JOURNAL OF CYBERSECURITY

Journal of Cybersecurity, 2025, tyaf003
https://doi.org/10.1093/cybsec/tyaf003

Research paper

The use of simulations in economic cybersecurity decision-making

Mazaher Kianpour [1,2,*] **and Ulrik Franke** [2,3,4]

[1]Department of Information Security and Communication Technology, Norwegian University of Science and Technology , Høgskoleringen 1, 7034 Trondheim, Norway
[2]RISE Research Institutes of Sweden, P.O. Box 1263, SE-164 29 Kista, Sweden
[3]Swedish Defence University, P.O. Box 278 05, SE-115 93 Stockholm, Sweden
[4]KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden

*Corresponding author. E-mail: mazaher.kianpour@ntnu.no

Received 12 December 2023; revised 29 August 2024; accepted 23 January 2025

Abstract

This paper presents an in-depth examination of the use of simulations in economic cybersecurity decision-making, highlighting the dual nature of their potential and the challenges they present. Drawing on examples from existing studies, we explore the role of simulations in generating new knowledge about probabilities and consequences in the cybersecurity domain, which is essential in understanding and managing risk and uncertainty. Additionally, we introduce the concepts of "bookkeeping" and "abstraction" within the context of simulations, discussing how they can sometimes fail and exploring the underlying reasons for their failures. This discussion leads us to suggest a framework of considerations for effectively utilizing simulations in cybersecurity. This framework is designed not as a rigid checklist but as a guide for critical thinking and evaluation, aiding users in assessing the suitability and reliability of a simulation model for a particular decision-making context. Future work should focus on applying this framework in real-world settings, continuously refining the use of simulations to ensure they remain effective and relevant in the dynamic field of cybersecurity.

Keywords: simulations; economics; decision-making under risk; decision-making under uncertainty; bias

Introduction

Cybersecurity is crucial for protecting tangible and intangible assets in modern society. However, achieving it is challenging, e.g. because of rapid technological development [1], cognitive limitations [2], lack of a skilled workforce [3], a plethora of stakeholders [4], and for many other reasons.

Precisely because cybersecurity depends on so many factors, it can be rewardingly approached from many different perspectives, including nontechnical ones, such as behavioral science [5], management (Dutt et al. [6]), or economics (Anderson and Moore [7]). (Conversely, cybersecurity research is impeded when disciplinary barriers *cannot* be overcome [8].) In this paper, we first and foremost adopt the perspective of economic decision-making in cybersecurity. While we readily acknowledge that this is but one out of many possible perspectives, we also believe that economic decision-making is a particularly rewarding perspective, because this is where many of the other perspectives meet and have to be traded-off against each other, e.g. in decisions, such as how much of a security budget to spend on better intrusion detection, training personnel, vetting vendors, improving SLAs, information sharing with others, or cyber insurance, respectively [9].

One particular such challenge when making such decisions is the scarcity and poor quality of the data available—leaving decision-makers to face risk, uncertainty, and ambiguity [10]. Partly, this is due to the interconnectedness of digital systems, the rapid evolution of technology, and the increasing sophistication of cyber adversaries (factors amenable to be studied in a positivist paradigm, where it is assumed that the phenomena of interest can be approached from an outside, objective point of view and explained, e.g. through inductive measurements or deductive logic.), but difficulties also arise from the intricate interplay between human, technological, organizational, and institutional factors (which may require more interpretive, so called *Verstehen*, methods, where it is assumed that the phenomena of interest are best understood by adopting

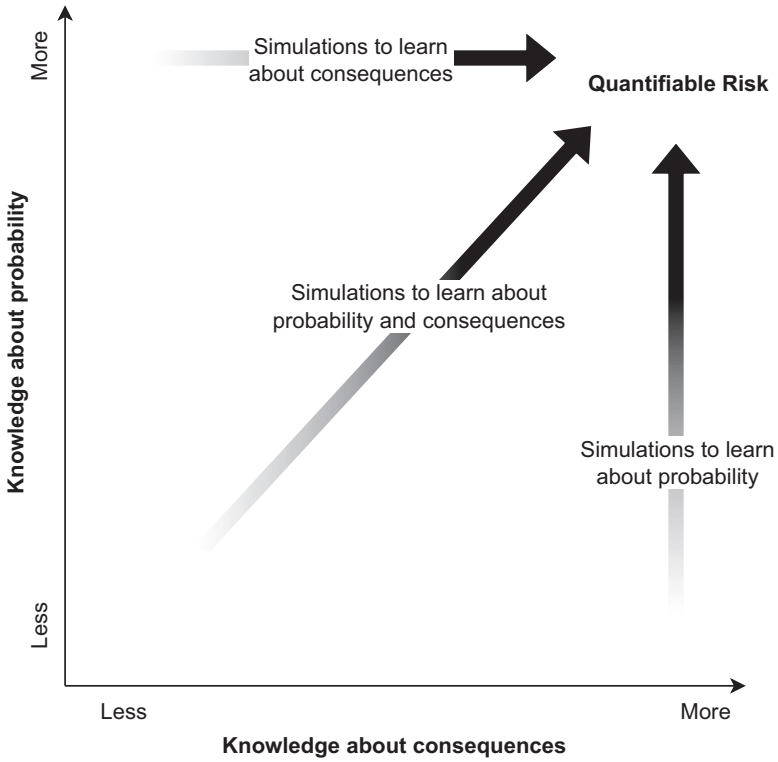

Figure 1 Knowledge about consequences and probabilities, constituting different kinds of decisions. Simulations can be a source of knowledge about the intricate relationship between probabilities, consequences, and their cumulative impact on quantifiable risk.

an inside perspective [11]) that collectively shape the cybersecurity landscape.

Lack of high-quality cybersecurity data is a long-standing and well-known problem (see, e.g. the reviews [12–14]). Proposals to mitigate it typically involve mechanisms, such as mandatory incident disclosure [15], Section 4.1.2, partnerships [16], and information sharing [17], but there remains a significant lack of accessible, high-quality data for economic cybersecurity decision-making. This hampers the ability of organizations to make rational, evidence-based decisions (see, e.g. [18–21]) to safeguard their digital ecosystems, even though we do not claim that all stakeholders need detailed quantitative data to make evidence-based policies.

It is against this background that this paper seeks to explore the potential of *simulations* as a methodological approach to generating additional data for economic cybersecurity decision-making[1]. Simulations offer a controlled and experimental environment, allowing for the systematic investigation of cyber threats, vulnerabilities, and the effectiveness of various defensive mechanisms [25]. By exploiting the power of simulations, it seems that detailed and rich datasets could be generated, and thus to some extent mitigate the scarcity of data. The purpose of this paper is to investigate this, delineating the prospects and calling attention to the pitfalls.

The rest of the paper unfolds as follows. In Section 2, some decision-theoretic background, which serves as a conceptual framework is provided. Section 3 then explores how new knowledge can be had from simulations, and perhaps more importantly, how and why this can fail. The paper is concluded with a discussion and outlook in Section 4.

Decision-making and two kinds of knowledge

Decision-making processes in cybersecurity must incorporate not only *known* information regarding states, threats, and vulnerabilities but also possess the flexibility to adapt to *new* information and evolving scenarios, not least due to rapid technological development [1]. To be effective, these processes need to be capable of considering and processing a range of potential outcomes, including those that are currently unknown or poorly understood. This necessitates a decision-making framework that is both dynamic and robust, capable of navigating the risks and uncertainties inherent in the cybersecurity domain [26,27].

Decision theory traditionally distinguishes between two kinds of knowledge needed for decision-making: knowledge of the possible *consequences* in different scenarios and knowledge of the *probabilities* of those scenarios. Despite its apparent simplicity, this framework, depicted in Fig. 1 enables rich and subtle analysis of different decision situations [28].

1 In this paper, we use the terms "model", "simulation model", and "simulation" somewhat interchangeably, although they denote distinct concepts. While the literature offers a range of definitions—from simple to complex—we define these terms as follows, drawing on the studies cited in this paper [22–24]. A model is a simplified representation of a system, which can be mathematical, computational, or conceptual, and is used to understand, analyze, or predict the behavior of that system. A simulation model specifically refers to a model designed for simulations, allowing exploration of a system's dynamics through one or more simulated scenarios. A simulation, in turn, is the process of using a model to replicate the behavior of a real-world system over time, running the model under various conditions to observe outcomes and derive insights. Throughout this paper, when we refer to a "model", we generally mean a "simulation model", assuming its use within the context of a simulation.

Knight, in his seminal work "Risk, Uncertainty, and Profit" [29], made the distinction between *risk* and *uncertainty* that has had a significant impact on economics and the study of decision-making under conditions of limited knowledge. From Knight's perspective, in decisions under risk, the scenarios are uncertain, but they are measurable and can be assigned probabilities based on past data or well-understood distributions. In other words, risks can be quantified, and insured against. In decisions under uncertainty, on the other hand, the probabilities of scenarios are not known and cannot be estimated. (To highlight the severe limitations on what is known under uncertainty, this case is sometimes instead called decisions under *ignorance*.) In these cases, unlike risk, uncertainties cannot be quantified or insured against. This distinction is crucial for Knight's economic theory because he argued that it is the presence of uncertainty (not merely risk) that provides opportunities for profit.[2]

Importantly, different decision rules are possible depending on how much is known about the problem at hand. For example, in a decision under uncertainty, where consequences but not their probabilities are known, only rules such as minimax or minimax regret can be applied (see, e.g. [28], Chapter 2, pp. 21–44). Such rules only involve a scenario analysis of the possible outcomes, without assuming any probabilities for the scenarios. For such a scenario analysis to be meaningful, however, the possible outcomes need to be reasonably accurate. If they turn out to be erroneous, the decision rules may fail and results may be worse than if decisions were just without rules.

In contrast, in a decision under risk, where both consequences and their probabilities are known, other rules become attractive, such as maximization of value or utility (see, e.g. Resnik [28], Chapter 3, pp. 45–80). In particular, for repeatedly recurring situations, such rules are attractive, because (as suggested by their names) they maximize the (accumulated) value or utility. If more cautious rules such as minimax are applied in such situations, the (accumulated) value or utility will be smaller. Thus, if it is possible to gain the additional knowledge (about probabilities) needed to transform a decision under uncertainty into one under risk, this is a significant improvement. On the other hand, if such a procedure fails, and the knowledge about probabilities turns out to be erroneous, then trying to maximize value or utility will also fail, and results may be worse than if only outcomes were considered, using rules such as minimax and its variants.

Theoretical economic models of cybersecurity often assume that decisions are made under (quantifiable) risk. For instance, a rational decision-maker typically invests in cybersecurity if the investment yields a positive return, or if the cost of the investment is lower than the risk it mitigates [31]. However, traditional methods of gaining knowledge about probability and consequences, such as using historical data, relying on the experience and intuition of experts, and case studies are—as noted in the introduction—often insufficient in cybersecurity decisions [12–14,32]. Decisions may be modeled as being made under risk, when they are actually made under uncertainty.

Having introduced this conceptual framework, we can now more precisely state the purpose of the paper: to explore the potential of simulations to improve our knowledge about probabilities and consequences in the decision-theoretic sense. Simulations, with their ability to model intricate systems and simulate diverse outcomes under various conditions, offer a promising avenue for gaining data beyond what is empirically available, thus offering the prospect of improving economic cybersecurity decision-making, where uncertainty and risk are prevalent. In the next section, this avenue is further explored.

The promise and peril of simulations

Simulation is a dynamic and continually evolving field, encompassing a wide range of methodologies and applications across various disciplines. Its core lies in creating models to replicate real-world processes, phenomena, or systems, allowing for in-depth analysis, experimentation, and prediction ([22], Chapter 4). Although simulations have been especially beneficial in fields, where direct experimentation is either impractical, costly, or poses significant risks, these days, the application of simulation has expanded beyond these constraints. It is now being increasingly adopted in a variety of sectors for its versatility and the depth of insight it provides [33]. In the current era, characterized by rapid technological advancements, complex systems, and uncertain environments, simulations have become a key tool in understanding the complexities of present systems and providing a landscape of plausible future scenarios [34].

The diversity of definitions and purposes attributed to simulations in various fields reflects their multifaceted nature and wide-ranging applicability. Simulations are not confined to a singular purpose or interpretation; rather, they serve a spectrum of objectives, from educational tools, entertainment, and research methodologies to decision-making aids, performing certain tasks, and predictive models.[3] While an in-depth exploration of each of these purposes is beyond the scope of this paper, we provide three examples related to the context of the paper's aim of using simulations to enhance knowledge about probabilities and consequences in decision-making:

Discovery: Among the seven purposes outlined by Axelrod [23] for which simulation can be used, employing simulation for discovery stands out as a source of new knowledge. Axelrod concurs that simulation is neither purely deductive nor inductive, and alternatively characterizes it as a third way of doing science aimed to uncover significant relationships and principles that may not be immediately apparent or accessible through traditional research methods. Like deduction, simulation begins with explicit assumptions. However, rather than proving theorems, it generates data for inductive analysis. The crucial difference is that the data stems from a predefined set of rules, rather than from direct measurement of the real world.

For example, Beresnevichiene et al. [37] present a methodology that seeks to aid information security managers in their decision-making processes regarding investments in security measures. The core of this methodology is a mathematical model of the system, which operates within dynamic threat and economic environments. By simulating this model in the presence of varying threat scenarios

2 Knight's distinction has been criticized, reconsidered, and expanded by several scholars. Stirling, for instance, expanded on the concepts of risk and uncertainty by identifying four situations of *incertitude*: risk, uncertainty, ambiguity, and ignorance [30].

3 For example, Axelrod [23] lists seven purposes of simulations: (i) prediction, (ii) performance, (iii) training, (iv) entertainment, (v) education, (vi) proof, and (vii) discovery. Winsberg [35] lists three: (i) heuristic purposes such as communicating knowledge to others or represent information to ourselves, (ii) predicting data that we do not have, and (iii) understanding data that we already have. Grüne-Yanoff and Weirich [36] list four: (i) representation, (ii) prediction, (iii) explanation, and (iv) policy decisions. While Humphreys [22] does not enumerate the purposes of simulation, he discusses at least five purposes such as: (i) representation, (ii) exploration, (iii) experimentation, (iv) prediction, and (v) explanation (epistemic opacity).

and security investments, the study reveals the potential operational consequences of different security choices. Such simulation-driven insights enable decision-makers to better comprehend the trade-offs between security performance and the ramifications of their decisions, making this study a source of knowledge on the consequences of security investment strategies.

Similarly, Caulfield and Pym [38] introduce a rigorous modeling framework grounded in semantically justified mathematical systems modeling, the economics of decision-making, and simulation. With an emphasis on the compositional nature of these models, the study emphasizes how complex systems can be understood as combinations of smaller, holistic models. By leveraging utility theory, the authors articulate the extent to which security managers' policies achieve their objectives under varied policy choices. The parametrization of these models based on real-world observations is particularly noteworthy, bridging the gap between theoretical constructs and tangible, real-world systems making them more ecologically valid ([39], Chapter 4.2).

These examples align with the idea that simulation begins with explicit assumptions but, unlike deduction, generates data for inductive analysis. By simulating complex systems and observing their behavior under various conditions, researchers can identify patterns, dependencies, and causal links that might remain hidden in real-world observations or theoretical analyses. This method of discovery through simulation extends beyond merely verifying existing hypotheses. It actively contributes to the generation of novel insights and the formulation of new theories, revealing unexpected dynamics and emergent properties of complex systems, and gaining a deeper understanding of the phenomena under study. This process is iterative, as the insights gained from one simulation can inform the design of subsequent models, leading to increasingly accurate and insightful results.

Boundary object: Star and Griesemer [40] introduced the concept of *boundary object* to describe an entity that is flexible enough to be interpreted in various ways by individuals from different domains, each possessing distinct expertise and goals. Despite these varied interpretations, a boundary object retains a consistent identity as it bridges across diverse fields. In this context, Tolk et al. [41] and Luna-Reyes et al. [42] argue that simulations can be aptly considered as boundary objects, as they embody both adaptability and stability. While adaptability allows simulations to be tailored to meet the specific requirements and objectives of different research fields, their stability provides a consistent framework for interdisciplinary collaborations. The importance and necessity of such collaborations in cybersecurity has been recognized as the technical perspective alone is not sufficient [7,8].

The utilization of simulation as a boundary object is exemplified in several studies. For example, Dutt et al. [6] integrate concepts from cognitive psychology (Instance-Based Learning Theory) with cybersecurity. By simulating and analyzing the decision-making processes of cybersecurity defenders against various attack strategies, the paper generates new insights into threat detection dynamics. These insights have practical applications in creating decision-support tools and training programs for cybersecurity professionals, effectively bridging the gap between theoretical research and practical cybersecurity strategies.

A study by Khan et al. [43] is another example of how simulation serves as a boundary object. It employs system dynamics-based Stock-and-Flow Model to analyze the complex interactions and indirect consequences of potential cybersecurity regulations in the context of connected and automated vehicles (CAVs). This simulation bridges various fields, such as public policy, cybersecurity, automotive engineering, and data analytics, providing a platform for these diverse domains to intersect and interact. It allows stakeholders from these areas (e.g. communication service providers, road operators, automakers, and CAV consumers) to engage with the model and understand the multifaceted implications of cybersecurity regulations, demonstrating both the adaptability and stability inherent in boundary objects.

Inspiration: Hartmann's categorization of five various functions of simulations offers a comprehensive perspective on their multifaceted roles in scientific inquiry. Hartmann [44] emphasizes specific aspects that highlight simulations' pivotal role for experimentalists. This includes generating new knowledge by supporting and enhancing empirical experiments, a process that entails inspiring new experiments, preselecting systems and setups, and conducting postexperimental analyses. This aligns with the concept of "thought experiment" discussed by Axelrod [23].

Digital twins exemplify Hartmann's characterization, serving as critical experimental tools in cybersecurity. Digital twin-based solutions, as discussed by Barricelli et al. [45] and Dietz and Pernul [46], provide platforms for exploring threats and vulnerabilities, studying cyber–physical systems under attack, and evaluating detection and mitigation measures. The paper by Murillo et al. [47] takes this a step further by developing a digital twin for water distribution systems that cosimulates physical processes and network data, creating realistic datasets and experiments that closely mirror real-world cybersecurity scenarios. This twin can effectively demonstrate attacks that are indistinguishable in network traffic yet produce vastly different physical outcomes.

Further expanding the scope of digital twins in cybersecurity, Nguyen [48] suggests human digital twins (HDT) as an innovative tool that integrates behavioral psychology with cybersecurity and simulates human behavior in cyber–physical systems. This tool significantly enhances the ability to predict and analyze adversary behaviors and tactics in cybersecurity contexts. The introduction of HDTs in cybersecurity simulations demonstrates an advanced application of simulation technology, offering new pathways for proactive defense strategies and embodying Hartmann's vision of simulations as a source of inspiration in scientific research.

It is crucial to understand that they are not mutually exclusive and often overlap in practical applications. The example of HDTs in cybersecurity showcases this effectively, blending discovery, boundary object functionality, and inspiration within a single simulation framework. HDTs not only inspire new methods and tools for cybersecurity, leading to innovative approaches in both defense and educational strategies, but also facilitate the discovery of human behavioral patterns in cyber systems, enhancing risk assessment and decision-making. Additionally, they serve as boundary objects by bridging the gap between human psychology and cybersecurity.

This overlap highlights the multifaceted nature and robust capabilities of simulations and underscores their influential role in comprehending complex digital ecosystems, as noted by Briscoe et al. [49]. In cybersecurity, as demonstrated by our examples, simulations enable the modeling of cyber attacks, defense mechanisms, and the dynamics between various cybersecurity elements such as human factors, technology, and organizational processes. This capability is essential for hypothesis testing, exploring potential security breach impacts, and evaluating countermeasures. Kavak et al. [25] identified five key research areas where simulation plays a pivotal role in cybersecurity. These areas are (i) representative environment building; (ii) test, evaluate, and explore; (iii) training and exercises; (iv) risk analysis and assessment; and (v) examining the role of people in the cybersecurity domain.

These areas collectively form a comprehensive approach to using simulations in cybersecurity. Given the unpredictable nature of cybersecurity, these areas often deal with scenarios where outcomes and probabilities are uncertain or unknown, such as unpredictable situations and ambiguous aspects, which pose challenges to informed decision-making. The utilization of simulation, therefore, becomes a vital tool in managing these challenges, offering various approaches and benefits for theoretical and empirical research and analysis in the cybersecurity domain. However, two critical questions arise: what specific insights do we gain from cybersecurity simulations, and how reliable and robust are those insights? While the latter is extensively addressed through studies on verification and validation methods [50,51], such as sensitivity analysis[4] [54,55], the former inquiry is a subset of a broader question concerning the knowledge we can derive from simulations in general. To examine this inquiry, we continue this section with a deeper exploration of two specific mechanisms through which simulations act as a source of new knowledge.

New knowledge from simulations

As pointed out by Herbert Simon ([24], pp. 14–17), since a simulation is no better than the assumptions built into it, it may seem that we could not learn anything new from it, but this is not the case. More precisely, Simon offers two ways in which simulations can be a source of new knowledge:[5]

Bookkeeping: Simon calls the first way obvious, and explains that "even when we have correct premises, it may be very difficult to discover what they imply. All correct reasoning is a grand system of tautologies, but only God can make direct use of that fact." ([24], p. 15). This is the kind of situation where we essentially know how the relevant phenomena work, but we cannot keep track of everything at once. Here, models and simulations become indispensable tools, as they allow us to systematically process and analyze extensive data, thereby uncovering the implications of interacting variables from intricate initial conditions. In essence, while we understand the basic mechanics, it is through computational methods that we can effectively manage and interpret large-scale phenomena.

A good cybersecurity example is a report by Lloyd's [56] investigating the consequences of service outages at major global cloud service providers such as Amazon, Google, and Azure. The method is essentially to use a database of which companies use which cloud service providers, use company turnover to estimate the losses incurred by an outage, sum the resulting losses, and use more statistics to see whether these losses are insured (for an introduction to how cyber-insurance of outages works, see [57], especially Section 4.6.). Conceptually, this is just simple arithmetic. The contribution of this simulation is to actually work out all the details, using credible statistics to find the right numbers. Another example is simulations using attack graphs to conduct a risk assessment of systems (for a few variations on this theme, see e.g. [58–60]). Again, the attack graph concept is relatively straightforward—essentially it is just using first-order logic to derive which attacks are possible on a specified system. The contributions of the simulations consist of actually building (vast) attack graphs (more or less automatically) for real systems and working out implications, such as which assets are the most vulnerable, which defenses are the most effective, and so on.

Abstraction: Simon calls the second way "more interesting and subtle" [24], p. 15. The key insight here is that we rarely need to simulate *all* the properties of systems; "we are usually interested only in a few properties abstracted from the complex reality." [24], p. 15. Many artificial systems are "particularly susceptible to simulation via simplified models" ([24], p. 16) because they are built to work in a certain way. A clock has to be able to tell time, so to simulate a clock, we do not have to simulate gears, springs, or pendulums, we just have to simulate time (pp. 5–6). A company has to be able to produce and sell things, so to simulate a company, we do not have to simulate factories, workers, and CEOs, we just have to simulate incomes, costs, and adaptive decision-making ([24], p. 8, p. 12). Simon's key point from these examples is that models and simulations can provide new knowledge because the ability to abstract means we do not need to have a complete understanding of every details of a system's workings in order to construct a viable model. Abstraction allows models to concentrate on pertinent aspects, avoid unnecessary complexity, and remain tractable.

In cybersecurity, Carfora and Orlando [61] is a good example, illustrating how very complicated data breaches, which in actuality included all kinds of technological, economic, and organizational subtleties can be usefully abstracted into just frequency data (number of breaches per day) and severity data (number of records breached), the statistics of which are investigated in an illuminating way. The same goes for many other numerical examples and simulations using investment models such as that by Gordon and Loeb (the original analytical model is given in [31]; some examples of simulations building on it include [62–64]), underpinned by precisely the kind of abstraction Simon talks about. Such models become useful precisely because they abstract many irrelevant details away, leaving only exactly what is relevant for the problem at hand. If done properly, very simple models can thus offer new insights about a system, particularly "if the aspects in which we are interested arise out of the *organization* of the parts, independently of all but a few properties of the individual components" ([24], p. 17), emphasis in original.

These two ways—bookkeeping and abstraction—represent different yet complementary approaches to gaining new knowledge and insights through the use of simulations. However, Simon's two ways also offer hints about how they could go wrong. Whereas the models

4 To ensure that simulations are robust and that approximations or assumptions do not lead to misleading conclusions, they must be rigorously verified and validated. A key method to achieve this is through sensitivity analysis. Sensitivity analysis is crucial for identifying, which variables most significantly impact the outcomes, allowing model designers and users to identify the most critical data points and better understand the shape of the problem space. For instance, Feng et al. [52], employed sensitivity analysis to assess the robustness of security investment strategies within managed security service providers (MSSPs). The study utilized a system dynamics model to simulate various investment strategies under different types of cyber-attacks, such as opportunistic and targeted attacks. Sensitivity analysis in this context helped reveal that investments in preventive measures had a stronger and more sustained impact on the business value of MSSPs compared to investments in detection and response strategies. This analysis validated the model's ability to reliably predict outcomes under varying conditions, thus providing actionable insights for MSSPs to optimize their security investments. Similarly, Behara et al. [53] demonstrated the importance of sensitivity analysis in their model, which was designed to simulate information security investments across different stages of the information security life-cycle. By applying sensitivity analysis, they were able to validate the stability and logical structure of their model, ensuring that the simulation results were consistent with real-world behaviors.

5 Simon does not himself use the term "bookkeeping"—or any other term—to describe his first road to knowledge. This label is ours. While Simon also does not use the term "abstraction" as a label, he does use it several times when describing his second road to knowledge, inspiring our use of the term as a label here.

underpinning simulations are accurate in benign environments, they may not be accurate in more taxing environments:

> A bridge, under its usual conditions of service, behaves simply as a relatively smooth level surface on which vehicles can move. Only when it has been overloaded do we learn the physical properties of the materials from which it is built ([24], p. 13).

Inspired by this method, we now proceed to investigate the failure modes of the two ways of learning from simulations—to find the taxing environments that challenge them.

How bookkeeping can fail

Bookkeeping fails if the books kept contain errors. To be sure, under benign conditions, errors may even out and cancel. Revisiting the simulation of insured cloud service downtime cost [56], it is reasonable to believe that numbers such as company turnover figures or insurance rates are not *exact*. If a large cloud service provider were to go down for a few days, company turnovers would of course not be *exactly* as reported in the last quarterly report, and the number of insured companies would not be *exactly* as reported in the last quarterly reports of the insurers, or filed with a regulator. However, if markets have not changed dramatically since the last measurement, these errors may even out and cancel: some companies have a larger turnover, some have a smaller; some companies are no longer insured, some have taken up insurance (this may be seen as robust external validity [39], Chapter 4: results are to some extent generalizable beyond the original population of companies). Something similar holds for the attack graphs [58–60]. The graphs may not be exactly accurate representations of the systems modeled (and addressing this uncertainty is indeed an important feature of some such work, see [65,66]) and the vulnerabilities incorporated in the model may not be exactly accurate representations of the vulnerabilities there are (that there may be false negatives is obvious, since new vulnerabilities are continuously discovered, but there may be false positives as well, both since some may have been patched away and since some may have been spurious in the first place). While on a fine-grained level such errors cannot exactly cancel, a simulation model aiming to find the most vulnerable assets or the most effective defenses may still give the same, reasonably accurate, answers even if some details are wrong.

But while *random* errors may cancel, *systematic* errors do not and such errors may thus constitute taxing conditions for bookkeeping. The simulation of insured cloud service downtime cost [56] may fail if there is an economic boom or bust pushing the turnovers of many companies in the same direction or if there is a rapid change to the cyber-insurance market (as has been the case in recent years, see, e.g. [67,68]), pushing the numbers of insured companies, or the terms they are offered, in systematic direction. Similarly, attack graph simulations may fail if new vulnerabilities affect not just a few components in an architecture, but many components at once, such as in the highly publicized cases of the Heartbleed (see, e.g. Zhang [69]) and Log4J (see, e.g. Srinivasa [70]) vulnerabilities.

More generally, it is well-known that much data relevant to cybersecurity economics—in particular, statistics on incident properties such as costs—may contain systematic errors. Surveys—the most popular measurement instrument in the absence of disclosure—are unreliable for several reasons [71,72]: first, there may be a systematic bias among respondents, where those affected may be more likely to respond than those not affected. Second, overestimates from individual respondents can have an excessive impact on the inferred averages if there is no way to check the plausibility of the responses. Third, estimating costs can be difficult even if one tries in earnest. Fourth, incentives are skewed in the sense that many who conduct incident cost studies have their own agendas [15].

Such problems do not pertain only to surveys. Getting truly representative samples in cybersecurity is difficult and requires careful deliberation (see, e.g. Metcalf and Spring [39], pp. 101–113). For, example, data from honeypots may be biased toward the less sophisticated attackers [73], penetration testers may bias data from their tests [74], and while social media can be used as a crowdsourced sensor to detect and characterize ongoing cyber-attacks [75] or new software vulnerabilities [76], such sources no doubt also exhibit their own systematic errors. Importantly, in our economic context, a source may be fine for some purposes (e.g. patching a particular vulnerability found) but not for others (e.g. drawing inferences about statistical characteristics of future vulnerabilities).

Another reason for the prevalence of systematic errors is that it is difficult to accurately estimate the statistical properties (e.g. means, medians, and variances) of very rare events. [77]. One obvious way in which this can happen is that sample sizes are too small to include the rarest event—too small to accurately represent the tails. Of course, there are ways to deal with this. Carfora and Orlando [61] represent one such way, when they show how the data on probabilities and consequences of breaches can be modeled in many ways: an empirical simulation assuming no models, an historical simulation assuming a statistical model of frequency (a negative binomial distribution), or a Monte Carlo simulation assuming models of both frequency and severity (a skew-normal distribution). The important insight offered is that such assumptions matter: the empirical simulations underestimate the risks compared to the more sophisticated methods, confirming the need for robust estimations of the full distributions involved in cyber risk. In short, Carfora use simulations to show how sensitive simulations can be to modeling assumptions. There is also a wealth of statistical resampling methods such as jackknife and bootstrap, which can be used to cleverly create new samples based on the observed one. Still, such methods are not perfect. Woods and Böhme [12] offer an illuminating observation of these difficulties in a frustrated comment on the statistics of data breaches:

> In 2016, Edwards et al. [77] estimated that the probability of seeing a breach of 200 million or more records in the next 3 years had a probability of around 0.1. Wheatley et al. [78] derived a maximum breach size of 200 million, growing by 50% in the 5 years following 2016. Yahoo! reported the loss of 3 billion customer records in the same year as both publications (albeit lost years earlier). What do we really know about data breaches when even methods designed for tail events like extreme value theory [78] set bounds that are exceeded by an order of magnitude within the same year (with multiple breaches exceeding 500 million in the last 3 years)? [12].

Subtle errors in what we call bookkeeping, thus blunt otherwise sophisticated models. One domain where substantial efforts, from governments and the private sector alike, are devoted to rigorously govern risks using state-of-the art models is the financial sector. Yet, there are indications that these methods systematically underestimate cyber risks [79,80]. If—when—this is indeed the case, we find ourselves in a situation where we thought that we were making decisions under risk, but we are closer to making them under uncertainty.

How abstraction can fail

Abstraction fails when a model is applied in contexts where its assumptions do not hold. There are plenty of such examples from natural sciences, where discussions about simplifications and the explana-

tory domains of models often are uncontroversial. For example, the ideal gas law works great for monatomic gases at low pressure and high temperature. Under these conditions, it does not matter that it neglects the size of the molecules or their interactions with each other—most of the gas is empty space. But with larger gas molecules, higher pressure and lower temperatures, molecular size, and interaction matters, and the (abstract simplification of the) ideal gas law no longer works. This is often explained at some length in physics and chemistry textbooks (see, e.g. Atkins and Jones [81], pp. 188–209 and Bowley and Sánchez [82], pp. 9–10), presumably not only in order to teach students something about the behavior of gases, but also to explain that all scientific models contain simplified assumptions, which make them appropriate in some circumstances, but inappropriate in others. The key to using such a model wisely is to understand the difference between these cases. The same point is emphasized by Giere [83], who argues that we should look not only at the dyadic relationship between the model and the world, but also include the purpose of the model: "Scientists use models to represent aspects of the world for specific purposes".

Cybersecurity models and simulations depend on many models which are similar to the ideal gas law. The following list is surely not exhaustive, but it aims to capture at least some common modeling abstractions—from technology and economics alike—which are appropriate in some contexts, but inappropriate in others:

Events being statistically independent is a common assumption in a wide range of models.

- In their simplest form, reliability engineering techniques such as fault trees assume that failure events are statistically independent. It is also possible to model events, which are not statistically independent—common cause events—but this is much more difficult. Oftentimes the independence assumption is reasonable ([84] is an example where the independence assumption seems to fit the empirical data), but when it is not, the model fails. A word of caution is given by [85], who believe that software fault trees will never be as precise as hardware fault trees precisely for this reason. A concrete cybersecurity example is two-factor authentication. Sometimes it may be reasonable to assume that compromise of the two factors are truly independent (because obtaining one factor is independent obtaining another), but a more detailed model would also include cases when it is not (e.g. a post-it note with a pin code on a smart card, so that obtaining the two factors are in fact highly correlated).
- In their simplest forms, insurance models assume that loss events are statistically independent. Just like in the fault tree case, it is also possible to model events which are not statistically independent (for an introduction see, e.g. Bahnemann [86], pp. 92–98), but again, this is more difficult. While for cyber insurance the independence assumption may sometimes be appropriate, it is well-known that the accumulation risks stemming from the fact that some cyber events are *not* independent is a major impediment to the cyber insurance industry [87], and as a result, models that assume independence may be too simplistic in other circumstances. (There are also models which retain some independence assumptions within models built to capture some element of overall but still attempt to capture accumulation risk and claims contagion; see [88,89].)

Attitudes to risk directly impact economic cybersecurity decision-making.

- In some common models, such as the Gordon–Loeb model [31], firms are risk-neutral. In other models, such as the cyber-insurance framework of Ogut et al. [90], risk-averse firms buy insurance from risk-neutral insurance companies. Empirical studies of practitioners suggest that they are risk-averse [91] or follow a variety of patterns [92]. Different models may be appropriate in different situations. Another complication is that large organizations may have different attitudes to risk in different departments, e.g. if an IT operations department is measured on availability (up-time) whereas a security department is measured on confidentiality and integrity. Sometimes modeling and simulation can help understand the trade-offs for the different teams or at least help explain the values—but then such complications need to be explicitly taken into consideration.
- As remarked above, insurers are often seen as risk-neutral in one-tier models, where risk-averse insureds by insurance from them. Sometimes, this is appropriate, and indeed, under idealized assumptions such as perfectly competitive insurance markets, it can be proven that insurers are risk-neutral [93]. But real insurance markets are not perfectly competitive—creating the need for reinsurance—and so in other circumstances, insurers should be modeled as risk-averse.

Equilibria are often studied in different kinds of models.

In microeconomic models, market equilibria are useful to characterize situations where economic forces such as supply and demand are balanced and will not change in the absence of external events. For example, under perfect competition, supply equals demand at the market price ([94], p. 219). On such markets, firms are price-takers, i.e. they have so small market shares that their behavior does not affect the overall market prices. But of course, competition is not always perfect, some markets are oligopolistic, and external events do happen, so equilibrium models are not always appropriate.

In reliability engineering, equilibria and steady states are used to characterize many different situations relevant for cybersecurity [95], such as reliability growth of software, loads in telecommunications networks, IT service availability, vulnerability arrival and patching rates, and so on. While models can be illuminating, for instance about average conditions, they do not shed light on nonequilibrium situations and transients between equilibria, which may give a poor appreciation of the variance of those average conditions [96]. For example, an equilibrium between vulnerability arrival and patching rates may be upset by the sudden arrival of more manpower, new technology, new legal requirements, and so on.

Perfect information is often a useful first approximation.

In many models, cyber incidents are immediately and perfectly detected. In some contexts, this is a reasonable assumption—availability incidents, for example, typically have immediate and obvious consequences. However, it is also well-known that stealthy cyber attacks may go unnoticed for a long time (see, e.g. [97]). What is assumed here has immediate consequences in many models. For example, if incidents are perfectly detected, risk-averse insureds buy full insurance protection, but if incidents are not perfectly detected, insureds buy less than full insurance [90].

It is often assumed that mandatory cyber incident reporting to governments would solve—or at least substantially alleviate—the information problem that incidents are unknown to relevant actors [15], Section 4.1.2. This is reasonable as a first approximation. However, empirical evidence from mandatory Network and Information Systems (NIS) reports suggests that the data quality of

such reporting may be so poor that the effect may be very small [98].

A useful perspective on poor cybersecurity is to look at it as an externality [7,99]. Thus, some models may assume that the imposition of Pigovian cyber taxes could get rid of the externality (and thus, of underinvestment in cybersecurity). However, it is also well-known that though the Pigovian tax is theoretically elegant, it requires the taxing authority to know exactly how the externality cost function looks, which is of course not the case in practice ([94], p. 434). Thus, a more realistic model would rather let the taxing authority impose an *imperfect* Pigovian tax, which may or may not match the actual externality. Similar reasoning holds for other fines and sanctions, such as those required for breach notification laws to work [100].

The common factor in all these examples is that modeling assumptions that may be reasonable and appropriate in some contexts (benign environments, in Simon's terminology) may be unreasonable and inappropriate in other contexts (taxing environments). Appropriate use of simulations demands that attention is paid to making the right assumptions in the right contexts—otherwise, abstraction will fail.

Why bookkeeping and abstraction can fail

In the previous sections, we have discussed *how* bookkeeping and abstraction can fail. But a key to avoiding such pitfalls is also to understand *why* they may fail. Not claiming to be exhaustive, the following list identifies some factors which exacerbate the risk of failure in cybersecurity simulation models used for decision-making:

Multidisciplinarity challenges:

- Accurately modeling the complex interaction of technical, human, and organizational elements in cybersecurity presents a significant challenge (e.g. with respect to assumptions about statistical independence and distributions, the behavior of actors and markets, or the applicability of equilibrium models, as discussed above). This complexity arises from the need to simultaneously consider the technical aspects of cybersecurity and the human behaviors and organizational dynamics that influence it. Moreover, modeling this interaction demands a balance between capturing the complexity of real-world interactions and maintaining the usability of the simulation models.
- The need to integrate diverse disciplines—particularly at the policy level, where all the different perspectives need to come together into a unified whole—heightens the risk of misunderstanding and misapplication in simulations. In general, building simulation models across disciplines is a challenging task, as it may require economists, for example, to use and understand technical models, and technical researchers to use and understand economic models. Compared to simulation models residing within single disciplines, this exacerbates the risk of failure, especially in the abstraction sense.

***Scarcity of quality data*:** limited, poor-quality, or biased data on cyber events [12–14] increases the risk of inaccurate simulations, especially in the bookkeeping sense. One difficulty, which persists even with mandatory reporting, is *survival bias*, where data from businesses which go bankrupt may be missing.

***Heavy-tailed distributions*:** here is a serious risk that rare but high-impact cyber events are underrepresented or misinterpreted in empirical data used for simulations.

***Adaptability challenges*:** difficulty in updating simulations to reflect the latest threat vectors, technological changes, and institutional transformations.

***Lack of transparency and reproducibility*:** inadequate documentation of the assumptions, methods, and mechanics of simulation models, hindering peer review and replication.

***Development in isolation*:** the process of creating simulation models is conducted without adequate collaboration or input from relevant stakeholders and users of the model. This can lead to significant gaps in the applicability and effectiveness of models, particularly when they fail to account for the complexities and nuances of the actual environment in which they are intended to be applied. As demonstrated by [101,102], the inclusion of diverse perspectives and expert inputs—such as those from security operations teams—into the modeling process and iteratively refining models ensure that the models remain relevant and practical for real-world decision-making.

Though this list could probably be made longer, it offers some illuminating starting points for how to manage the perils of cybersecurity simulations for decision-making. In particular, it is worth noting with respect to the first item on the list, multidisciplinarity, that simulations which are limited in scope to single disciplines may beg the question at hand. Thus the multidisciplinarity can often not be *avoided*, but rather has to be properly *managed*.

Similar reflections can be made for the other items (e.g. that using only solid empirical data rather than simulated may defeat the very purpose of the simulation). Indeed, the factors listed overlap significantly with the very benefits that make simulations an attractive tool in the first place. This observation shows the subtle complexity inherent in the use of simulations for cybersecurity. Recognizing this dual nature——where strengths can also manifest as risks——is crucial in navigating the use of simulations effectively.

A framework of considerations

To make the most of simulations for economic cybersecurity decision-making and to avoid the pitfalls, we propose a framework comprised of a series of organized questions intended for consideration *before* deploying a simulation model in decision-making processes. Unlike a rigid checklist that requires affirmative answer to every question, this framework is designed to prompt critical thinking and thoughtful evaluation. It acknowledges that not all questions will receive positive responses, and that this is acceptable and expected. In fact, recognizing and documenting negative responses is a vital part of understating the model's limitations. Documentation, transparency, impact assessment, and communication with users ensures that the model's limitations are understood and prevents misuse. The George Box quote. "all models are wrong, but some are useful", is particularly relevant here, as it underscores the importance of using models as tools for insight rather as infallible predictors. Hence, this framework functions as a guide to help model designers and and users[6] understand the strength, limitations, and applicability of the model in a specific context. The framework is a tool to aid in determining the appropriateness of a simulation, ensuring it is utilized

6 While this framework is primarily aimed at model designers during the development phase, it is also valuable for users of the model—such as system stakeholders or decision-makers—before deployment. These users can engage with the framework to understand the model's assumptions, limitations, and relevance to their specific context. By doing so, they can make informed decisions about whether the simulation model is appropriate for their needs and how to interpret its results effectively.

Table 1. A set of guiding questions to be considered before employing a simulation model in economic cybersecurity decision-making processes. The table is a summary of the preceding sections, where more details are found.

Upholding data integrity and avoiding bias	• Is the data representative and collected using robust methods? • Are there biases in the data collection process? • How does the model handle incomplete or uncertain data? • What mechanisms are in place to update data sources and ensure ongoing relevance?
Precision and scope of data	• Does the data align with the scope of the decision problem? • Is the data granular enough for the specific aspects being simulated? • Are there processes to periodically reassess the relevance and accuracy of the data? • How does the model accommodate data from different and potentially conflicting sources?
Assumptions about statistical distributions and independence	• Are the assumptions about distributions and event independence justified? • How do these assumptions impact the simulation outcomes? • Is there a process for regularly reviewing and updating these assumptions? • How are outliers and anomalies in data handled in the model? • Does the model support causal or evidential decision-making? Is the model transparent enough to answer this question?
Modeling asymmetries and actor behavior	• Are there unjustified asymmetries in how different actors or market dynamics are modeled? • Is the model's representation of actor behavior (risk-neutral, risk-averse) appropriate? • Does the model account for potential changes in actor behavior or preferences over time? • Are there considerations for unexpected or sudden market changes?
Equilibrium and dynamic modeling	• Does the model appropriately account for nonequilibrium conditions and transitions? • Are steady-state assumptions justified?
Information assumptions	• Are the assumptions about the availability and accuracy of information within the model realistic? • How do these assumptions affect the simulation results?
Transparency and reproducibility	• Is the model and its underlying mechanics documented in a way that allows for reproducibility? • Are the sources of data and the methodologies used clearly stated? • Are there guidelines for interpreting the results of the model?
Contextual appropriateness	• Is the level of abstraction appropriate for the cybersecurity context being simulated? • Does the model account for the specificities of the cyber ecosystem under study? • How does the model integrate interdisciplinary knowledge (e.g. technological, sociological, and psychological)?
Adaptability and updating	• Can the model be easily updated to reflect new data or changes in the cyber threat landscape? • How flexible is the model in adapting to new scenarios or information? • Is there a feedback mechanism for users to suggest improvements or report issues?
Validation and testing	• Has the model been validated against real-world data or scenarios? • Are there mechanisms in place for continuous testing and improvement of the model?
Deployment	• Is the model's readiness and relevance for deployment and operational use assessed? • Are the model's limitations and their impacts documented and communicated to the users of the model? • What training and knowledge transfer is required for model's users to effectively interpret and apply model's results? How will this training be delivered?

with good judgment, especially when handling any identified limitations or negative responses.

Table 1 presents this framework, categorized into 11 distinct yet interrelated components, designed to prompt careful consideration and evaluation of various facets of simulation modeling, from the initial stages of data collection to the final stages of model implementation and review. While this framework provides a broad-strokes outline, it covers a range of aspects essential to advance and improve the use of simulations in economic cybersecurity decision-making. It is designed as a foundational guide to address the key areas, such as data quality and management, modeling techniques, multidisciplinarity, and iterative improvements, that are crucial for enhancing the effectiveness, accuracy, and applicability of simulations. By methodically applying this framework, we can more effectively utilize the potential of simulations while concurrently addressing and mitigating the inherent risks associated with their use. This structured approach, which emphasizes iterative feedback loops and collaboration between modelers and users in different categories, significantly enhances the reliability and validity of simulation outcomes. By incorporating continuous validation and expert input, the framework ensures that simulations are not only theoretically sound but also practically applicable and directly relevant to economic cybersecurity decision-making processes.

Discussion and conclusion

Our investigation into the use of simulations in economic cybersecurity decision-making reveals both promising avenues and potential pitfalls. Simulations offer the possibility of creating rich, detailed datasets that can significantly mitigate the current scarcity of high-quality cybersecurity data. By providing a controlled environment, simulations allow for thorough examination of cyber threats, vulnerabilities, and the effectiveness of defense strategies. One of the key strengths of simulations lies in their capacity to incorporate multiple perspectives, fostering multidisciplinary, multiparadigmatic dialogues and experiments. This approach allows for a broader and more intricate understanding of cybersecurity challenges, integrating technical, human, and organizational factors into a cohesive analysis. It offers a platform for testing hypotheses and exploring scenarios that may be impractical or impossible to replicate in real-world set-

tings due to ethical, logistical, or financial constraints. The discussion in the previous sections illustrates this potential, in particular with respect to providing the information required for decision-making. The studies cited offer a set of good examples of how economic cybersecurity decision-making can be improved and made more robust, using simulations to fill information gaps as needed to navigate the complex landscape of cyber threats.

However, the effective use of simulations also demands an appropriate awareness of their limitations and potential risks. When we use cybersecurity simulations for representation, prediction, and explanation *without* an aspect of policy decision-making (to use the list of applications from Grüne-Yanoff and Weirich [36]), the stakes are limited. Of course, it is bad if we misrepresent how things work, if our predictions are off the mark, and if our explanations are not the best.[7] But it is worse—possibly, much worse—if, for example, misrepresentations of how things work lead us to impose unrealistic requirements or undeserved fines on some actors, erroneous predictions lead us to policy responses, which cause overshoots or oscillations (to borrow imagery and terminology from control theory), or if faulty explanations lead us to ignore potential policy tools. In short, the higher stakes connected to real world implications are the difference between the academic seminar room on the one hand and the corporate boardroom or national government offices on the other.

These higher stakes are part of the motivation for using cybersecurity simulations in the first place. If the stakes are so high, must we not use all possible means to gain knowledge so as to make the best decisions? But the higher stakes are also part of the motivation for treating cybersecurity simulation results with caution. If the stakes are so high, must we not discard dubious knowledge from the basis upon which we make decisions? The intuitive answers to both questions are "yes", but as our reasoning has illustrated, there is a tension between the two affirmations. For example, some simulations may, at first sight, seem to position us firmly in a decision under risk, but with hindsight, it may be that we were, still, actually closer to a decision under uncertainty.

An interesting practical example of a policy seemingly intended to force extra prudence in decision-making is the Swedish rules for protective security [104]. Protective security (Swedish *säkerhetsskydd*) refers to protective measures against espionage, sabotage, terrorism, and other threats against national security. At the heart of this work is a protective security analysis process, and interestingly, the instructions for how to conduct such an analysis explicitly mandates that *only consequences, not probabilities*, must be taken into account. Due to the high stakes involved in national security, it has been decided that protective security must not be treated as a decision under risk, but should instead be treated as a decision under uncertainty. Of course, there are many conceivable reasons for this. But one such reason—arguably a good one—is that decision-makers may be overconfident about how good their probability estimates are (e.g. because these estimates come from simulations that seem credible). Such overconfidence may engender more risky behavior than intended: a decision which looks risk-neutral (or mildly risk-averse) from the perspective of the consequences and probabilities thought to be known may in fact be risk-seeking (or at least less risk-averse than intended) from the perspective of the actual consequences and probabilities. Another possible reason for this policy, pointed out by an anonymous reviewer, is that in protective security there is an intelligent adversary, whereas in many safety contexts, there is not. Whether a natural disaster will topple a bridge is not analyzed the same way as whether terrorists will blow it up (though in practice, this distinction is not always upheld [105]). With adversaries, the analysis typically focuses on their capability, access, and intent. If an adversary has all three of those, probability of attack is by definition 1. Otherwise, it is 0. On this interpretation, taking only consequences, but not their probabilities, into account simply amounts to assuming there are motivated adversaries who will bring those consequences about if they can. Sometimes, this approach makes perfect sense. The US Cybersecurity and Infrastructure Security Agency maintains a Known Exploited Vulnerability (KEV) catalog.[8] Since the vulnerabilities listed there are known to have been exploited, it is prudent to assume that they will be exploited again with a probability of 1. But observe that disregarding probabilities would prevent us from patching the vulnerabilities listed in the KEV catalog before patching vulnerabilities not listed there (at least if doing so was based on their higher probability of exploitation). And observe also that though this may seem like a bad idea, there is some logic to it, since an intelligent adversary may deliberately go for vulnerabilities not listed in the KEV catalog, if the strategy of KEV-guided patching is known. (This may be rather be an argument for keeping strategies secret than for disregarding probabilities.)

In conclusion, by advancing the discourse on the effective use of simulations, we have emphasized their role in generating new knowledge about probabilities and consequences in the cybersecurity domain. This is a crucial aspect in our understanding of risk and uncertainty, and significantly influences the decision-making process. This exploration specifically contributes to the broader field of risk management in cybersecurity, where understanding and differentiating situations of risk and uncertainty are vital. In such a complex field, recognizing and appropriately responding to these distinct contexts is key to sound decisions and effective strategy formulation and implementation.

Furthermore, our discussion around the concepts of "bookkeeping" and "abstraction" in simulations highlights their critical roles in ensuring that simulations provide meaningful, actionable insights that can inform policy and strategy. Bookkeeping, with its focus on the detailed and accurate representation of data, ensures that simulations are grounded in reality and reflect the true complexity of the cybersecurity environment. Abstraction, on the other hand, allows us to simplify this complexity into manageable models, making it possible to explore various scenarios and outcomes without getting lost in the minutiae. Together, these concepts facilitate a balanced approach to simulation, enabling us to model cybersecurity systems in ways that are both comprehensive and comprehensible.

This paper, therefore, not only sheds light on the multifaceted nature of simulations in cybersecurity but also provides guidance on their thoughtful and effective application. By understanding and leveraging the strengths of simulations while being mindful of their limitations, we can significantly enhance our capabilities in cybersecurity risk management. As the cybersecurity landscape continues to evolve, our approach to simulations must also adapt, ensuring that they remain relevant and effective tools in our ongoing efforts to understand and mitigate cyber risks.

Building upon the foundations laid in this paper, future work could focus on the practical application, testing, and refinement of the framework we outlined in Section 3.5, moving beyond theoretical frameworks to their concrete implementations and evaluations in real-world scenarios and applications, such as cyber ranges, critical infrastructure simulations, and other practical cybersecurity simulation environments. Therefore, this study calls for a comprehensive

7 For a recent collection of essays on inference to the best explanation, see McCain and Poslon [103].

8 https://www.cisa.gov/known-exploited-vulnerabilities

and structured approach to incorporating these considerations into the life-cycle of cybersecurity simulations, encompassing their design, execution, and continuous refinement.

Author contributions

Mazaher Kianpour (Conceptualization, Data curation, Investigation, Methodology, Project administration, Visualization, Writing–original draft, Writing–review & editing), Ulrik Franke (Conceptualization, Data curation, Investigation, Methodology, Project administration, Visualization, Writing–original draft, Writing–review & editing)

Conflict of interest: None declared.

Funding

This work was supported by H2020-SU-DS02-2020, grant number 101020259 (M.K.) and the Swedish Foundation for Strategic Research, grant number SM22-0057 (U.F.). The funding for open access publication is covered by Norwegian University of Science and Technology (NTNU).

References

1. Lewallen J. Emerging technologies and problem definition uncertainty: the case of cybersecurity. *Regul Governance* 2021;**15**:1035–52.
2. Johnson P, Ekstedt M. The Tarpit – a general theory of software engineering. *Inform Software Tech* 2016;**70**:181–203.
3. Blažič BJ. The cybersecurity labour shortage in Europe: moving to a new concept for education and training. *Technol Soc* 2021;**67**:101769.
4. Bronk C, Conklin WA. Who's in charge and how does it work? US cybersecurity of critical infrastructure. *J Cyber Policy* 2022;**7**:155–174.
5. Lebek B, Uffen J, Neumann M. *et al.* Information security awareness and behavior: a theory-based literature review. *Manag Res Rev* 2014;**37**:1049–92.
6. Dutt V, Ahn YS, Gonzalez C. Cyber situation awareness: modeling detection of cyber attacks with instance-based learning theory. *Hum Fact* 2013;**55**:605–618.
7. Anderson R, Moore T. The economics of information security. *Science* 2006;**314**:610–613.
8. Falco G, Eling M, Jablanski D. *et al.* Cyber risk research impeded by disciplinary barriers. *Science* 2019;**366**:1066–1069.
9. Kianpour M, Kowalski SJ, Øverby H. Systematically understanding cybersecurity economics: a survey. *Sustainability* 2021;**13**:13677.
10. Moore T, Dynes S, Chang FR. Identifying how firms manage cybersecurity investment. In: *Workshop on the Economics of Information Security (WEIS)*. Dallas, TX, USA: Darwin Deason Institute for Cyber Security, Southern Methodist University, 2016, 1–27.
11. Grimm S. Understanding. In: Zalta EN (ed.), *The Stanford Encyclopedia of Philosophy*. Summer 2021 edn. Stanford, CA: Metaphysics Research Lab, Stanford University, 2021.
12. Woods DW, Böhme R. SoK: quantifying cyber risk. In: *Proceedings of the 2021 IEEE Symposium on Security and Privacy (SP)*. Piscataway, NJ: IEEE, 2021, 211–28.
13. Woods DW, Walter L. Reviewing estimates of cybercrime victimisation and cyber risk likelihood. In: *Proceedings of the 2022 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*. New York, NY: IEEE, 2022, 150–62.
14. Cremer F, Sheehan B, Fortmann M. *et al.* Cyber risk and cybersecurity: a systematic review of data availability. *Geneva Pap Risk Insur Iss Pract* 2022;**47**:698–736.
15. Moore T. The economics of cybersecurity: principles and policy options. *Int J Crit Infr Prot* 2010;**3**:103–17.
16. Rodin DN. The cybersecurity partnership: a proposal for cyberthreat information sharing between contractors and the federal government. *Public Contract Law J* 2015;**44**:505–28.
17. Skopik F, Settanni G, Fiedler R. A problem shared is a problem halved: a survey on the dimensions of collective cyber defense through security information sharing. *Comput Secur* 2016;**60**:154–76.
18. De Bruijn H, Janssen M. Building cybersecurity awareness: the need for evidence-based framing strategies. *Gov Inform Quart* 2017;**34**:1–7.
19. Dean B, McDermott R. A research agenda to improve decision making in cyber security policy. *Penn State J Law Int Aff* 2017;**5**:29.
20. Valeriano B. The need for cybersecurity data and metrics: empirically assessing cyberthreat. *J Cyber Policy* 2022;**7**:140–54.
21. Woods DW, Seymour S. Evidence-based cybersecurity policy? A meta-review of security control effectiveness. *J Cyber Policy* 2023;**8**:1–19.
22. Humphreys P. *Extending Ourselves: Computational Science, Empiricism, and Scientific Method*. Oxford: Oxford University Press, 2004.
23. Axelrod R. Advancing the art of simulation in the social sciences. In: *Simulating Social Phenomena*. New York, NY: Springer, 1997, 21–40.
24. Simon HA. *The Sciences of the Artificial*. 3rd edn. Cambridge, MA: The MIT Press, 1996.
25. Kavak H, Padilla JJ, Vernon-Bido D. *et al.* Simulation for cybersecurity: state of the art and future directions. *J Cybersecur* 2021;**7**:tyab005.
26. Shreeve B, Gralha C, Rashid A. *et al.* Making sense of the unknown: how managers make cyber security decisions. *ACM T Softw Eng Meth* 2023;**32**:1–33.
27. Shreeve B, Hallett J, Edwards M. *et al.* "So If Mr Blue Head Here Clicks the Link..." risk thinking in cyber security decision making. *ACM Trans Priv Secur* 2020;**24**:1–29.
28. Resnik MD. *Choices: An introduction to Decision Theory*. Minneapolis, MN: University of Minnesota Press, 1987.
29. Knight FH. *Risk, Uncertainty and Profit*. Vol. **31**. Boston, MA: Houghton Mifflin, 1921.
30. Stirling A. On the economics and analysis of diversity. *Sci Pol Res Unit* 1998;**28**:1–156.
31. Gordon LA, Loeb MP. The economics of information security investment. *ACM Trans Inf Syst Secur* 2002;**5**:438–57.
32. Chai S, Kim M, Rao HR. Firms' information security investment decisions: stock market evidence of investors' behavior. *Decis Support Syst* 2011;**50**:651–61.
33. Wagg D, Worden K, Barthorpe R. *et al.* Digital twins: state-of-the-art and future directions for modeling and simulation in engineering dynamics applications. *ASCE-ASME J Risk Uncertainty Eng Syst Part B Mech Eng* 2020;**6**:030901.
34. Bankes SC, Lempert RJ, Popper SW. Computer-assisted reasoning. *Comput Sci Eng* 2001;**3**:71–7.
35. Winsberg E. Computer Simulations in Science. In: Zalta EN, Nodelman U (eds), *The Stanford Encyclopedia of Philosophy*. Winter 2022 edn. Stanford, CA: Metaphysics Research Lab, Stanford University, 2022.
36. Grüne-Yanoff T, Weirich P. The philosophy and epistemology of simulation: a review. *Simul Gam* 2010;**41**:20–50.
37. Beresnevichiene Y, Pym D, Shiu S. Decision support for systems security investment. In: *Proceedings of the 2010 IEEE/IFIP Network Operations and Management Symposium Workshops*. Piscataway, NJ: IEEE, 2010, 118–25.
38. Caulfield T, Pym D. Modelling and simulating systems security policy. *EAI Endorsed Trans Secur Saf* 2016;**3**:e3–e3.
39. Metcalf L, Spring J. *Using Science in Cybersecurity*. Singapore: World Scientific, 2021.
40. Star SL, Griesemer JR. Institutional ecology, translations' and boundary objects: amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Soc Stud Sci* 1989;**19**:387–420.
41. Tolk A, Harper A, Mustafee N. Hybrid models as transdisciplinary research enablers. *Eur J Oper Res* 2021;**291**:1075–90.
42. Luna-Reyes LF, Black LJ, Ran W. *et al.* Modeling and simulation as boundary objects to facilitate interdisciplinary research. *Syst Res Behav Sci* 2019;**36**:494–513.
43. Khan SK, Shiwakoti N, Stasinopoulos P. *et al.* Modelling cybersecurity regulations for automated vehicles. *Accident Anal Prev* 2023;**186**:107054.
44. Hartmann S. The world as a process: simulations in the natural and social sciences. In: *Modelling and Simulation in the Social Sciences from the Philosophy of Science Point of View*. New York, NY: Springer, 1996, 77–100.

45. Barricelli BR, Casiraghi E, Fogli D. A survey on digital twin: definitions, characteristics, applications, and design implications. *IEEE Access* 2019;**7**:167653–71.
46. Dietz M, Pernul G. Unleashing the digital twin's potential for ics security. *IEEE Secur Priv* 2020;**18**:20–7.
47. Murillo A, Taormina R, Tippenhauer N. *et al.* Co-simulating physical processes and network data for high-fidelity cyber-security experiments. In: *Proceedings of the Sixth Annual Industrial Control System Security (ICSS) Workshop*. New York, NY: ACM, 2020, 13–20.
48. Nguyen TN. Toward human digital twins for cybersecurity simulations on the metaverse: ontological and network science approach. *JMIRx Med* 2022;**3**:e33502.
49. Briscoe G, Sadedin S, De Wilde P. Digital ecosystems: ecosystem-oriented architectures. *Nat Comput* 2011;**10**:1143–94.
50. Bayarri MJ, Berger JO, Paulo R. *et al.* A framework for validation of computer models. *Technometrics* 2007;**49**:138–54.
51. Midgley D, Marks R, Kunchamwar D. Building and assurance of agent-based models: an example and challenge to the field. *J Bus Res* 2007;**60**:884–93.
52. Feng N, Wang M, Li M. *et al.* Effect of security investment strategy on the business value of managed security service providers. *Electron Commer Res Appl* 2019;**35**:100843.
53. Behara R, Huang CD, Hu Q. A system dynamics model of information security investments. In: *Proceedings of the Fifteenth European Conference on Information Systems (ECIS)*. AIS Electronic Library, 2007.
54. Borgonovo E, Plischke E. Sensitivity analysis: a review of recent advances. *Eur J Oper Res* 2016;**248**:869–87.
55. Christopher Frey H, Patil SR. Identification and review of sensitivity analysis methods. *Risk Anal* 2002;**22**:553–78.
56. Lloyd's Cloud Down: Impacts on the US economy. Technical report. London: Lloyd's of London, 2018.
57. Franke U. The cyber insurance market in Sweden. *Comput Secur* 2017;**68**:130–44.
58. Sheyner O, Haines J, Jha S. *et al.* Automated generation and analysis of attack graphs. In: *Proceedings 2002 IEEE Symposium on Security and Privacy*. Piscataway, NJ: IEEE, 2002, 273–84.
59. Wang L, Singhal A, Jajodia S. Toward measuring network security using attack graphs. In: *Proceedings of the 2007 ACM workshop on Quality of Protection*. New York, NY: ACM, 2007, 49–54.
60. Johnson P, Lagerström R, Ekstedt M. A meta language for threat modeling and attack simulations. In: *Proceedings of the 13th International Conference on Availability, Reliability and Security*. New York, NY: ACM, 2018, 1–8.
61. Carfora MF, Orlando A. Quantile based risk measures in cyber security. In: *Proceedings of the 2019 International Conference on Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*. Piscataway, NJ: IEEE, 2019, 1–4.
62. Naldi M, Flamini M. Calibration of the Gordon-Loeb models for the probability of security breaches. In: *Proceedings of the 2017 UKSim-AMSS 19th International Conference on Computer Modelling and Simulation (UKSim)*. Piscataway, NJ: IEEE, 2017, 135–40.
63. Gordon LA, Loeb MP, Zhou L. Integrating cost–benefit analysis into the NIST Cybersecurity Framework via the Gordon–Loeb Model. *J Cybersecur* 2020;**6**:tyaa005.
64. Skeoch HR. Expanding the Gordon–Loeb model to cyber-insurance. *Comput Secur* 2022;**112**:102533.
65. Hacks S, Katsikeas S, Ling E. *et al.* powerLang: a probabilistic attack simulation language for the power domain. *Energy Inform* 2020;**3**:1–17.
66. Katsikeas S, Johnsson P, Hacks S. *et al.* VehicleLang: a probabilistic modeling and simulation language for modern vehicle IT infrastructures. *Comput Secur* 2022;**117**:102705.
67. Mott G, Turner S, Nurse JR,. *et al.* Between a rock and a hard (ening) place: cyber insurance in the ransomware era. *Comput Secur* 2023;**128**:103162.
68. Woods DW. A turning point for cyber insurance. *Commun ACM* 2023;**66**:41–4.
69. Zhang L, Choffnes D, Levin D. *et al.* Analysis of SSL certificate reissues and revocations in the wake of Heartbleed. In: *Proceedings of the 2014 Conference on Internet Measurement Conference*. New York, NY: ACM, 2014, 489–502.
70. Srinivasa S, Pedersen JM, Vasilomanolakis E. Deceptive directories and "vulnerable" logs: a honeypot study of the LDAP and log4j attack landscape. In: *Proceedings of the 2022 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*. Piscataway, NJ: IEEE, 2022, 442–47.
71. Florêncio D, Herley C. Sex, lies and cyber-crime surveys. In: Schneier B (ed.). *Economics of Information Security and Privacy III*. New York, NY: Springer, 2013, 35–53.
72. Anderson R, Barton C, Böhme R. *et al.* Measuring the cost of cybercrime. In: Böhme R (ed.). *The Economics of Information Security and Privacy*. Berlin, Heidelberg: Springer, 2013, 265–300.
73. Morishita S, Hoizumi T, Ueno W. *et al.* Detect me if you...oh wait. An internet-wide view of self-revealing honeypots. In: *Proceedings of the 2019 IFIP/IEEE Symposium on Integrated Network and Service Management (IM)*. Piscataway, NJ: IEEE, 2019, 134–43.
74. Gutzwiller RS, Rheem H, Ferguson-Walter KJ. *et al.* Exploratory analysis of decision-making biases of professional red teamers in a cyber-attack dataset. *J Cognit Eng Decision Mak* 2024;**18**:37–51.
75. Khandpur RP, Ji T, Jan S. *et al.* Crowdsourcing cybersecurity: cyber attack detection using social media. In: *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*. New York, NY: ACM, 2017, 1049–57.
76. Sabottke C, Suciu O, Dumitra T. Vulnerability disclosure in the age of social media: exploiting Twitter for predicting {Real-World} exploits. In: *Proceedings of the 24th USENIX Security Symposium (USENIX Security 15)*. Berkeley, CA: USENIX, 2015, 1041–56.
77. Edwards B, Hofmeyr S, Forrest S. Hype and heavy tails: a closer look at data breaches. *J Cybersecur* 2016;**2**:3–14.
78. Wheatley S, Maillart T, Sornette D. The extreme risk of personal data breaches and the erosion of privacy. *Eur Phys J B* 2016;**89**:1–12.
79. Eling M, Schnell W. Capital requirements for cyber risk and cyber risk insurance: an analysis of Solvency II, the US risk-based capital standards, and the Swiss Solvency Test. *North Am Actuar J* 2020;**24**:370–92.
80. Peihani M. Regulation of cyber risk in the banking system: a Canadian case study. *J Financial Regul* 2022;**8**:139–61.
81. Atkins P, Jones L. *Chemistry: Molecules, Matter, and Change*. New York, NY: W. H. Freeman and Company, 1999.
82. Bowley R, Sánchez M. *Introductory Statistical Mechanics*. Oxford: Clarendon Press/Oxford University Press, 1999.
83. Giere RN. How models are used to represent reality. *Philos Sci* 2004;**71**:742–52.
84. Närman P, Franke U, König J. *et al.* Enterprise architecture availability analysis using fault trees and stakeholder interviews. *Enterp Inf Syst* 2014;**8**:1–25.
85. Leveson NG, Harvey PR. Software fault tree analysis. *J Syst Softw* 1983;**3**:173–81.
86. Bahnemann D. *Distributions for Actuaries*. CAS Monograph Series. Arlington County, VA: Casualty Actuarial Society, 2015. https://www.casact.org/sites/default/files/2021-02/02-Bahnemann.pdf. (10 November 2023, date last accessed).
87. Eling M, Schnell W. What do we know about cyber risk and cyber risk insurance?. *J Risk Finance* 2016;**17**:474–91.
88. Franke U, Draeger J. Two simple models of business interruption accumulation risk in cyber insurance. In: *Proceedings of the 2019 International Conference on Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*. Piscataway, NJ: IEEE, 2019, 1–7.
89. Hillairet C, Lopez O, d'Oultremont L. *et al.* Cyber-contagion model with network structure applied to insurance. *Insur Math Econ* 2022;**107**:88–101.
90. Ogut H, Raghunathan S, Menon NM. *Information Security Risk Management Through Self-Protection and Insurance*. Third version, dated October 2005. Richardson, TX: The University of Texas at Dallas, 2005. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.705.5910&rep=rep1&type=pdf.

91. Mersinas K, Hartig B, Martin KM. *et al.* Are information security professionals expected value maximizers?: an experiment and survey-based test. *J Cybersecur* 2016;**2**:57–70.
92. Franke U, Buschle M. Experimental evidence on decision-making in availability service level agreements. *IEEE T Netw Serv Man* 2015;**13**:58–70.
93. Blazenko G. The economics of reinsurance. *J Risk Insur* 1986;**53**:258–77.
94. Varian HR. *Microeconomic Analysis*. New York, NY: W. W. Norton & Company, 1992.
95. Gupta RC. Role of equilibrium distribution in reliability studies. *Probab Eng Inform Sci* 2007;**21**:315–34.
96. Franke U. Optimal IT service availability: shorter outages, or fewer?. *IEEE T Netw Serv Man* 2011;**9**:22–33.
97. Goodin D. Hackers spent 2+ years looting secrets of chipmaker NXP before being detected. Ars Technica. 2023. https://arstechnica.com/security/2023/11/hackers-spent-2-years-looting-secrets-of-chipmaker-nxp-before-being-detected/. (10 November 2023, date last accessed).
98. Franke U, Turell J, Johansson I. The cost of incidents in essential services—data from Swedish NIS reporting. In: *Proceedings of the International Conference on Critical Information Infrastructures Security*. New York, NY: Springer, 2021, 116–29.
99. Gordon LA, Loeb MP, Lucyshyn W. *et al.* Externalities and the magnitude of cyber security underinvestment by private sector firms: a modification of the Gordon–Loeb model. *J Inf Secur* 2014;**6**:24.
100. Laube S, Böhme R. The economics of mandatory security breach reporting to authorities. *J Cybersecur* 2016;**2**:29–41.
101. Baldwin A, Beres Y, Duggan GB,. *et al.* Economic methods and decision making by security professionals. In: *Economics of Information Security and Privacy III*. New York, NY: Springer, 2013, 213–38.
102. Beres Y, Griffin J, Shiu S. *et al.* Analysing the performance of security solutions to reduce vulnerability exposure window. In: *Proceedings of the 2008 Annual Computer Security Applications Conference (ACSAC)*. Piscataway, NJ: IEEE, 2008, 33–42.
103. McCain K, Poslon T. *Best Explanations: New Essays on Inference to the Best Explanation*. Oxford: Oxford University Press, 2017.
104. Säkerhetspolisen. *Säkerhetskyddsanalys – Vägledning i säkerhetsskydd*. version from January 2023. Solna, 2023. https://www.sakerhetspolisen.se/download/18.3baf70bf187108c7cf04b7/1681802201089/Sa%CC%88kerhetskyddsanalys_anpassad.pdf. (10 November 2023, date last accessed).
105. Varga S, Brynielsson J, Franke U. Cyber-threat perception and risk management in the Swedish financial sector. *Comput Secur* 2021;**105**:102239.