



Lancaster University  
Management School

---

## Research on the Economics of Cybersecurity

February 19, 2026

**Kim Kaivanto**

TRIPLE-ACCREDITED, WORLD-RANKED



## Today's talk

CISO question: 'How much should we spend on information security?'

- Gordon & Loeb (2002) on optimal expenditure on information security.
- But: complex externality patterns, trade-offs between stakeholder groups

Psychology of humans facing cyberattacks

- Prospect Theory vs. EV in an SDT model of security behaviour (Kaivanto 2014)
- move beyond single-criterion decision-making models to multi-criteria models of security behaviour (Embrey Kaivanto 2023)

Cyberattacks as potential sources of systemic risk

- Warren Kaivanto Prince (2018)

TRIPLE-ACCREDITED, WORLD-RANKED



Lancaster University  
Management School

## **Trade-off involved choice of how much to spend on information security**

Gordon & Loeb's (2002) "Economics of information security investment" formalisation:

Define  $z$  is the amount spent on information security ( $z \geq 0$ )

$L$  captures the potential loss (in £) from breach of the information set

$v$  is the vulnerability;  $v=0$  completely invulnerable;  $v=1$  completely vulnerable,  
probability of breach under current conditions, prior to investment  $z$

$vL$  is the expected loss in the absence of any infosec investment

$S(v,z)$  is the security-breach probability function

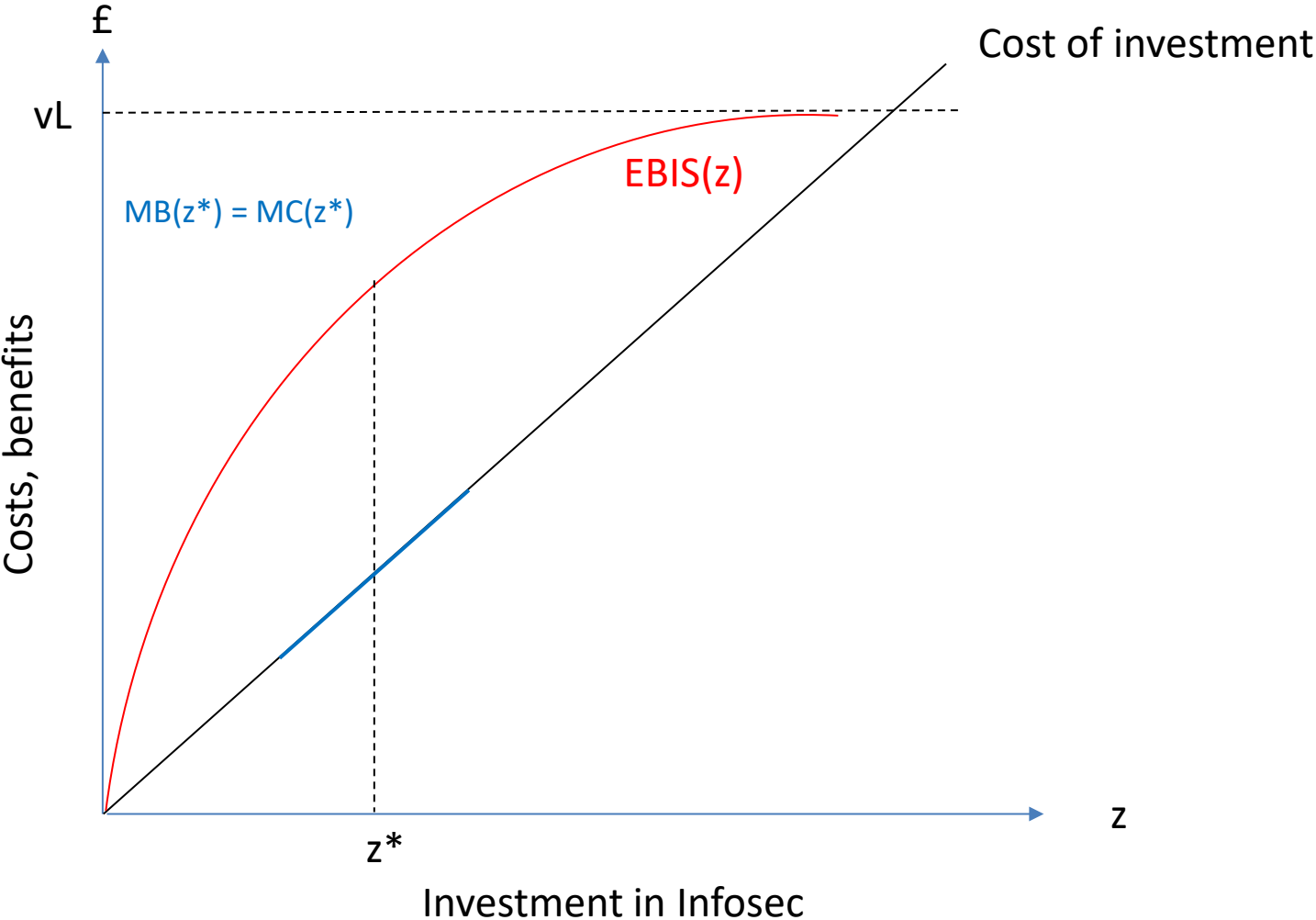
$EBIS(z) = [v - S(v,z)] * L$  is the expected benefit of infosec investment  $z$

TRIPLE-ACCREDITED, WORLD-RANKED



Lancaster University  
Management School

# Trade-off involved choice of how much to spend on information security



TRIPLE-ACCREDITED, WORLD-RANKED



Lancaster University  
Management School

## Trade-off involved choice of how much to spend on information security

Suppose  $S(v,z) = \frac{v}{(\alpha z + 1)^\beta}$  for  $\alpha \geq 0, \beta \geq 1$

or  $S(v,z) = v^{\alpha z + 1}$  for  $\alpha > 0$

then  $z^*(v) < vL/e$  where Euler's number is 2.718... and therefore

$z^*(v) < 36.79\%$  of expected loss in absence of infosec investment.

- this has been extended to a large class of security-breach probability functions  $S(v,z)$
- depends on the form and shape of  $S(v,z)$ !
- also need to determine the rest of the parameters!

TRIPLE-ACCREDITED, WORLD-RANKED



Lancaster University  
Management School

## Trade-off involved choice of how much to spend on information security

Extended to account for externalities:  $L^{SC} = L^P + L^E$

$$z^{SC*}(v) < (1/e)[1 + (L^E / L^P)]vL^P$$

This relies on ‘ensemble averaging’.

Different results would be obtained with ‘time averaging’ in ‘ergodicity economics’.

TRIPLE-ACCREDITED, WORLD-RANKED



Lancaster University  
Management School

## **Trade-off involved choice of how much to spend on information security**

Gordon et al. (2016): when firm has *multiple information sets*

1. Estimate L for each information set
2. Estimate probability that each information set will be breached
3. Create a grid of all possible combinations of 1. and 2. steps
4. Derive the level of cybersec investment by allocating funds to protect the information sets.

Ensure that:

incremental benefits from additional investments  $\geq$  incremental costs of investment

## Complex externality trade-offs

If firm A under-invests in infosec protection, it creates a soft entry point for APT within its supply chain

- a *negative externality* to supply chain as whole (but vulnerable magnet)
- a *positive externality* on firms B, C, ... in its sector (draws attention away from them)

If firm A over-invests in infosec protection above others, firms B, C, ... in its sector remain less-well protected

- a *positive externality* to supply chain as a whole (others now vulnerable magnets)
- a *negative externality* from firm A onto B, C, ... in its sector (now they get targeted)

If industry standards at level  $X$  are introduced, then firms A, B, C, ... are breached at  $X+\epsilon$

- Standardisation can lead to *systemic risk*

## Complex externality trade-offs

There are subtle differences between

- what is good for the country
- what is good for the sector
- what is good for a particular supply chain
- what is good for the firm
- what is good for firm owners & executives vs. employees

There is a role for standards and best-practice guidelines.

But there also needs to be discretion and judgement.

# Kaivanto (2014) Effect of Decentralized Behavioral Decision Making on System-Level Risk

*Risk Analysis, Vol. 34, No. 12, 2014*

DOI: 10.1111/risa.12219

## **The Effect of Decentralized Behavioral Decision Making on System-Level Risk**

**Kim Kaivanto\***

---

TRIPLE-ACCREDITED, WORLD-RANKED



Lancaster University  
Management School

## Kaivanto (2014) Effect of Decentralized Behavioral Decision Making on System-Level Risk

In computer networks, system-level risk depends on the actions and choices of a collection of 'lay' users.

Q. How should we model the decision making of these lay users ?

A. PT-SDT with psychology of deception effects.

Q. Does it matter whether our modeling assumptions reflect normative rationality or heuristics & biases?

A. YES. (See comparative statics and simulation results)

# Kaivanto (2014) Effect of Decentralized Behavioral Decision Making on System-Level Risk

Precis of work

Core of model: classical SDT under normative rationality

Add Behavioral factors

- From the decision-making literature: CPT
- From the phishing & deception literatures

Re-derivation of optimal cutoff threshold under CPT-SDT

- Using T&K92 probability weighting function
- Using neo-additive probability weighting function
- Incorporating the psychology of deception

TRIPLE-ACCREDITED, WORLD-RANKED



Lancaster University  
Management School

# Kaivanto (2014) Effect of Decentralized Behavioral Decision Making on System-Level Risk

## Classical SDT

The optimal cutoff threshold is a function of

- a misclassification cost matrix
- the base rate odds of (email) being non-malicious
- risk preferences

n.b. There is no reason for the misclassification costs to be the same for the user as for the organization

n.b. Classical SDT admits that costs can be replaced by their utilities, but in fact proceeds (exclusively) with minimizing expected cost, i.e. assuming *risk neutrality*.

n.b. In the literature, the ‘optimal classifier’ is computed under risk neutrality (!)

# Kaivanto (2014) Effect of Decentralized Behavioral Decision Making on System-Level Risk

## Behavioral factors

Descriptively, behavioral decision makers display

- reference dependence, framing effects
- non-linear probability weighting
- loss aversion
- ambiguity aversion
- four-fold pattern of risk aversion

All incorporated in (cumulative) Prospect Theory (PT)

Deception plays employ

- peripheral-route persuasion
  - authority, scarcity, similarity & identification, reciprocity, consistency, social proof
- visceral emotions
- urgency
- contextual cues

TRIPLE-ACCREDITED, WORLD-RANKED



Lancaster University  
Management School

# Kaivanto (2014) Effect of Decentralized Behavioral Decision Making on System-Level Risk

## PT-SDT comparative statics

PT-SDT is more *conservative* than classical SDT!

The difference between classical SDT and PT-SDT optimal trade-offs entails that the bias of incorrectly assuming normative rationality is larger for agents with a lower  $d'$ , i.e. a lower ROC curvature and AUC.

PT-SDT shifts the optimal cutoff and the optimal operating point more for agents with a lower ROC curvature and AUC.

The psychology of deception magnifies the effect of behavioral decision making under risk and uncertainty.

TRIPLE-ACCREDITED, WORLD-RANKED



Lancaster University  
Management School

# Kaivanto (2014) Effect of Decentralized Behavioral Decision Making on System-Level Risk

## Comparative simulation results

*Are the individual-level behavioral effects quantitatively consequential at the level of the whole network?*

- M0 Classical SDT model
- M1 PT-SDT model
- M2 PT-SDT with psych of deception,

We simulate (ABM, NetLogo) a 3-week spear-phishing attack on an organization with 100 users.

Each user receives 250 emails per working week.  $1/250=0.004$  of emails are malicious.

During an attack, a user may be fooled at most once. The users *learn* from their mistakes.

TRIPLE-ACCREDITED, WORLD-RANKED

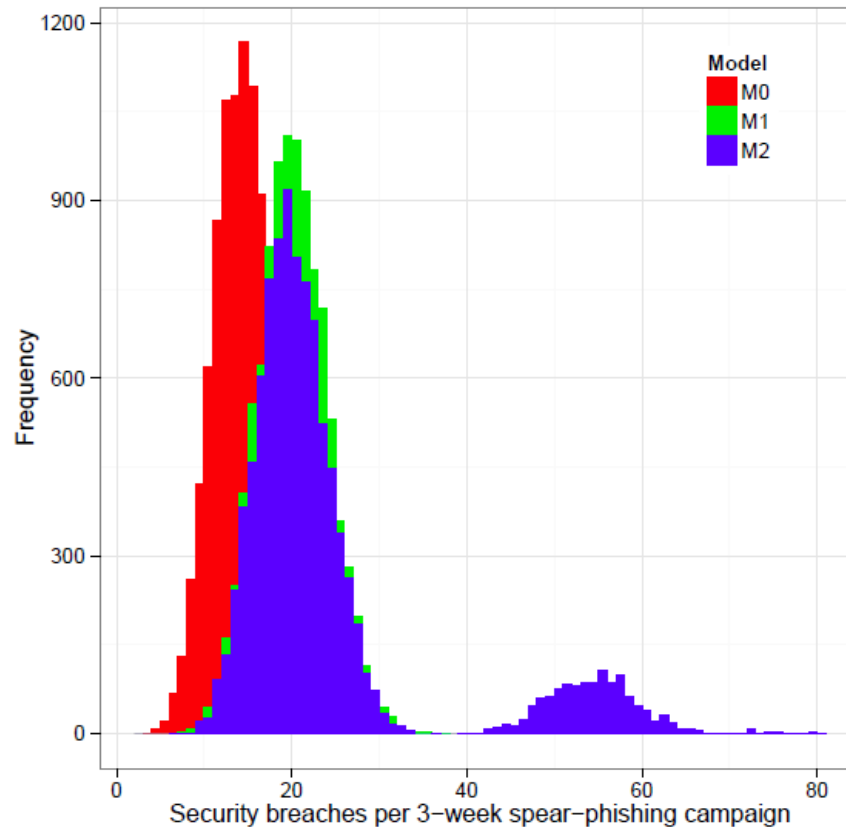


Lancaster University  
Management School

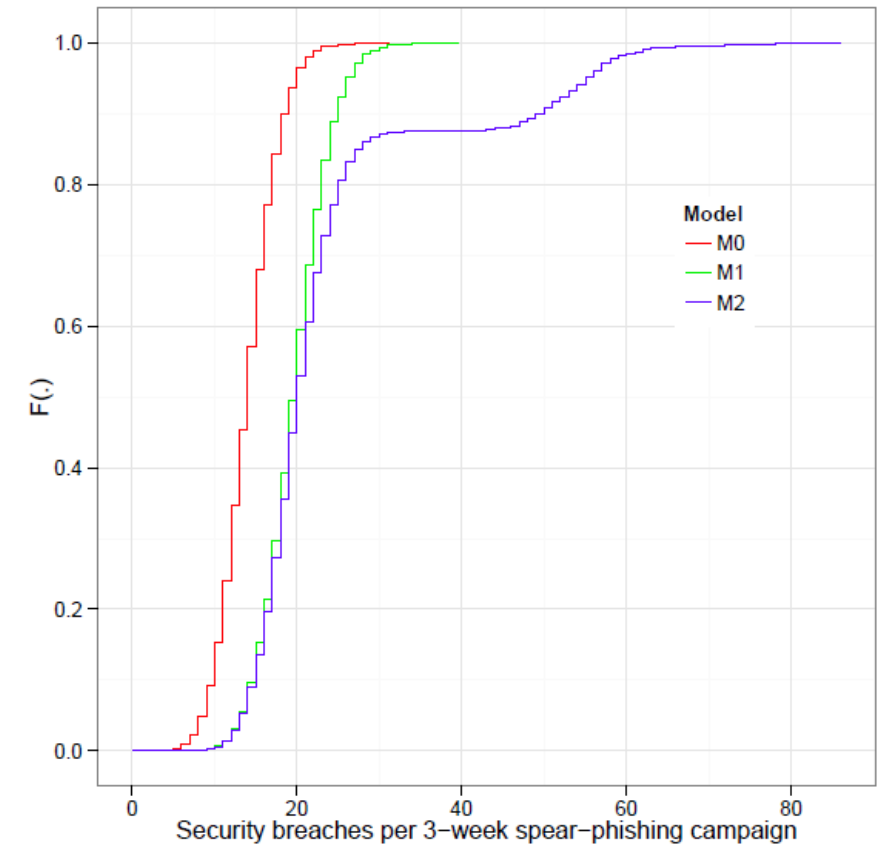
# Kaivanto (2014) Effect of Decentralized Behavioral Decision Making on System-Level Risk

Distribution of security breaches in 10,000 repetitions

(a) Frequency distributions



(b) Empirical CDFs



TRIPLE-ACCREDITED, WORLD-RANKED

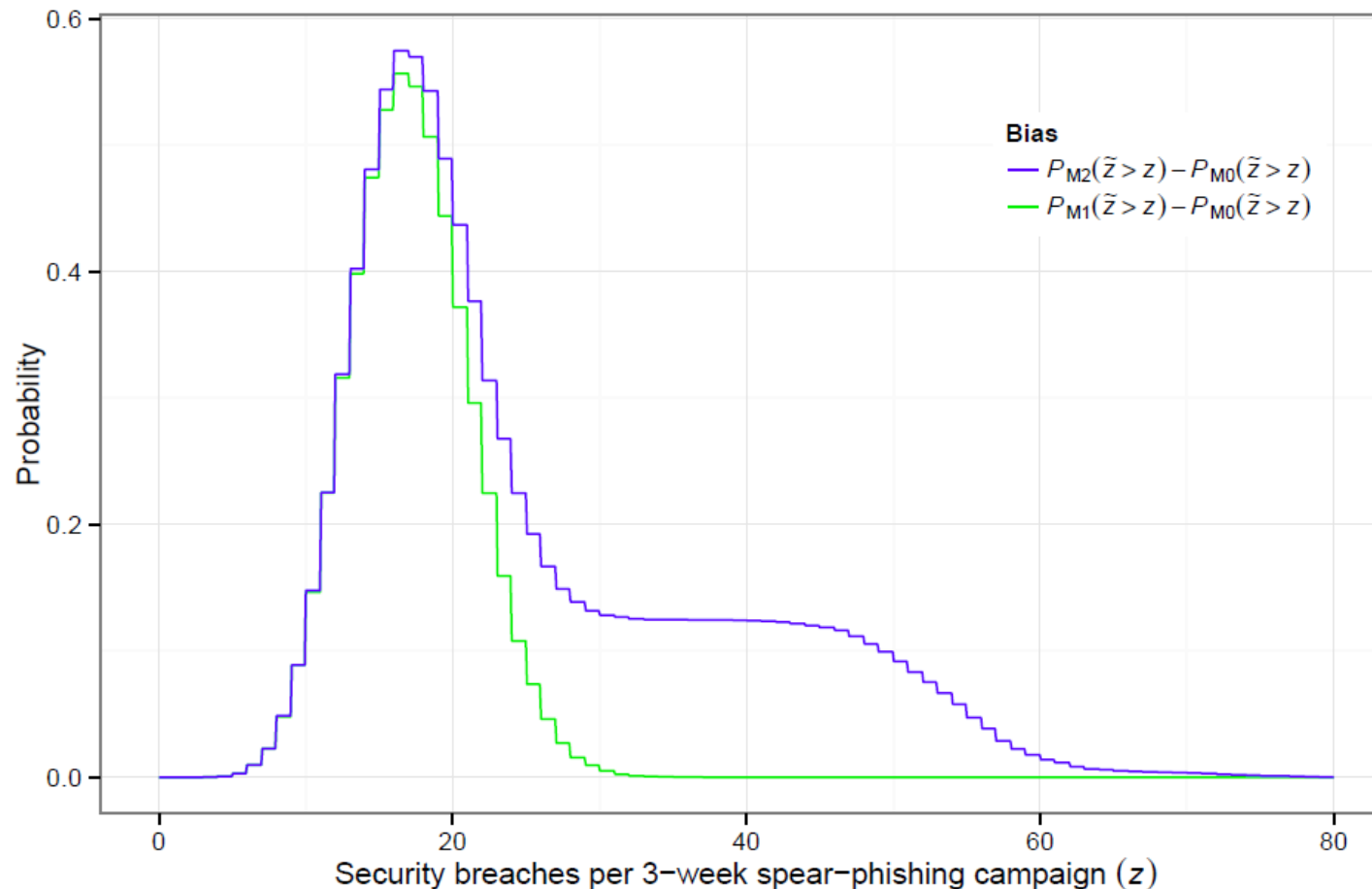


Lancaster University  
Management School

# Kaivanto (2014) Effect of Decentralized Behavioral Decision Making on System-Level Risk

Conclusion: Individual-level behavioral effects matter for system-level risk!

Figure 4: Magnitude of under-estimate (bias) in calculating  $P_{M_0}(\tilde{z} > z)$  when in fact the descriptively accurate model is M1 (green line) or M2 (blue line).



# Embrey & Kaivanto (2023) Many phish in the C: A coexisting-choice-criteria model of security behavior, *Risk Analysis*

ORIGINAL ARTICLE

## Many phish in the C: A coexisting-choice-criteria model of security behavior

Iain Embrey<sup>1</sup>  | Kim Kaivanto<sup>2</sup> 

TRIPLE-ACCREDITED, WORLD-RANKED



Lancaster University  
Management School

## Embrey & Kaivanto (2023) Many phish in the C: A coexisting-choice-criteria model of security behavior, *Risk Analysis*

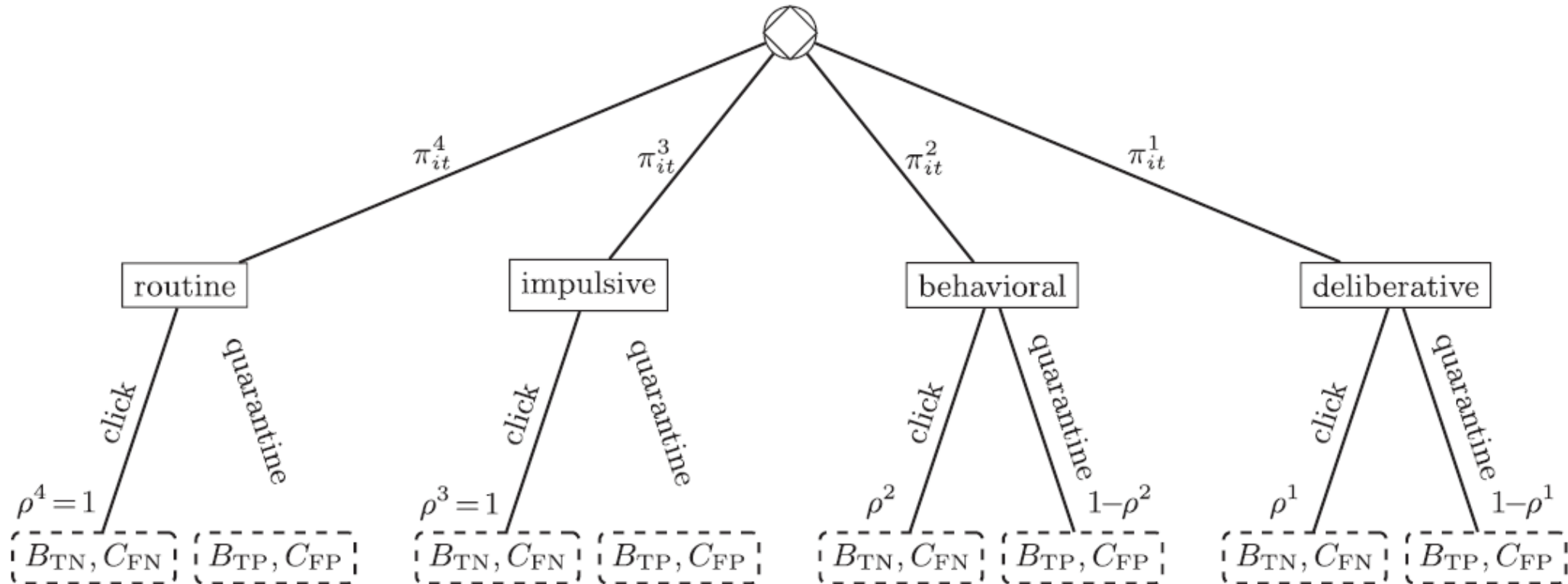
Social-engineering attack vectors operate through emotions, peripheral-route persuasion

Rather than a single choice-criterion (e.g. EU, SEU, CPT), consider the set  $C$  of choice criteria  
- e.g. dual-process theory, *Thinking Fast vs. Thinking Slow*

Let  $|C| \geq 2$

A decision (e.g. click/don't click on a link) is taken via one of multiple possible choice criteria

# Embrey & Kaivanto (2023) Many phish in the C: A coexisting-choice-criteria model of security behavior, *Risk Analysis*



**FIGURE 1** An agent's stochastic state-of-mind response to an email

*Note:* Ex ante the agent is uncertain about an email's true nature. The payoff at each terminal node is therefore either a benefit due to correct classification (True Positive or True Negative), or a cost due to incorrect classification (FP or FN)

TRIPLE-ACCREDITED, WORLD-RANKED

# Embrey & Kaivanto (2023) Many phish in the C: A coexisting-choice-criteria model of security behavior, *Risk Analysis*

## Email recipients' coexisting choice criteria

- c = 1 Normative deliberation: characterized by the internal-consistency axioms of completeness, transitivity, independence of irrelevant alternatives (iia), continuity, Bayesian updating, and time consistency (i.e. exponential discounting).
- c = 2 Behavioral: characterized by the weakening of iia, Bayesian updating, and time consistency (i.e. to hyperbolic discounting), as per the behavioral decision making literature.
- c = 3 Impulsively click through: characterized by dominance of visceral emotions, which suppress and displace deliberative reasoning; the remaining consistency axioms are abandoned.
- c = 4 Routinely click straight through: characterized by routinization and automaticity; again, the remaining consistency axioms are abandoned.

TRIPLE-ACCREDITED, WORLD-RANKED



Lancaster University  
Management School

## Embrey & Kaivanto (2023) Many phish in the C: A coexisting-choice-criteria model of security behavior, *Risk Analysis*

Highly routinised, efficient practices increase vulnerability. It is vital that organizational culture supports the precautionary verification steps

-ISOs should actively engage with wider aspects of organizational culture & practices

A strategic attacker will seek to target  $c=3, c=4$ .

Training that focuses on  $c=1$  will be of limited effectiveness.

- Training should focus on reducing criterion-selection probabilities  $\pi^3, \pi^4$

Best be achieved by helping employees to understand:

- (i) their inherent vulnerability to phishing when making choices either Routinely, Impulsively;
- (ii) the psychological ploys by which attackers may induce Impulsive or Routine SoM.

TRIPLE-ACCREDITED, WORLD-RANKED



Lancaster University  
Management School

# Embrey & Kaivanto (2023) Many phish in the C: A coexisting-choice-criteria model of security behavior, *Risk Analysis*

Experiment at Lancaster University

Baseline Treatment (TB): University's standard email security/anti-phishing training.

Routine-Interrupt Treatment (TRI): 7-min multimedia interactive training module with periodic understanding verification questions. Designed to cause the student to slow down, think consciously, and pre-empt "routine" and "automatic" processing of email information.

Impulse-Interrupt Treatment (TII): 8-min multimedia interactive training module with periodic understanding verification questions. Designed to (a) alert the student to those features of an email that aim to elicit visceral emotions and thereby trigger an impulsive response, and to (b) provide strategies for processing emails to reduce the likelihood of impulsive clicks.

TRIPLE-ACCREDITED, WORLD-RANKED



Lancaster University  
Management School

## Embrey & Kaivanto (2023) Many phish in the C: A coexisting-choice-criteria model of security behavior, *Risk Analysis* *Procedures.*

The training package and follow-up test were administered to participants via the university's Virtual Learning Environment (Moodle).

The follow-up test consisted of a landing & introduction page, a consent page, six test-email pages (see Table 3), and a final outro page.

Each test-email page contains not only the html-formatted test email complete with embedded links, but also a further section asking the participant to rate “How likely would you be to fall for an attack like this in real life?” with radio buttons for 95%, 75%, 50%, 25%, and 5%.

---

TRIPLE-ACCREDITED, WORLD-RANKED



Lancaster University  
Management School

## Embrey & Kaivanto (2023) Many phish in the C: A coexisting-choice-criteria model of security behavior, *Risk Analysis*

Description	Weakness targeted	Authenticity
justWink e-card	impulsive response	genuine
facebook notification	routine response	genuine
library renewal reminder	routine response	fake
student union discount card offer	impulsive response	fake
law enforcement fraud alert	impulsive response	fake
anti-phishing project collaboration invitation	routine response	fake

TRIPLE-ACCREDITED, WORLD-RANKED



Lancaster University  
Management School

## Embrey & Kaivanto (2023) Many phish in the C: A coexisting-choice-criteria model of security behavior, *Risk Analysis*

Table 4: Linear probability model parameters and SEs.

participants assigned to TRI  
correctly classified nearly 10%  
more emails than those assigned  
to TB

TRI participants identified genuine  
emails at a much higher rate (16%)  
than TB

	Outcome Measures			
	<i>score</i>	<i>true pos</i>	<i>true neg</i>	<i>confidence</i>
<i>TB</i>				
<i>TRI</i>	0.0965** (0.0323)	0.0650 (0.0411)	0.1593*** (0.0491)	-0.0724* (0.0352)
<i>TII</i>	0.0336 (0.0312)	0.0523 (0.0405)	-0.0037 (0.0493)	0.0371 (0.0318)
<i>constant</i>	0.5278*** (0.0226)	0.5486*** (0.0299)	0.4861*** (0.0345)	0.6433*** (0.0239)
<i>n</i>	332	332	332	328
<i>R</i> <sup>2</sup>	0.0273	0.0087	0.0413	0.0334

One, two, and three asterisks denote statistical significance at the 5%, 1%, and 0.1% levels respectively.

## Embrey & Kaivanto (2023) Many phish in the C: A coexisting-choice-criteria model of security behavior, *Risk Analysis*

*TRI* appears to be successful in stimulating metacognitive processes, and relatedly, in reducing (over)confidence.

Working practices in most commercial, voluntary, and public-sector organizations presume that links and email attachments are benign when sent from within the organization or by customers, suppliers, or partner organizations.

This is a major vulnerability that is as much a reflection of organizational culture as it is a reflection of explicit security protocols (or absence thereof).

# Warren, Kaivanto, Prince (2018) Could a cyber attack cause a FS systemic impact? *BoE QB*



BANK OF ENGLAND

## Quarterly Bulletin

2018 Q4

Topical article

Could a cyber attack cause a systemic impact in the financial sector?

TRIPLE-ACCREDITED, WORLD-RANKED



Lancaster University  
Management School

# Warren, Kaivanto, Prince (2018) Could a cyber attack cause a FS systemic impact? *BoE QB*

a systemic impact is triggered via a shock<sup>(9)</sup> (eg a firm failure);

its causes can gradually build up<sup>(10)</sup> (eg via a credit bubble or the neglect of tail risk);<sup>(11)</sup>

a significant part or parts<sup>(12)</sup> of the sector are impacted;

the event propagates through and is amplified by the interconnected<sup>(13)</sup> nature of the affected business environment;

there is a lack of substitutability<sup>(14)</sup> to contain the disturbance;

human behaviour fuels the impact as consumers react to changes in confidence and trust in the financial sector (eg hoarding or flight);<sup>(15)</sup>

the consequence is a failure of the provision of services<sup>(16)</sup> (eg access to credit); and the impact is felt in the real economy<sup>(17)</sup> (eg economic growth or welfare).

Common  
features  
of  
systemic  
risk

From the BoE's standpoint, systemic risks have to translate into a significant reduction to economic activity (i.e. GDP impact, recession).

There is persistent disagreement whether cyber attacks can be the *cause* of systemic crisis rather than merely the *trigger*.

TRIPLE-ACCREDITED, WORLD-RANKED



Lancaster University  
Management School

## Warren, Kaivanto, Prince (2018) Could a cyber attack cause a FS systemic impact? BoE QB

There are no current examples of systemic cyber risk crystallising and impacting the real economy but this does not prove an absence of risk.

- Deloitte's Independent Review of RTGS Outage on 20 October 2014
  - RTGS outage for 9h; resulted from a routine CHAPS config change
  - Payments delay impact (see next slide)

- Financial Impact:

As of 20 March 2015, 36 individuals had contacted the Bank directly enquiring about redress and were directed to their banks for further information. The Bank had paid 9 claims totalling £4,056.89 and was not expecting any substantial further claims at the time of publication.

TRIPLE-ACCREDITED, WORLD-RANKED



Lancaster University  
Management School

# Warren, Kaivanto, Prince (2018) Could a cyber attack cause a FS systemic impact? *BoE QB*

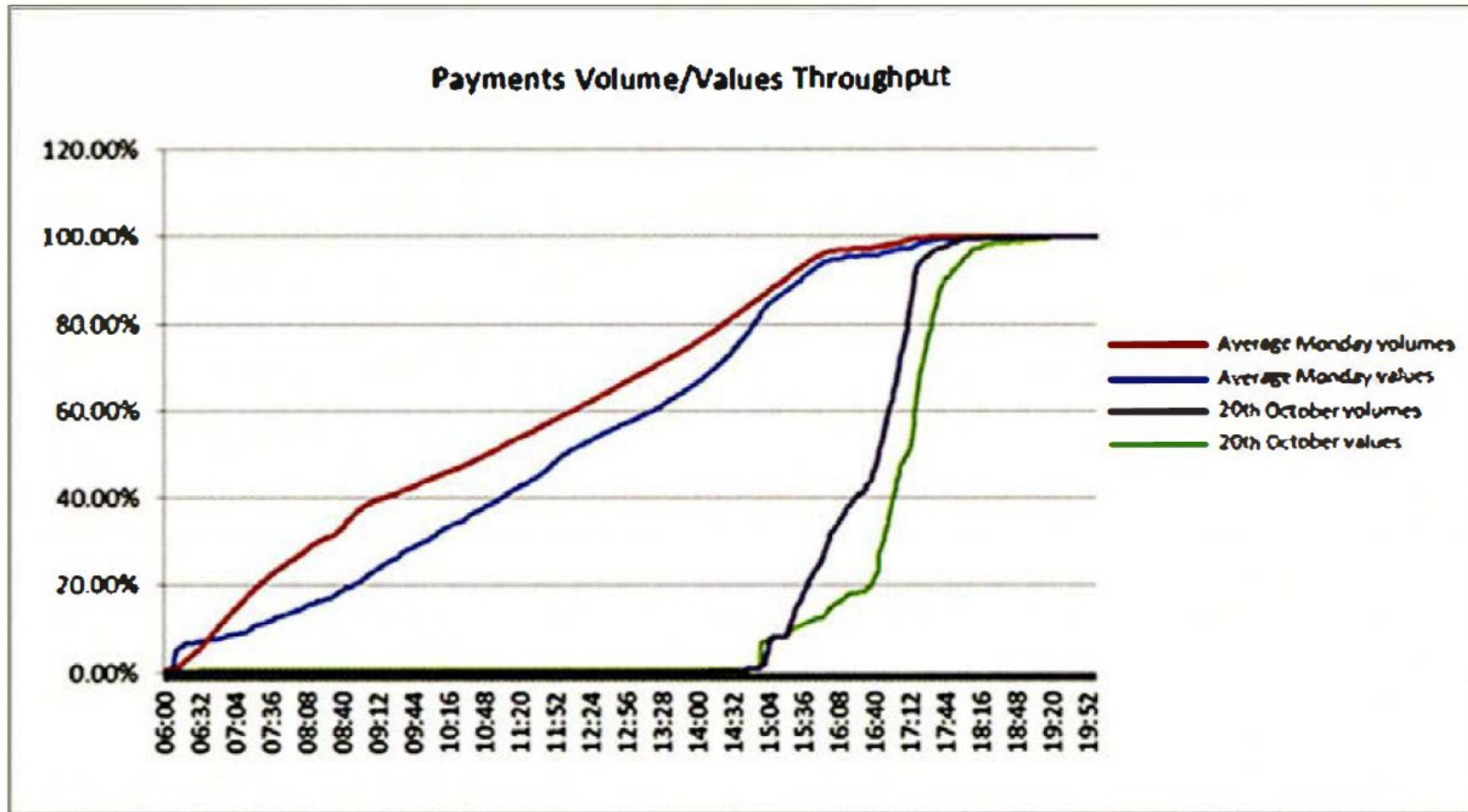


Figure1: Payments Volume/Values Throughput (including October 20 incident)

TRIPLE-ACCREDITED, WORLD-RANKED



Lancaster University  
Management School

## Warren, Kaivanto, Prince (2018) Could a cyber attack cause a FS systemic impact? *BoE QB*

There are no current examples of systemic cyber risk crystallising and impacting the real economy but this does not prove an absence of risk.

- Deloitte's Independent Review of RTGS Outage on 20 October 2014
- 2016 Bangladesh Bank cyber heist
  - US\$1 billion from the Bangladesh FRBNY account
  - US\$81 million was successfully transferred to accounts in the Philippines.
  - Philippine law exempted casino transactions from the scrutiny of the country's Anti-Money Laundering Council (AMLC) → loophole subsequently closed
  - No noticeable recession

TRIPLE-ACCREDITED, WORLD-RANKED



Lancaster University  
Management School

## Warren, Kaivanto, Prince (2018) Could a cyber attack cause a FS systemic impact? *BoE QB*

There are no current examples of systemic cyber risk crystallising and impacting the real economy but this does not prove an absence of risk.

- Deloitte's Independent Review of RTGS Outage on 20 October 2014
- 2016 Bangladesh Bank cyber heist
- BUT: Bulgarian bank runs: Corporate Commercial Bank (KTB), First Investment Bank (FiB)
  - Corporate feud, continued through online campaigns (e.g. facebook)
  - Led to genuine bank runs
  - Bulgaria asked EU Commission for exception the EU State Aid rules to provide backstop
  - A genuine 'systemic event'
  - Crisis led to long-term closure of KTB, a significant budgetary hit, and a loss of public confidence

TRIPLE-ACCREDITED, WORLD-RANKED



Lancaster University  
Management School

# Warren, Kaivanto, Prince (2018) Could a cyber attack cause a FS systemic impact? *BoE QB*



CIA Triad:



Most thinking around cyberattacks views them through the lens of ‘availability’ or ‘confidentiality’.

The Bulgarian Bank Run demonstrated the effect of a PERCEPTION OF INTEGRITY ATTACK.

TRIPLE-ACCREDITED, WORLD-RANKED



Lancaster University  
Management School