ratio fraudsters*' to '*vuln activation time*'.

Figure 13 displays the simulated total fraud for both ATM cases. In (qualitative) agreement with the facts, the typical European bank was for quite long time a passive observer of an avalanche of fraud while the typical US bank did not face the rising wave of dissatisfaction and complaints that ultimately forced the European banks to change the rules of the game: they had to invest much more to fix the deplorable security of their ATMs in addition to losing face and have ultimately to compensate angry customers who to begin with were suspected of having effected the fraud transactions.

Figures 15-19 provide insight in the processes going on in the two cases.

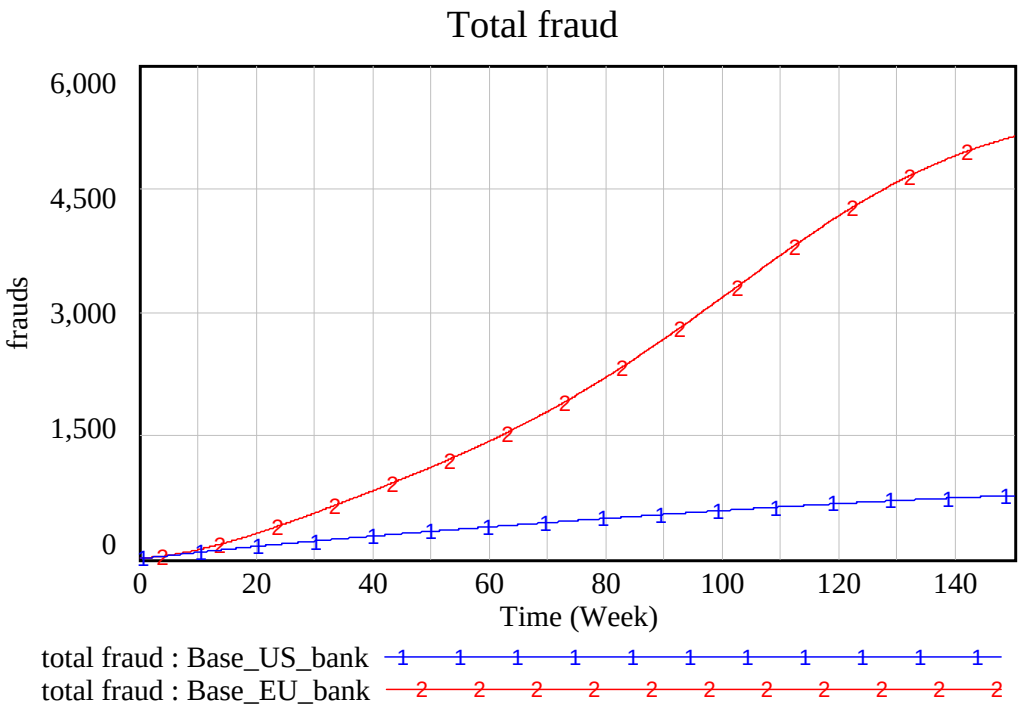

Figure 13 Simulated total fraud for the European and the US ATM case

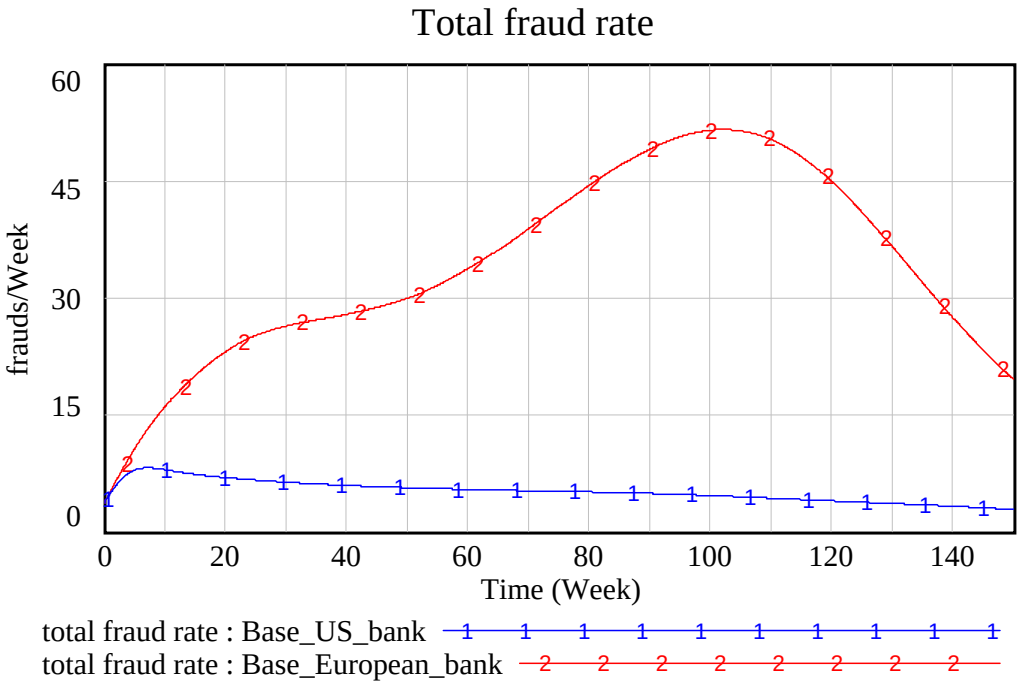

Figure 14Simulated total fraud rate for the European and the US ATM case

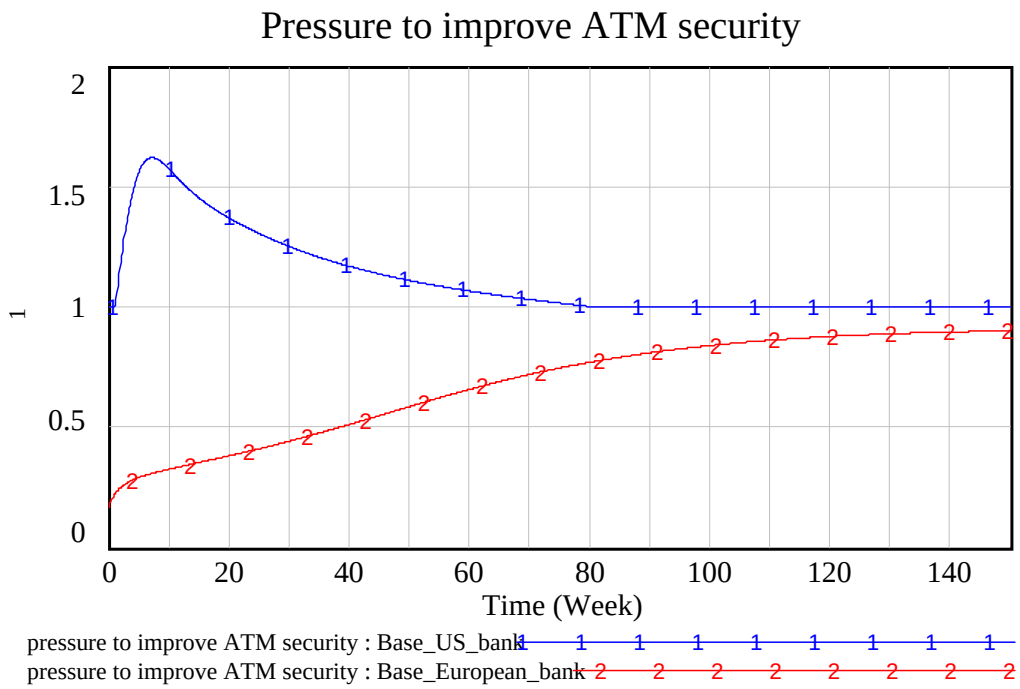

Figure 15 Pressure to improve ATM security for the European and the US ATM case

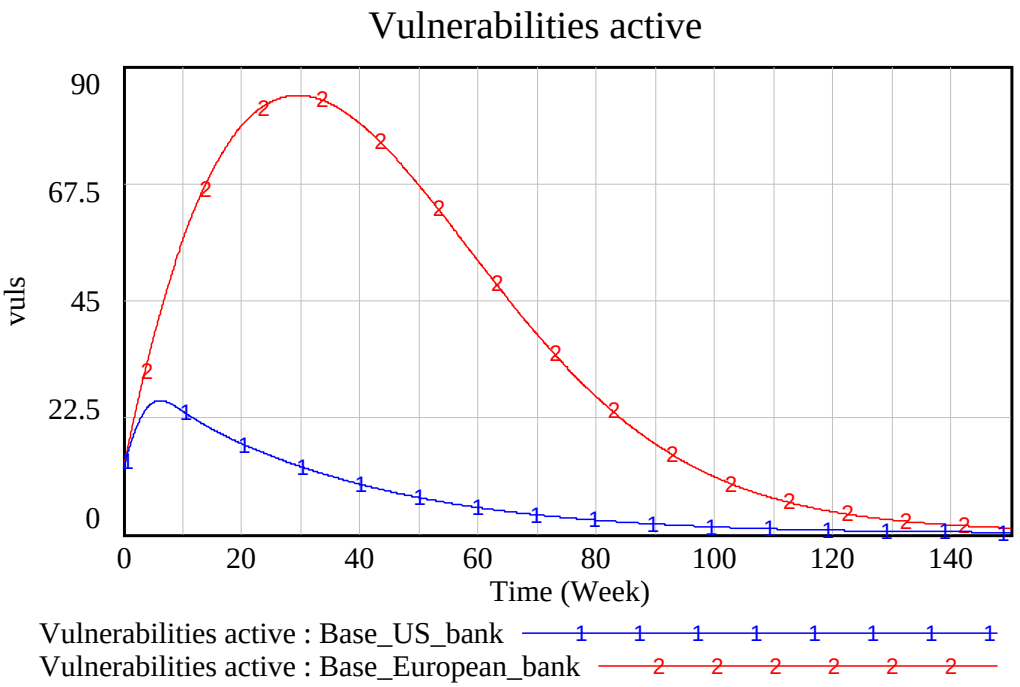

Figure 16 Active vulnerabilities for the European and the US ATM case

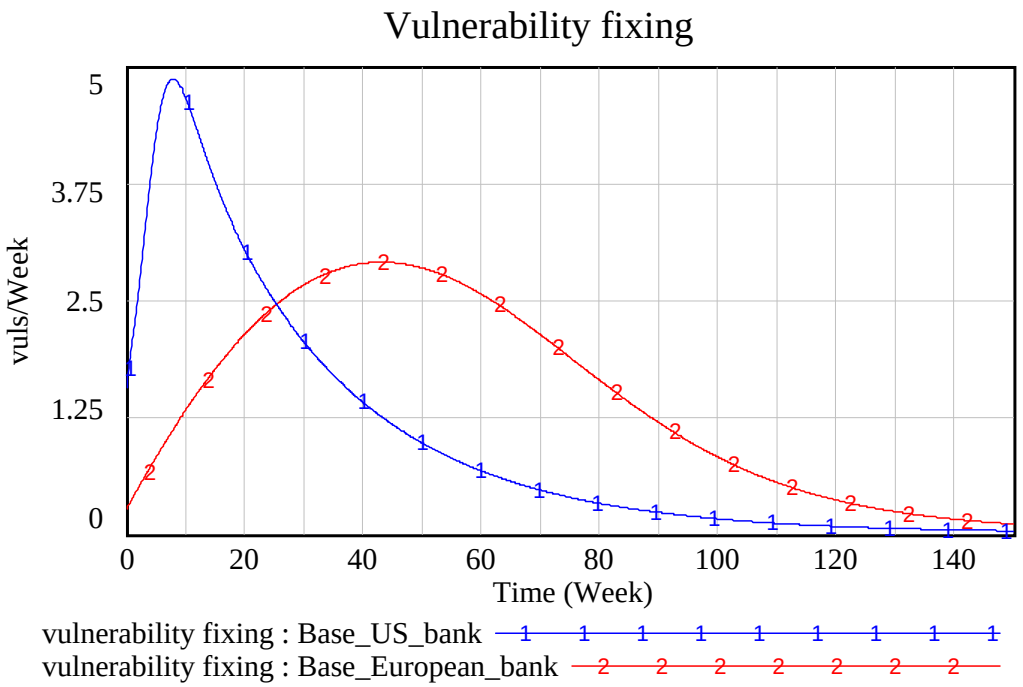

Figure 17 Rate of fixing active vulnerabilities for the European and the US ATM case

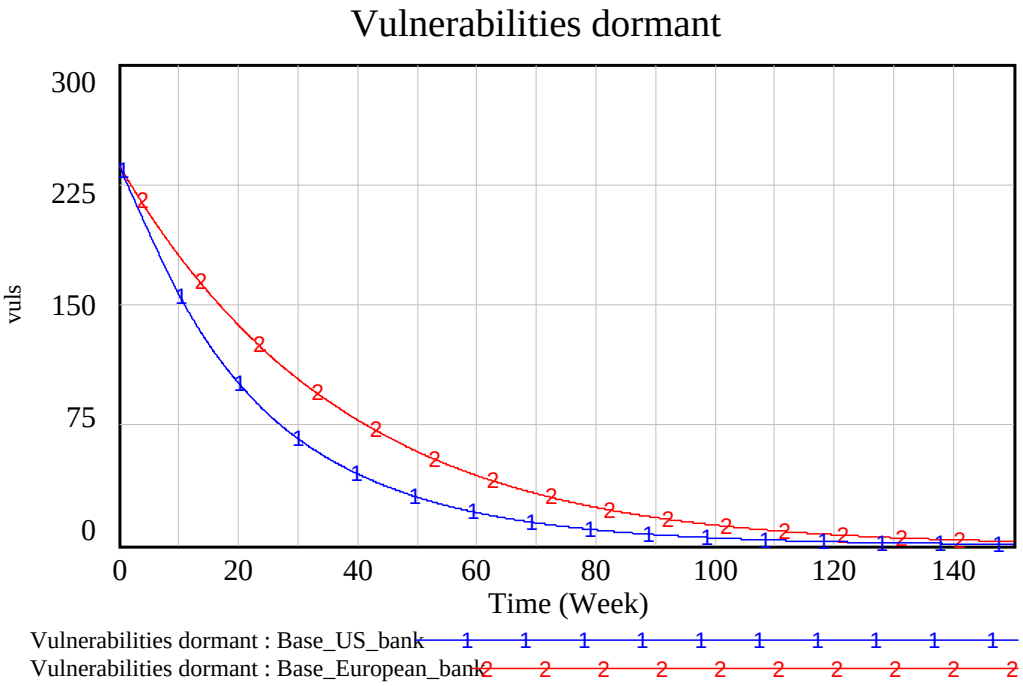

Figure 18 Dormnt vulnerabilities for the European and the US ATM case

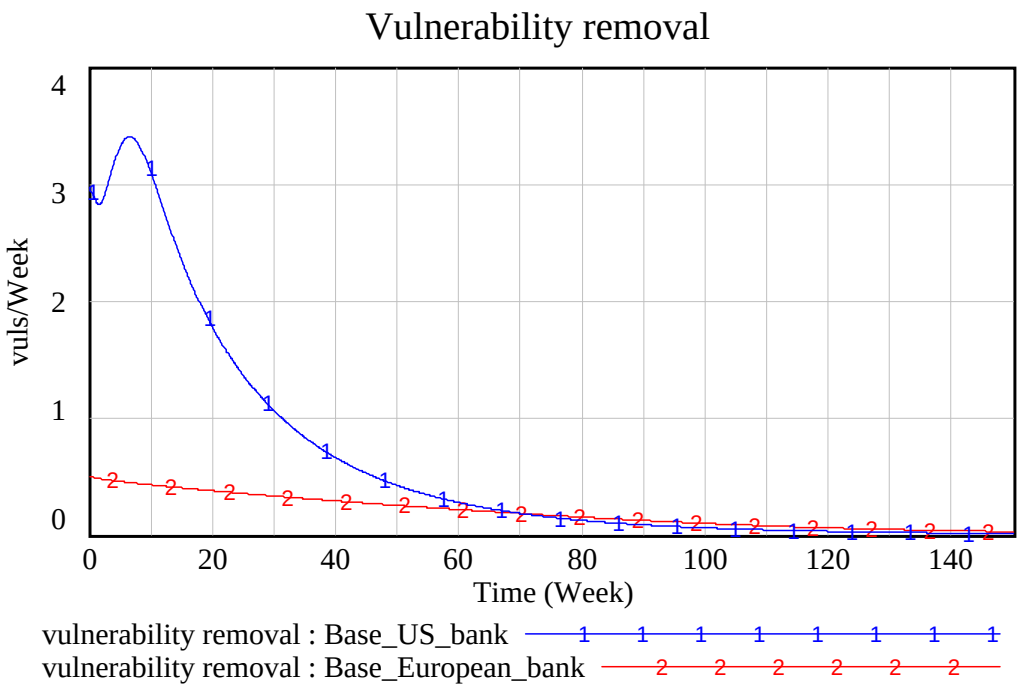

Figure 19 Removal of dormant vulnerabilities for the European and the US ATM case

6 Concluding remarks

Archetypes like those shown on Figures 3-4 and concept simulation models like those shown on Figures 6-7 have low cost and take short time to develop. In particular, expert modellers with expertise in system dynamics are quick in identifying feedback that is likely to compromise the intended outcome of interventions.

Assume that the European banks, instead of trying a costly strategy without analysing its unintended consequences, had invested some thousand euros in developing such archetypes and a concept model so as to understand the impact of different strategic choices. Presumably, the European banks would have got second thoughts and rather opted for a better security solution?

In several application areas it has been shown that investing in simulation models for strategy analysis cost very little compared to the cost of failures done by bad decisions. There is still a long way to go concerning the availability of quality data in information security. Regrettably, organizations making their data available for analysis and simulation are very scarce. By teaching system archetypes and system dynamics in courses of security management we are hoping to increase the awareness of the benefit of systems thinking to anticipate

and prevent the impact of misaligned incentives in information security. Their long term effects can act as boomerangs upon the party who tries to pass the consequences of bad security on their parties.

7 References

[1] Anderson, R., and Moore, T., "The Economics of Information Security", Science, 314(5799), 2006, pp. 610-613.
[2] Moore, T., "The Economics of Cybersecurity: Principles and Policy Options", International Journal of Critical Infrastructure Protection, 3(3–4), 2010, pp. 103-117.
[3] Trcek, D., Managing Information Systems Security and Privacy, Springer, Berlin, Heidelberg, 2006.
[4] Sterman, J.D., Business Dynamics : Systems Thinking and Modeling for a Complex World, Irwin/McGraw-Hill, Boston, 2000.
[5] Wolstenholme, E.F., "Towards the Definition and Use of a Core Set of Archetypal Structures in System Dynamics", System Dynamics Review, 19(7), 2003, pp. 7-26.
[6] Anderson, R.J., "Why Cryptosystems Fail", Proceedings of the First ACM Conference on Computer and Communications Security, 1993, pp. 215-227.
[7] Wiik, J., Gonzalez, J.J., Lipson, H.F., and Shimeall, T.J., "Dynamics of Vulnerability - Modeling the Life Cycle of Software Vulnerability", The 22nd International Conference of the System Dynamics Society July 20-24., 2004
[8] Richardson, G.P., "Concept Models in Group Model Building", System Dynamics Review, 29(1), 2013, pp. 42-55.