Ransomware Insurance*

Zachary Breig† Claudio Mezzetti‡ Zarina Vakhitova§

December 29, 2025

Abstract

Ransomware insurance can improve recovery (restoring systems, data, business operations) by making ransom payments feasible, but it may raise ransom demands and encourage attacks. We use mechanism design to characterize the welfare-maximizing structure of ransomware insurance under asymmetric information about victim's budgets. The optimal contract is a single-cap (sub-limit) design that bunches victims at a common post-attack payment capacity, thereby inducing lower equilibrium ransom demands. Under the optimal design, victim and total surplus increase without increasing attacker payoffs, whereas poorly designed coverage can reduce welfare. A laboratory experiment comparing no insurance to sophisticated and naive single-cap contracts supports these predictions.

Keywords: insurance, ransomware, kidnapping, experiment, mechanism design.

JEL-Classification: C90; D47; D82; D86; G22; K24.

*The paper has benefited from comments and feedback from seminar audiences at the University of Sydney, Monash, and Deakin. We thank Max Grossmann for sharing excellent code for generating sliders in oTree. The experiment reported in this study received ethical approval from UQ's BEL LNR Committee (2023/HE000101) and was preregistered in the AEA RCT registry (AEARCTR-0013035).

†School of Economics, University of Queensland, Level 6, Colin Clark Building 39, Brisbane St. Lucia, QLD 4072, Australia. Email: z.breig@uq.edu.au

‡School of Economics, University of Queensland, Level 6, Colin Clark Building 39, Brisbane St. Lucia, QLD 4072, Australia. Email: c.mezzetti@uq.edu.au.

§School of Social Sciences, Monash University, 20 Chancellors Walk, VIC 3800, Australia. Email: zarina.vakhitova@monash.edu

1 Introduction

Ransomware is a form of malware that infects computers using security flaws to gather, encrypt, or erase original data files. Ransomware attacks can cause temporary or permanent loss of sensitive or proprietary information, disruption to regular operations, financial losses to restore systems and files, third-party liability losses, and damage to an organization's reputation. The primary objective of a ransomware attack is extortion, i.e., extracting payments (Kharraz et al., 2015; Paquet-Clouston et al., 2019); attackers demand a ransom payment from the victims – often in cryptocurrency – in exchange for restoring access or preventing data disclosure (e.g., in the dark web). In recent years, ransomware attacks have emerged as one of the most significant cyber-threats facing organizations, particularly in the healthcare, financial services, higher education, telecommunication, and energy sectors. According to the US Office of the Director of National Intelligence (2025), there were approximately 4,591 ransomware incidents in 2023 and 5,289 in 2024. Victims paid a record-breaking US$1 billion in ransoms in 2023 and approximately US$813 million in 2024 (Chainalysis Team, 2024, 2025). A recent global survey estimated that 49% of affected organizations reported paying ransoms in 2024, with a median payment of around US$400,000 (Sophos, 2025).

Against this backdrop, cybersecurity insurance has emerged as an important risk management tool to mitigate the damage from ransomware attacks. Cyber insurance policies that cover first-party costs such as incident response, ransom payments, and business interruption, as well as third-party liabilities, have grown in popularity. The global volume of cyber-insurance premiums has reached approximately USD 16.6 billion in 2024, driven by increased extortion-related claims (Risk & Insurance Editorial Team, 2025). As ransomware losses have escalated, insurers have increased premiums and tightened underwriting standards, often requiring compliance with baseline cybersecurity controls, such as offline backups, or multifactor authentication, as a precondition for coverage (Geneva Association, 2022; CRO Forum, 2023). Importantly, they have also introduced sub-limits for ransom payments. Sub-limits are caps within a cyber-insurance policy that limit how much the insurer will reimburse specifically for ransom-related costs, even if the overall policy limit is much higher.

The interplay between ransomware and insurance raises fundamental questions about incentives for attackers and governance. On one hand, by facilitating the payment of the ransom, insurance may enable organizations to recover data that might otherwise be lost or leaked in the dark web. It may lead to a faster resumption of critical services (e.g., in hospitals, utilities, logistics companies, etc.), and a reduction in harm to the public. On the other hand, critics warn that it might increase ransom payments and the profitability of ransomware, thus inducing attackers to enter the ransomware market and increase attack volume. A central question in the policy debate that we study in this paper is whether the availability of insurance benefits victims and society or whether it should be banned — an issue that speaks directly to moral hazard and the social externalities of insurance.[1] We aim to shed light on several policy questions: (i) Does ransomware insurance raise the ransom requests of attackers? (ii) Does it raise the likelihood that victims will pay the ransom? (iii) Does it facilitate the recovery of data? (iv) Does it induce entry and further attacks by attackers? (v) Does it induce cybersecurity complacency?

We use a mechanism design approach to study the ransomware insurance market. We impose several natural constraints on the "insurance designer" – including the constraint that the insurance companies compete against one another – and derive the insurance policy menu that maximizes the potential victim's expected payoff. A mechanism recommends insurance coverage levels to potential victims; given victims' take-up and payment decisions, the attacker chooses a profit-maximizing ransom. The recommended coverage must be incentive compatible for victims.

The novel, key feature of our modeling approach is to treat agents' budget for paying ransoms as private information, unknown to both the attacker and the insurance designer. When a ransomware attack hits, a victim without insurance may not have enough funds to pay the requested ransom. Insurance increases the funds available, but attackers are aware of the existence of an insurance market and modify their ransom requests.

We find that optimal insurance contracts involve the use of the caps and sub-limits that we see in practice. In particular, they take the form of "single-cap insurance", which bunches

[1] See Block and Tinsley (2008), for a libertarian, legal argument against banning ransom payments: "such a law ends up punishing the victim instead of the kidnapper."

a mass of potential victims at a given protection level (the budget available to pay the ransom after an attack), and by doing so induces the attacker to select a ransom that can be paid by the mass of insured agents. Victims benefit both from recovering their data more often *and* by paying lower ransoms. Thus, the sub-limits insurance companies are increasingly using do not just limit insurers' losses; they benefit the insured! We show that in our model, attackers cannot lose from the presence of insurance, but they do not gain if the policy is chosen optimally. Under the optimal (single-cap) menu, total surplus increases; under poorly designed menus, insurance can raise ransom demands and reduce surplus.

To demonstrate the applicability of our results, we implement a laboratory experiment that tests how single-cap insurance contracts shape outcomes in ransomware markets. Subjects act as both attackers (choosing ransom demands) and victims (purchasing insurance and deciding whether to pay). We compare a no-insurance benchmark to two single-cap contracts: "sophisticated" insurance, designed to increase victims' expected payoffs, and "naive" insurance, which is predicted to reduce them. The results largely match the theoretical predictions: sophisticated insurance concentrates ransom demands at the payment cap and increases victims' average payoffs relative to no insurance, whereas naive insurance reduces victims' payoffs.

Finally, we examine how the optimal menu changes under alternative market features. In a benchmark in which agents can purchase unlimited actuarially fair coverage, insurance benefits victims only when the value of data recovery is sufficiently high. We then study how the optimal menu varies with uncertainty about the attacker's darknet valuation and with victims' ability to invest in protection. We also show that the optimality of single-cap contracts is robust to introducing risk aversion, victim bargaining power, and a social cost of darknet data sales.

Although we focus on ransomware, the mechanism we study is more general. The same logic extends to environments in which a counterparty with market power sets a price for a remedy, agents are liquidity constrained, and the outcome is binary: pay and recover versus do not pay and suffer a loss. In such settings, insurance (or any financing arrangement) relaxes liquidity constraints and can therefore change the counterparty's price offers and the welfare effects of trade. Kidnapping and maritime or land-based hijacking are leading examples; on the rise of kidnappings of people, see Control Risks (2022) and on the rise of maritime piracy, see Jones (2025). The same logic can also be relevant in less literal "extortion" settings, such as disaster repair or specialized medical services, when urgent need and limited local capacity confer pricing power and some agents are screened out by liquidity constraints.

Related Literature There is a large literature on insurance stemming from the classic work on the demand for coverage under risk aversion of Arrow (1963) and Mossin (1968). Starting with Rothschild and Stiglitz (1976), a main strand of this literature has focused on adverse selection and the existence of a competitive equilibrium under accident risk heterogeneity.[2] We differ substantially from this literature by focusing on budget heterogeneity in an extortion setting, and by using a mechanism design approach.

The literature on cyber risk (or kidnapping) and insurance is small. Selten (1988) and Lapan and Sandler (1988b) are the starting point of the literature on kidnapping. In Selten (1988), a hostage is taken, and the kidnapper threatens to kill the victim unless a ransom is paid by the hostage's family. The focus of the paper is on the credibility of the threat to kill the hostage. If, as it is reasonable, the kidnapper does not gain from killing, then the threat is not credible in subgame-perfect equilibrium, and thus kidnapping should not occur. Lapan and Sandler (1988b) study a model of hostages taken by a terrorist organization and focus on when the government gains from committing not to pay the ransom. They show that committing not to pay is only credible in limited circumstances. In Konrad and Skaperdas (1997), a criminal gang must make an unobservable up-front investment in its ability to extort from a sequence of potential victims. They show that as the number of potential victims becomes large, the only subgame-perfect equilibria have extortion and expected loss from violence.

As Selten (1988), Fink and Pingle (2014) model kidnapping as a full-information two-person game between a kidnapper and a family, but their focus is on the impact of kidnapping insurance. They model the insurance industry as a standard competitive market, and families as risk-averse decision makers choosing the level of insurance to purchase. They find that the

[2]See Stiglitz (1977) and Chade and Schlee (2012) for the analysis of a monopolistic market environment.

existence of an insurance market increases the ransom families are willing to pay and might increase the probability of an attack. Unlike this paper, they do not study the impact of different insurance policies and do not allow for asymmetric information about the victims' ability to pay.

A more recent literature applies and extends kidnapping models to cyber-extortion, especially ransomware. Laszka et al. (2017) develop a game-theoretic model, focused on the attackers' decision on whom to attack and the potential victims' security investment decisions for mitigation. Cartwright et al. (2019) adapt the models in Selten (1988) and Lapan and Sandler (1988a) to digital "kidnapping" of encrypted files. They focus on the spillover effects across victims and their effect on the ransom that criminals charge and the role of deterrence through preventive measures.

While we are not aware of quantitative, empirical papers studying cybersecurity insurance, there are several qualitative studies. Baker and Shortland (2023) interviewed industry insiders and found that cybersecurity insurance helps to contain liabilities and restore affected IT systems, but security decisions are typically left to the insured. Using data from surveyed or interviewed professionals with expertise in cyber insurance, cybersecurity and/or ransomware, Cartwright et al. (2023) found that: "perceptions are divided on whether victims with insurance are more (or less) likely to pay a ransom." Finally, we should also mention the sizeable literature on cyber insurance management, risk quantification and pricing; for a recent survey, see Carannante and Mazzoccoli (2025).

Organization The remainder of the paper is structured as follows: Section 2 introduces our theoretical model; Section 3 analyzes the model in the benchmark case in which insurance is not available; Section 4 describes our mechanism design approach, derives the optimal insurance policy, and studies its properties; Section 5 describes and present the results of a lab experiments in which we tested the prediction of the theory developed in Section 4; Section 6 studies several important extensions to our model and derives additional policy results; Section 7 offers some conclusions. Proofs are in Appendix A; additional empirical results are in Appendix B, and the experiment screenshots are in Appendix C.

2 The Model

We study the optimal design of a competitive market for ransomware insurance. There are a continuum of agents with mass one, who are potential victims of a ransomware attack, a continuum with mass $\alpha < 1$ of attackers, and a competitive insurance industry. Each attacker successfully attacks an agent. Thus, we can think that α is the probability that any given agent becomes a victim of a ransomware attack.[3]

An agent attaches value v to the data that could be stolen in a ransomware attack; the agent has liquidity, or a budget, b, that could be used to pay the ransom if an attack takes place. The agents are assumed to be risk neutral, but if they are victims of an attack, they can only recover the data if they pay the ransom. Thus, if the ransom request exceeds their budget but is below the value of the data, then agents would benefit from being insured, because without insurance they would not be able to pay the ransom and recover the data. Naturally, this benefit must be traded off against the cost of insuring.[4]

The data of the agent, if stolen, could be sold in a competitive darknet market; it has a darknet market value of $v_d > 0$. We assume that the darknet value of the data is less than the value of the data to the agent, $v_d < v$. If it were not the case, the data would never be bought back by the agent. Indeed it would be efficient for the attacker to keep the data. The values of the data to an agent and in the darknet market are common knowledge, while the budget b is private information of the agent.[5] When solving the mechanism designer's problem below, we refer to b as the agent's *type*. Let $F(b)$ be the distribution of the agent's budget; it has a density $f(b)$ that is positive everywhere on the support $[\underline{b}, \bar{b}]$, with $\bar{b} \leq v$, and an increasing hazard rate $h(b) = \frac{f(b)}{1-F(b)}$.

The timeline of our model is as follows. In the first stage, the market designer selects a menu of insurance policies; each policy in the menu must yield zero profit to each insurance company, as each company in the competitive insurance market must find it optimal not to deviate and offer a different menu of policies. If an insurance policy selected by the designer

[3]Equivalently, we could think that the mass of attackers is μ and that the probability that an attack is successful is σ. Letting $\alpha = \sigma\mu$, the analysis would be unchanged.

[4]In Section 6.3 we show that our main result extends to the case of risk-averse agents.

[5]For our results, it is sufficient that the insurance companies (not the agent) know the darknet value v_d. We believe this to be a realistic assumption. To check the robustness of our results, in Section 6.2 we extend our analysis to the case where the darknet value is private information of the attacker.

yielded a positive profit to the companies, a company would have an incentive to capture the entire insurance market by offering a policy menu with a lower profit that agents would accept. In the second stage, each agent decides whether to buy insurance and if he buys, which of the policies on offer to buy. In the third stage, each attacker is randomly assigned to an agent and successfully gains access to the agent's data. The attacker then chooses a ransom request r to maximize his expected payoff. In the fourth stage, the victim (i.e., the agent who was attacked) decides whether to pay the ransom, but paying the ransom is only feasible when the victim's insurance policy yields a final budget that is sufficient to cover the ransom payment.

3 No Insurance Benchmark

In this section, we begin our analysis by considering the benchmark case in which insurance is not available (or is forbidden).

If there is no insurance and the attacker asks for a ransom r, then the victim finds it optimal to accept to pay and recover the data if $r \leq b$, as we have assumed that the value of the data is at least as high as the highest possible budget, $v \geq \bar{b}$. On the other hand, the victim cannot pay the ransom if $r > b$. It follows that the attacker's maximization problem is:

$$\max_r (r - v_d)(1 - F(r)) + v_d. \tag{1}$$

Note first that if $v_d \geq \bar{b}$, then the solution is $r_0 = v_d$ or, equivalently, if the darknet value is higher than the maximum budget available to the victim, then the attacker sells the data on the darknet market with probability one.

If, instead, $v_d < \bar{b}$, then the first order condition for an interior solution $r_0 \in (\underline{b}, \bar{b})$ of the maximization problem (1) is:

$$1 - F(r) - (r - v_d)f(r) = 0$$

which yields the solution:

$$r_0 = v_d + \frac{1 - F(r_0)}{f(r_0)}.$$

Since the second order condition is satisfied, we may summarize this result in the following proposition.[6]

Proposition 1. *Without an insurance market, if $v_d < \bar{b}$, then the attacker asks for the ransom r_0 that solves:*

$$r_0 = v_d + \frac{1}{h(r_0)}. \tag{2}$$

If $v_d \geq \bar{b}$, then the attacker asks for the ransom $r_0 = v_d$ and thus sells the data in the darknet market with probability one.

Note that the assumption of an increasing hazard rate implies that, even at an interior solution, the ransom request is increasing in v_d.

Letting $\mathbb{1}_{r\leq b}$ be the indicator function equal to 1 if $r \leq b$ and equal to zero otherwise, and $\mathbb{1}_{r>b}$ the indicator function equal to 1 if $r > b$ and equal to zero otherwise, we may denote the attacker's expected payoff under no insurance as follows:

$$\pi_0 = \int_{\underline{b}}^{\bar{b}} \left[r_0 \mathbb{1}_{r_0\leq b} + v_d \mathbb{1}_{r_0>b}\right] dF(b). \tag{3}$$

4 Optimal Insurance

We take a mechanism design approach and study what menu of insurance policies maximizes the agents' expected payoff in a setting with a competitive insurance industry. Such an approach allows us to see what policy measures would be ideal to implement and thus can provide guidance to policymakers and industry experts. Indeed, we show that capping the amount of insurance agents can buy (but not forbidding insurance) is optimal. In Section 6.1 we compare our results with the results in a setting in which agents can "fully insure"; that is, they can buy as much insurance as they want at a fair premium.

An insurance policy specifies the premium $p \geq 0$ that an agent must pay upfront in exchange for the insurance company committing to pay x to the agent if the agent is attacked. If the victim pays ransom r to the attacker and $x - r > 0$, then this can be viewed as a top-up payment; if $x - r < 0$, then we can view $r - x$ as a copayment. More formally, a

[6] The assumption that the hazard rate $h(r) = \frac{f(r)}{1-F(r)}$ is increasing in r implies that the second order condition is satisfied, since $[1 - F(r)]\,[1 - (r - v_d)h(r)]$ is decreasing in r at $r = r_0$.

menu of insurance policies $\langle p, x\rangle$ specifies two functions, with the function $p : [\underline{b}, \bar{b}] \to \mathbb{R}_+$ mapping types into premium payments and the function $x : [\underline{b}, \bar{b}] \to \mathbb{R}_+$ determining the top-up payout or co-payment of each type.

Note that if there is an attack, an insurance company's outlay when the agent is type b is $x(b)$ regardless of the ransom request. Since an attack occurs with probability α and the insurance market is assumed to be competitive, we say that a menu is feasible if the insurance companies make *zero profit* on each agent; that is, the following must hold for all $b \in [\underline{b}, \bar{b}]$:

$$p(b) - \alpha x(b) = 0. \tag{4}$$

The zero profit condition guarantees that there is no menu of insurance policies that would yield a positive profit to a deviating company that offered it when all other companies follow the designer recommendation and offer the menu of policies $\langle p, x\rangle$.

Our assumption that agents' utility is linear in money implies the key value that insurance provides is to allow for the payment of ransoms.[7] Given the initial budget and insurance policy $(p(b), x(b))$, we define the money available for an agent of type b to pay a ransom in the case of an attack, or final budget, as

$$\beta(b) \equiv b + x(b) - p(b). \tag{5}$$

Without loss of generality we ignore the case in which the final budget is higher than the value of the data, setting $\beta(b) \leq v$.

Putting together (4) and (5), we can define the premium and compensation as functions of the final budget:

$$x(b) = \frac{\beta(b) - b}{1 - \alpha}, \tag{6}$$

$$p(b) = \frac{\alpha[\beta(b) - b]}{1 - \alpha}. \tag{7}$$

Thus, a menu of insurance policies in a competitive market is fully defined by the function β, and it is this function that we focus on choosing to solve the designer's problem. Letting r_* be the ransom request chosen by the attacker, we can write the payoff of an agent of type

[7] We discuss how risk aversion affects our analysis in Section 6.3.

b as:

$$(1-\alpha)v + b + \alpha(v - r_*)\mathbb{1}_{r_* \le \beta(b)}. \tag{8}$$

4.1 Incentive Compatibility

The insurance designer's problem will be to find an insurance policy that maximizes agents' payoff given their incentive constraints as well as the behavior of the attacker, which may change in response to the menu of insurance contracts. Thus, we define a mechanism as $\langle \beta, r_* \rangle$, which recommends that an agent of type b choose the insurance contract $\beta(b)$ and that the attacker makes the ransom request r_*. This section describes the incentive compatibility constraints that we require such a mechanism to satisfy.

The first condition that a menu of insurance policies must satisfy is *affordability*; for each type of agent, the premium cannot be higher than the agent's available initial budget; that is, for all $b \in [\underline{b}, \overline{b}]$ it must be $p(b) \le b$. Using (7), this condition can be written as:

$$\beta(b) \le \frac{b}{\alpha}. \tag{9}$$

The second is a *participation* condition: the agent must be willing to take up insurance rather than not to insure. Since, by the zero profit condition, the expected gross payment $\alpha x(b)$ received by an agent is equal to the premium $p(b)$, the agent expected monetary payoff from insuring is zero. Hence, the agent will benefit from participating only if the final budget available to pay the ransom is at least as high as the initial budget, making it at least as likely to recover the data; that is, it must be that for all $b \in [\underline{b}, \overline{b}]$:

$$b \le \beta(b). \tag{10}$$

Given that the menu of insurance policies β must satisfy $\beta(b) \le v$ for all $b \in [\underline{b}, \overline{b}]$, any agent with budget b such that $r \le \beta(b)$ is both willing and able to pay the ransom r. We define the attacker's expected payoff from choosing r conditional on policies β as

$$\pi(r; \beta) = \int_{\underline{b}}^{\overline{b}} \left[r\mathbb{1}_{r \le \beta(b)} + v_d \mathbb{1}_{r > \beta(b)} \right] dF(b), \tag{11}$$

Thus, the third requirement that the mechanism must satisfy is the attacker's incentive compatibility constraint: he is maximizing his expected payoff by choosing r_* such that

$$r_* \in \arg\max \pi(r; \beta). \tag{12}$$

The final requirement that an insurance policy menu must satisfy is the agent's incentive compatibility constraint. For each type b, the policy $\beta(b)$ recommended by the designer to type b must be the best policy for b among all policies $\beta(b')$ that b can afford to buy. Given the zero-profit condition, this is equivalent to the following: if, under the policy $\beta(b)$ intended for type b, the agent cannot afford to pay the ransom r_* induced by the mechanism $\langle \beta, r_* \rangle$ (i.e., chosen by the attacker), then there is no alternative policy $\beta(b')$ that type b can afford and that would allow him to pay the ransom. That is, noting that the final budget of a type b that selects policy $\beta(b')$ is $b + \beta(b') - b'$, it must be that for all $b, b' \in [\underline{b}, \overline{b}]$:

$$\beta(b) < r_* \leq \beta(b') \quad \Rightarrow \quad \text{either } \; b < p(b') = \frac{\alpha[\beta(b') - b']}{1-\alpha} \; \text{ or } \; b + \beta(b') - b' < r_*. \tag{13}$$

4.2 The Optimal Insurance Policy Menu

The goal of the insurance designer is to choose the menu of policies β that maximizes the ex-ante expected payoff of the agent. We can write it as a constrained maximization problem with constraints (9)-(13) and choice variables $\beta(b)$ and r_*. Integrating the expression in (8) over the set of types b, the expected payoff of the agent can be written as:

$$(1-\alpha)v + \mathbb{E}[b] + \alpha \int_{\underline{b}}^{\overline{b}} \left(v - r_*\right) \mathbb{1}_{r_* \leq \beta(b)} dF(b) \tag{14}$$

Note that the choice of a policy menu β by the designer does not affect the constant term $(1-\alpha)v+\mathbb{E}[b]$ or the multiplicative constant α in the expected payoff of the agent. In writing the objective function of the insurance designer, we may thus ignore these terms and use the definition of $\pi(r;\beta)$. Also ignoring constraint (13) (we will check later that it is satisfied at the solution), the insurance designer's maximization program can be written as follows:

$$\max_{\beta} \int_{\underline{b}}^{\overline{b}} \left[\left(v - v_d\right) \mathbb{1}_{r_* \leq \beta(b)}\right] dF(b) - \pi(r_*;\beta) + v_d \tag{15}$$

$$\text{subject to} \quad b \leq \beta(b) \leq \min\left\{v, \frac{b}{\alpha}\right\}, \quad \text{and}$$

$$r_* \in \arg\max_{r} \; \pi(r;\beta)$$

The constraint $\beta(b) \geq b$ requires that the final budget of an agent of any type b under any menu of insurance policies be at least as high as the initial budget b. It follows that by choosing the same ransom request r_0 as under no insurance, an attacker can guarantee himself an expected payoff at least as high as his expected payoff π_0 with no insurance; all types b that pay such a ransom request under no insurance are also able to pay under any insurance policy $\beta(b)$. This is formalized in the next result.

Proposition 2. *The attacker's expected payoff under no insurance, π_0, is a lower bound on the attacker's payoff under the optimal (indeed any) insurance policy.*

We now introduce a class of simple policy menus and then later show that the optimal policy belongs to this class.

Define a *single-cap insurance policy menu* β_r as follows:

$$\beta_r(b) = \begin{cases} r & \text{if} \quad b \in \left[\alpha r, r\right] \\ b & \text{otherwise} \end{cases} \tag{16}$$

With a single-cap policy menu, the designer offers a policy consisting of a constant final budget $\beta_r(b) = r \leq v$ to any agent that can afford to buy it and has an initial budget at most equal to r; that is, to all types $b \in [\alpha r, r]$. We say that the policy offers protection to these types. Types that cannot afford to buy the single-cap policy ($b < \alpha r$) or that have an initial budget b above r, on the other hand, are not offered any insurance protection, as the policy gives them a final budget equal to their initial budget, $\beta_r(b) = b$.

The assumption that $v > v_d$ and Proposition 2 suggest that the solution of (15) requires that the insurance designer maximizes the range of types b for which $\beta(b) \geq r_*$, while guaranteeing the attacker a payoff of π_0. As Proposition 3 shows, this is accomplished by carefully choosing the cap r in a single-cap policy menu,

To see how the cap should be optimally chosen by the designer, consider the best reply of the attacker when the policy menu is β_r. First, recall that asking for a ransom below the darknet value of the data v_d is a dominated strategy. The attacker's payoff if he asks for a ransom $\hat{r} \geq v_d$ when the insurance policy menu is β_r is:

$$\pi(\hat{r}; \beta_r) = \begin{cases} (\hat{r} - v_d)[1 - F(\alpha r)] + v_d & \text{if} \quad \hat{r} \in [\alpha r, r] \\ (\hat{r} - v_d)[1 - F(\hat{r})] + v_d & \text{if} \quad \hat{r} \notin [\alpha r, r] \end{cases}$$

Note first that if $\alpha r > \bar{b}$, then the final budget of each agent is the same as the initial budget. Insurance has no effect as no agent can afford to increase its budget by insuring. It is thus optimal for the attacker to set $\hat{r} = r_0$. Second, if $\alpha v_d > \overline{b}$, then no single-cap policy menu can provide protection to any agent type, as the attacker will never ask for a ransom less than $r_0 = v_d$ and no agent type can afford a policy with a cap as high as v_d.

Thus, let us focus on the case in which $\overline{b} \geq \alpha v_d$ and $r \leq \frac{\overline{b}}{\alpha}$. It is immediate that it is never optimal for the attacker to set $\hat{r} \in [\alpha r, r)$, because the attacker's payoff is strictly increasing in $\hat{r}$ in this interval; hence, either $\hat{r} = r$ or $\hat{r} \notin [\alpha r, r]$ are optimal. An upper bound on the payoff that the attacker can make by setting $\hat{r} \notin [\alpha r, r]$ is the profit under no insurance π_0, which he can obtain if $\hat{r} = r_0 \notin [\alpha r, r]$. This is because in such a case, the types that are able to pay the ransom $\hat{r}$ are the same that could pay it without an insurance market; all types with $b \geq \hat{r}$. Thus, when the policy menu is β_r, it is optimal for the attacker to choose $\hat{r} = r$ if the profit from doing so is higher than the profit under no insurance; that is, if the following inequality holds:

$$\big(r - v_d\big)\big[1 - F(\alpha r)\big] + v_d \geq \pi_0 \qquad \text{or equivalently}$$

$$\Delta\pi(r, r_0) = \big(r - v_d\big)\big[1 - F(\alpha r)\big] - \big(r_0 - v_d\big)\big[1 - F(r_0)\big] \geq 0 \tag{17}$$

Note that the agent's incentive compatibility constraint is automatically satisfied by a single-cap insurance policy $\beta_r(b)$ that induces the attacker to choose r as the ransom. To see this observe that $\beta_r(b) < r \leq \beta_r(b')$ requires that $b' \geq \alpha r > b$. There are two cases. First, if $b' > r$, then $\beta_r(b') = b'$ and $x(b') = p(b') = 0$; hence $b + x(b') - p(b') = b < r$. Second, if $b' < r$, then $\beta_r(b') = r$ and $x(b') = \frac{r-b'}{1-\alpha}$; hence $b + x(b') - p(b') = b + r - b' < r$. In both cases, the agent's incentive compatibility constraint is satisfied.

The next lemma characterizes the values of the cap $r \leq \frac{\overline{b}}{\alpha}$ for which the attacker earns more by setting the ransom equal to r under the single-cap policy β_r than under the no-insurance benchmark payoff π_0, ignoring the constraint $r \leq v$; that is, ignoring that the cap cannot be above v as the agent would not want to pay a ransom above the value of the data. It shows that the attacker's profit is at least as high under the single-cap policy β_r than under no insurance if and only if the cap r belongs to a closed interval which includes the ransom request r_0 that is optimal under no insurance.

Lemma 1.

(*i*) *Suppose* $\bar{b} \leq v_d \leq \frac{\bar{b}}{\alpha}$, *and hence* $r_0 = v_d$. *Then the solutions to* $\Delta\pi(r, r_0) = 0$ *such that* $r \leq \frac{\bar{b}}{\alpha}$ *are* $r_1 = r_0 = v_d$ *and* $\bar{r} = \frac{\bar{b}}{\alpha}$. *If* $v_d < \frac{\bar{b}}{\alpha}$, *then* $\Delta\pi(r, r_0) > 0$ *for all* $r \in \left(v_d, \frac{\bar{b}}{\alpha}\right)$.

(*ii*) *Suppose* $v_d < \bar{b}$ *and hence* $v_d < r_0 < \bar{b}$. *There are two solutions to* $\Delta\pi(r, r_0) = 0$ *such that* $r < \frac{\bar{b}}{\alpha}$, *denoted by* r_1 *and* $\bar{r}$ *with* $r_1 < r_0 < \bar{r}$. $\Delta\pi(r, r_0) > 0$ *if* $r \in (r_1, \bar{r})$ *and* $\Delta\pi(r, r_0) < 0$ *if* $r \notin [r_1, \bar{r}]$.

Since $r_1 \leq r_0$ and $r_0 \leq \max\{\bar{b}, v_d\} \leq v$ it follows that $r_1 < v$ and thus it is feasible to set a single-cap policy menu β_{r_1} with cap r_1. On the other hand, there is no guarantee that $\bar{r}$ satisfies the feasibility constraint $\bar{r} \leq v$; that is, it may not be feasible for the designer to set a cap as high as $\bar{r}$. For this reason, it is convenient to define $r_2 = \min\{\bar{r}, v\}$ as the highest cap in a single-cap policy menu β_r that is feasible and induces the attacker to choose r.

We now argue that in the class of single-cap policy menus, there is no loss to the designer in focusing on policy menus β_r that induce the attacker to choose the insurance cap r as the ransom; that is, policy menus β_r with $r \in [r_1, r_2]$. To see this, note that to increase the expected payoff of an agent above the payoff with no insurance, the single-cap insurance policy menu β_r must induce the attacker to choose a ransom $r_* \leq r$, as otherwise the agent types able to pay the ransom and recover the data are the same as under no insurance. Moreover, if the chosen ransom is $r_* < r$, then the designer could increase the agents' expected payoff by offering the policy menu β_{r_*} instead of β_r; that is, by lowering the single-cap to r_* and thus allowing a larger number of agent types to be able to afford to pay to have a final budget equal to the cap.

It then follows that the policy menu β_{r_*} with cap $r_* = r_1$ (the smallest solution of the equation $\Delta(r_*, r_0) = 0$) is the the optimal menu for the designer in the class of feasible single-cap policy menus. The next proposition shows that this is the optimal policy menu of the designer among all feasible policy menus, not just the single-cap menus. With this policy menu the attacker charges $r_* = r_1$ and makes the same profit as if he charged r_0 under no insurance.

Proposition 3. *The single-cap insurance policy menu* β_{r_*} *described in (16), with* r_* *equal to the lowest solution of the equation* $\Delta(r_*, r_0) = 0$ *(i.e.,* $r_* = r_1$*), solves the insurance* *designer's maximization problem (15).*

4.3 Single-cap Insurance vs No Insurance

In this section, we compare the outcome under a single-cap insurance policy menu, including the optimal policy menu, and the outcome when the agent has no access to insurance coverage. We begin with a corollary stating that under the optimal insurance policy the ransom request is lower than if insurance is not available. The corollary follows immediately from Proposition 3, according to which the optimal policy menu is β_{r_1}, and Lemma 1.

Corollary 1. *The attacker's ransom request is lower under the optimal insurance policy menu than with no insurance: $r_1 \leq r_0$, with the inequality being strict if $v_d < \bar{b}$.*

Recall that by design, the attacker's expected payoff under the optimal insurance policy menu is the same as his expected payoff with no insurance coverage. By Proposition 2, any other single-cap insurance menu with cap $r \in [r_1, r_2]$ must yield the attacker an expected payoff at least as high. It remains to determine the effect of a single-cap insurance policy menu β_r, with $r \in [r_1, r_2]$, on the agent's expected payoff.

As we noted in the previous section, if $v_d > \frac{\bar{b}}{\alpha}$, then $r_1 = r_0 = v_d$ and selling in the darknet market is so valuable that no agent type b can afford to pay the zero-profit premium $p(b) = \frac{\alpha[v_d - b]}{1-\alpha}$ associated with the cap $r_1 = v_d$; that is, the optimal insurance policy menu is no insurance, $\beta_{r_*}(b) = b$ for all $b \in [\underline{b}, \bar{b}]$. Hence the agent trivially obtains the same expected payoff $(1-\alpha)v + b$ when insurance is and is not available. We now focus on the interesting case in which $v_d \leq \frac{\bar{b}}{\alpha}$ and consider single-cap insurance policies with cap $r \in [r_1, r_2]$.

The expected payoff with menu β_r of an agent with initial budget b is

$$(1-\alpha)v + b + (v-r)[1 - F(\alpha r)]$$

As the next proposition shows, as long as $r < r_0$ all agent types are weakly better off when the single-cap insurance policy β_r is on offer than when no insurance policy is available. Furthermore, types $b \geq \alpha r$ are strictly better off. If the cap is below the ransom request attackers would make with no insurance, the policy achieves the double benefit of lowering the ransom request and expanding the set of types who are able to pay the ransom and recover the data.

Even if the cap leads to a higher ransom payment than with no insurance, $r > r_0$, the cost to the types that are able to pay with and without insurance may be outweighed by the benefit of expanding the set of types that recover the data, which occurs as long as $\alpha r < r_0$. Thus, the benefits to the agent of a single-cap insurance policy are not knife-edge.

Proposition 4. *Suppose $v_d \leq \frac{\bar{b}}{\alpha}$. Let $r \in [r_1, r_2]$ be the cap in the single-cap insurance policy menu β_r selected by the insurance designer.*

1. *If $r < r_0$, then no agent type is worse off and types $b \geq \alpha r$ are strictly better off with the insurance policy β_r than when no insurance is available.*
2. *If $\alpha r < r_0 < r$, then agent types $b \in [\alpha r, r_0)$ are strictly better off and types $b \geq r_0$ are strictly worse off with the insurance policy β_r than when no insurance is available.*
3. *If $r_0 < \alpha r$, then no agent type is better off and types $b \geq r_0$ are strictly worse off with insurance policy β_r than when no insurance is available.*
4. *There exists a threshold value of the cap $r_t > r_0$ such that the ex-ante expected payoff of the agent is higher with the insurance policy β_r than with no insurance if and only if $r < r_t$.*

To test our theory that a single-cap policy can be beneficial even if the cap is not set optimally, as long as it is not set at too high a level, in the next section we report the results from an experiment in which subjects play a discretized version of the following example.

Example 1. *The initial budget b is uniformly distributed in $[1, 25]$; that is, $F(b) = \frac{b-1}{24}$. The darknet value is $v_d = 6$, the agent's value is $v = 36$ and the probability of an attack is $\alpha = 0.5$. With no insurance, the ransom request is $r_0 = 15.5$ and the optimal insurance cap is approximately $r = 10.6$.*

In the experiment we use two single-cap policies. We call the treatment with the smaller cap, $r_s = 18$, "sophisticated" insurance, and the treatment with the larger cap, $r_n = 36$, "naive" insurance. In evaluating the two policies, we focus on the expected agent (or *target*) surplus and the expected total surplus; that is, the expected payoff gains, conditional on an attack, over the payoff when not paying the ransom. Thus, the agent's surplus is the value of the data to the agent minus the ransom payment if the ransom is paid, and zero if it is not; total surplus is the value of the data to the agent minus the darknet value if the ransom is paid, and zero if it is not.

The ranking of the three regimes depends on the probability of paying the ransom, which is highest with sophisticated insurance $\left(\frac{16}{24}\right)$ and lowest with naive insurance $\left(\frac{7}{24}\right)$. With no insurance it is $\left(\frac{9.5}{24}\right)$.[8] With no insurance, conditional on an attack taking place, expected agent and total surpluses are \$8.11 and \$11.88 respectively. Naive insurance reduces agent surplus to zero (because the equilibrium ransom request under naive insurance is the value of the data) and expected total surplus by \$3.13 to \$8.75. Sophisticated insurance raises the expected agent surplus by \$3.89 to \$12 and expected total surplus by \$8.12 to \$20.

5 The Experiment

5.1 Experimental Design

We ran an in-person laboratory experiment with 8 sessions at UQ's Center for Unified Behavioural and Economic Sciences laboratory in February and March 2024 with a total of 179 subjects.[9] The experiment was coded using oTree (Chen et al., 2016). Screenshots for all parts of the experiment can be found in Appendix C.

The experiment used a mixed within- and between-subject design. The within-subject variation changed whether insurance was available to victims. The availability of insurance was randomized so that 50% of the groups in each round had access to insurance. The between-subject variation changed the type of insurance that was available: either *naive* or *sophisticated*. This variation was assigned at the session level.

In the experiment, subjects play 25 rounds of a game. At the start of each round, subjects are randomly assigned to groups of three. Within each group, one subject is assigned the role of Attacker and the other two the role of Targets. Players' roles were randomized each round, and groups were formed using stranger matching. Each round consists of three stages:

- In **Stage 1**, Targets are informed of their initial budget, which is randomly drawn and

[8] For completeness, note that the expected agent and total surpluses would be highest with the optimal insurance cap, as the ransom paid would be lower and the probability of paying the ransom higher $\left(\frac{19.7}{24}\right)$.

[9] 180 subjects were originally recruited, but one subject withdrew consent and left the session a few minutes after the experiment started.

uniformly distributed over integer values from \$1 to \$24.[10] If insurance is available in the round and their budget is sufficient, Targets may purchase it at this stage.[11]

- In **Stage 2**, the Attacker submits a ransom request to one randomly selected Target (see Appendix Figure C.13). The Attacker is informed of whether insurance was available in the round but does not know the Target's budget or whether insurance was purchased. The ransom request is made by moving a slider, which was uninitialized and had a range of \$6 to \$36 in all treatments.
- In **Stage 3**, final budgets are updated based on insurance decisions and whether the Target was chosen to receive the ransom request. If a Target purchases insurance and is selected to receive the ransom request, her *final budget* (the insurance price cap) is set to \$36 under naive insurance and to \$18 under sophisticated insurance. If no insurance is purchased, her final budget equals her initial budget.[12] If the selected Target's final budget is at least as large as the ransom, they may decide whether to pay. If it is smaller, the ransom is automatically rejected.

Payoffs are determined as follows. The Target who is not selected to receive a ransom request earns her initial budget minus the price paid to insure (if he has chosen to insure), plus the "data value" of \$36. The Target who receives the ransom request receives her final budget. If she accepts the request, she pays the ransom amount but also receives the data value of \$36. The Attacker receives the ransom amount if it is accepted, and otherwise earns the "data darknet value" of \$6. In addition, all subjects receive a baseline payment of \$25. After each round, subjects are shown their payoffs and how these were determined. An example of this feedback for the case of a victim that purchased naive insurance can be found in Appendix Figure C.18. In each round, regardless of treatment, insurance is available to 50% of the groups within a session.

All sessions followed the same procedure. Upon arrival, subjects were seated at computer terminals, where they read a participant information sheet and signed a consent form.

[10]Throughout the paper, the symbol \$ denotes Australian dollars.

[11]When Targets decided whether to purchase insurance and when Attackers chose a ransom, they had access to an interactive tool that displayed payoffs across possible scenarios. The tool included two sliders—one for a hypothetical initial budget and one for a hypothetical ransom demand—and a payoff table that updated accordingly. The table showed the resulting payoffs under each combination of insurance and ransom-payment decisions, or indicated that a given outcome was impossible.

[12]Note that both naive and sophisticated insurance are single-cap policies.

Instructions were read aloud and were also available on their screens and in a printed copy. After reviewing a set of examples, subjects completed a quiz, earning \$1 for each correct answer. They were then shown the correct answers. Once all participants had finished the quiz, the experiment proceeded to the main rounds. After all rounds were completed, a single round was randomly chosen to be the one that counted for subjects' payments. Subjects were reminded of their payments for this round. They then completed a survey including demographic information, a cognitive reflection test (Frederick, 2005), and feedback on the experiment. Summary statistics for demographic variables can be found in Appendix Table B.1. Each session took no longer than two hours and the overall average payment was \$55.25.

5.2 Hypotheses

We now briefly discuss the empirical hypotheses that we tested in our experiment and link them to the theory developed in earlier sections.

Our first hypothesis is about the insurance purchasing behavior of targets. In defining the incentive compatible insurance policies in Section 4.1, we require that agents buy insurance when it is affordable, as it maximizes their payoff. Our initial hypothesis is that targets do indeed do so. As subjects' behavior in experiments often deviate from monetary reward maximization, reflecting a range of underlying behavioral motives, we view this as a test on the practical relevance of our theory.

Hypothesis 1. *Targets buy insurance when it is available to them.*

Our second hypothesis considers the behavior of Attackers. As shown in Example 1, the theory predicts the ransom request without insurance to be 15.5 when the initial budget is uniformly distributed in $[1, 25]$. In the discretized version of the example we use in the experiment, ransom requests of both 15 and 16 are optimal. Naive and sophisticated insurance are single-cap policies designed so that Attackers have an incentive to make ransom requests exactly equal to the cap: 18 with sophisticated insurance and 36 with naive insurance.

Experimental evidence, however, suggests that the exact theory predictions of the subgame perfect equilibrium of games with players assumed to be risk-neutral are "too tight". In particular, small-stakes risk aversion is often observed in laboratory settings and risk-averse Attackers might want to lower their request in order to increase the likelihood that it is accepted by Targets. Attackers, for example, might foresee that Targets are pushed towards refusal by behavioral motives, (e.g., the wish to reciprocate negatively and reject a ransom request perceived as aggressive). Hence we formulate our hypothesis as follows.

Hypothesis 2. *Attackers make ransom requests such that:*

1. *Requests are concentrated at the level of insurance coverage when insurance is available;*
2. *Average ransom requests are highest with naive insurance and lowest without insurance.*

Our final hypothesis examines how different insurance policies affect the distribution of the surplus; that is, the expected payoff gains, conditional on an attack, over the payoff when not paying the ransom. The second paragraph after Example 1 reports how different levels of single-cap policies are expected to affect surplus.

Hypothesis 3. *Relative to no insurance, naive insurance decreases expected victim and total surplus, while sophisticated insurance increases both.*

5.3 Results

Following our preregistration, the analysis that we discuss in this section uses data *only* from rounds 6-25 of the experiment.[13] In our pre-analysis plan, we preregistered the exact regressions that are found in Table 1.[14] The other analyses found in this section were not preregistered and thus can be considered exploratory in nature.

Our first result regards the behavior of targets when they are offered insurance and is consistent with Hypothesis 1.

Result 1. *Targets purchase insurance approximately 80% of the time when it is available. Purchase rates are increasing in the level of the initial budget.*

The evidence for Result 1 can be found in Figure 1. The figure shows the relationship between the target's initial budget and the rate of purchasing insurance for the two insurance

[13] As noted in Footnote 9, one subject left the experiment before it was completed. In this section, we report the results omitting the data from all groups that the subject was a part of. Appendix Table B.3 reproduces Table 1 including the data from rounds in which the subject who left was a member of the group but was not the Attacker. The quantitative results are almost identical.

[14] Appendix Tables B.4, B.5, and B.6 reproduce the results from Table 1 for all rounds and without round and initial budget fixed effects. The results are largely unchanged.

treatments. Recall that in the Naive treatment, insurance protection could only be purchased by those who had an initial budget of at least \$18. In the Sophisticated treatment, insurance protection could only be purchased when initial budgets were between \$9 and \$17. Insurance was purchased 77.4% of the time when it was available in the Naive treatment and 77.8% of the time when it was available in the Sophisticated treatment. Our hypothesis was that *all* targets would purchase insurance when it was available to them, and thus we had no ex-ante reason to believe that insurance purchase rates would be related to the initial budget, conditional on it being purchasable. However, the pattern of purchase rates being increasing in the initial budget is clear from Figure 1, and we confirm that the relationship is statistically significant in Appendix Table B.2.

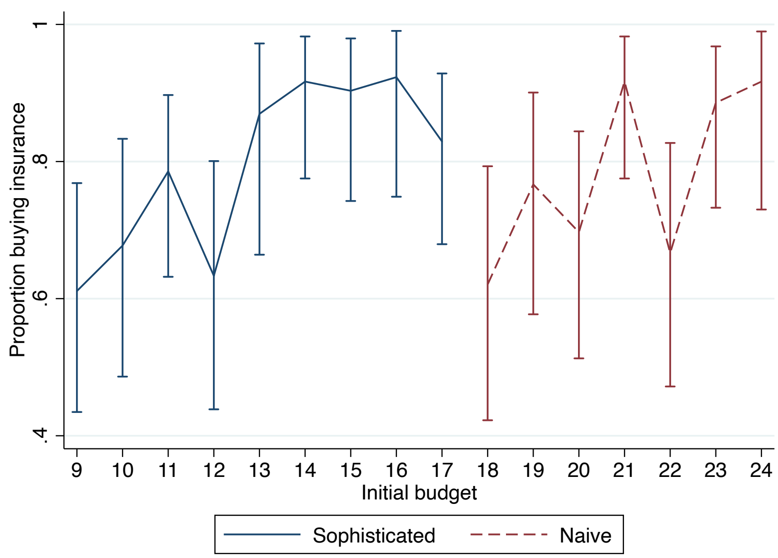


Figure 1: Rates of Purchasing Insurance. Vertical bars indicate 95% confidence intervals for the purchase proportion at each budget level, computed using exact binomial (Clopper-Pearson) intervals.

We now discuss how ransom requests differed between the treatments.

Result 2. *Ransom requests responded systematically to the availability and type of insurance.*

1. *When insurance was unavailable, ransom requests clustered at or just below equilibrium levels (54% of requests are on the range \$12-\$16). The empirical average was just over \$14 and was not statistically different from the theoretical benchmark.*
2. *The introduction of sophisticated insurance increased the average ransom request slightly*

(by roughly \$1.5 relative to no insurance), and concentrated mass at or just below the insurance cap of \$18.

3. *The introduction of naive insurance raised the average ransom request by more than \$6 and substantially increased the variance of requests.*

Figure 2 shows the empirical CDFs of ransom requests, split by treatment. The first notable feature is that the presence of sophisticated insurance affects the distribution of ransom requests. Specifically, it shifts probability mass from requests on the range from \$10 to \$17 to \$18, which was the level covered by sophisticated insurance. Second, relative to no insurance being available, the naive insurance shifts the distribution of ransom requests upward, but without any substantial spike at a single value. Finally, the distributions of ransom requests in the naive and sophisticated treatments in the rounds when insurance was *not* available are almost indistinguishable (a Pearson's chi-squared test does not reject the null hypothesis of the distributions being the same, $\chi^2(28) = 28.7$, $p = 0.43$). This suggests that experience with the two types of insurance did not differentially affect behavior when the insurance was not available.

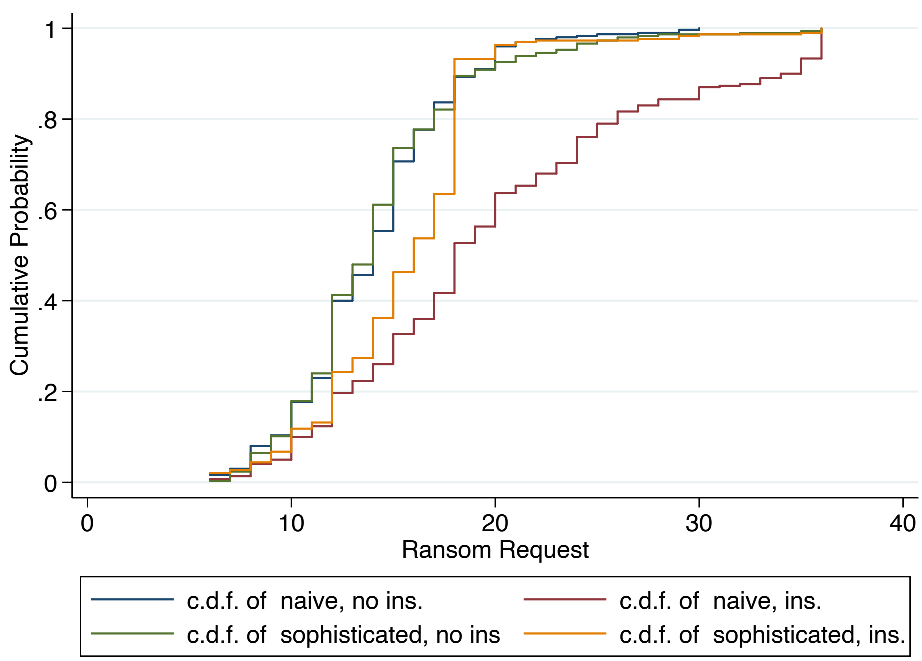


Figure 2: CDFs of Ransom Requests

Column (1) of Table 1 complements Figure 2 with random effects regression analysis of

Ransom Requests on treatment variables. Consistent with the observation that the distributions of ransom requests were very similar across the two treatments when insurance was not available, the coefficient on Sophisticated Treatment is small and not statistically significant. Thus, the average ransom request when insurance is not available was just over $14, which is similar to the theoretically predicted levels of 15 or 16. The coefficient on Insurance Available provides the effect of naive insurance on ransom requests. The estimated coefficient shows that naive insurance increases the average ransom request by over $6, and that this increase is statistically significant at the 1% level. However, this increase is substantially less than the predicted increase of $21. Finally, the effect of sophisticated insurance is the sum of the coefficients, implying an overall increase in average ransom requests of roughly $1.5. This effect is statistically significantly different from both $0 and $3 (the theoretical prediction) at the 1% level.

Overall, victims were overwhelmingly likely to pay the ransom when they had the resources to do so. Of the 657 ransom requests that were below the victim's final budget, 631 were accepted. Payment rates, however, declined as ransom amounts increased: 55 of the 65 requests strictly greater than 18 but no larger than the final budget were paid, while only 2 of the 4 feasible requests at the maximum level of 36 were accepted.

Thus, to summarize, our results support the hypothesis that ransom requests are highest with naive insurance and lowest without insurance. There is also concentration of requests with sophisticated insurance, but at a level lower than the theoretical prediction. With naive insurance, on the contrary, no such concentration is present. This is perhaps due to the extreme nature of the naive insurance policy, under which asking for a ransom as high as the value of the data is the equilibrium outcome. The equilibrium ransom requests with naive insurance thus resemble the equilibrium offer in the ultimatum game, where it is well known that offers typically differ from the equilibrium predictions (e.g., see Güth and Kocher, 2014).

Finally, we consider the effect of insurance on surplus.

Result 3. *Relative to no insurance, naive (sophisticated) insurance decreases (increases) average total surplus and average victim surplus.*

Table 1: Primary Outcomes: Ransom Requests and Surplus

	(1) Ransom Request	(2) Total Surplus	(3) Target Surplus
Soph. Treat	0.26	-0.72	-0.50
	(0.53)	(0.92)	(0.80)
Insurance Available	6.14***	-3.17***	-3.71***
	(0.75)	(0.77)	(0.69)
Soph. Insurance Available	-4.70***	7.22***	5.68***
	(0.86)	(1.18)	(1.01)
Constant	14.2***	0.95	1.37
	(0.76)	(1.45)	(1.19)
Initial Budget FE	No	Yes	Yes
Observations	1180	1180	1180

Notes: Linear regression with Attacker random effects, round fixed effects, and standard errors clustered at the Attacker level. Significance indicated by: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

The evidence for Result 3 can be found in Columns (2) and (3) of Table 1. The interpretation of these coefficients is the same: the effect of introducing naive insurance is the coefficient on Insurance Available, and the effect of sophisticated insurance is the sum of the coefficients on Insurance Available and Soph. Insurance Available. The results show that naive insurance decreases total surplus by $3.17 on average, but that sophisticated insurance increases total surplus by $4.05. These results are quite close to the theoretical predictions that naive insurance ought to reduce average total surplus by $3.13 and sophisticated insurance ought to raise it by $3.89.[15] The results for average target surplus also go in the same direction as the theoretical predictions, but the effects are smaller, especially for naive insurance, which decreased target surplus by $3.71 on average, as opposed to the theoretical prediction of a reduction by $8.11. This is likely due to the fact that attackers did not make the equilibrium offer that, like in an ultimatum game, would extract all the surplus from the target. Sophisticated insurance increased target surplus by $1.97, about half of the theoretical prediction of an increase of $3.89.[16] Thus, overall, Result 3 is consistent with the predictions in Hypothesis 3.

[15] As we pointed out while presenting Example 1, in this setting ransom payments are pure transfers and therefore do not directly affect total surplus. Consequently, differences in total surplus correspond exactly to differences in the proportion of cases where the ransom was paid, and the data were recovered.

[16] Some simple arithmetic shows that targets capture 1.97/4.05≈ 49% of the surplus created by sophisticated insurance. They suffer *more* than 100% of the costs of naive insurance.

6 Extensions

In this section, we analyze several extensions of our model: maximal insurance coverage, uncertainty about the ransom request, risk-averse agents, bargaining about the ransom, attackers entry, choosing protection, and the social cost of darknet market sales.

6.1 Maximal Insurance

Consider an alternative competitive insurance market setting in which insurance companies offer coverage at a fixed per-unit price α, which guarantees that they make zero profit. Agents can buy as much insurance as they want and given that they are risk neutral find it optimal to use all their budget to buy insurance.[17] We can think of this market setting as corresponding to the case of the designer choosing a *maximal insurance* policy that sets $p(b) = b$, and $\beta(b) = x(b) = \frac{b}{\alpha}$. It is clear that if $v_d \geq \frac{\bar{b}}{\alpha}$, then insurance will have no impact as the attacker will sell the data in the darknet market whether there is or there is not an insurance market. Thus, from now on, we assume $v_d < \frac{\bar{b}}{\alpha}$.

Under a maximal insurance policy, the distribution of the final budget available to pay the ransom is given by $F(\alpha b)$. To reduce the number of cases to consider, it is convenient to assume that all agents that can afford to pay the ransom r will find it optimal to pay it. This is the case if the value of the data is sufficiently high, $v \geq \frac{\bar{b}}{\alpha}$. This assumption implies that the probability that the agent pays a ransom r is $1 - F(\alpha r)$ and the attacker expected profit when asking ransom r_m is

$$\pi_m = \int_{\underline{b}}^{\bar{b}} \left[r_m \mathbb{1}_{\alpha r_m \leq b} + v_d \mathbb{1}_{\alpha r_m > b} \right] dF(b) = r_m[1 - F(\alpha r_m)] + v_d F(\alpha r_m) \tag{18}$$

An argument similar to the one used to derive Proposition 1 leads to the following proposition.

Proposition 5. *Suppose $v_d < \frac{\bar{b}}{\alpha} \leq v$. With a maximal insurance policy the attacker asks for the ransom r_m that solves:*

$$r_m = v_d + \frac{1}{\alpha h(\alpha r_m)}. \tag{19}$$

[17]Indeed, maximal insurance coverage would be strictly optimal if there were a probability $\epsilon > 0$, no matter how small, that the attacker chooses the ransom randomly with positive probability density in the interval $[v_d, v]$.

We next compare the attacker's ransom request and payoff without insurance and with maximal insurance coverage and show that from the agent's point of view there is a trade-off: with insurance the ransom request is higher, but the set of types that can recover the data is larger.

Proposition 6. *Suppose $v_d < \frac{\bar{b}}{\alpha} \leq v$. The attacker's ransom request is higher with maximal insurance coverage than with no insurance, $r_m > r_0$. However, the set of agent types that is able to recover the data is higher with a maximal insurance policy than with no insurance, $\alpha r_m < r_0$. The attacker's expected payoff is greater with maximal insurance coverage than with no insurance.*

It follows from the proof of Proposition 6 that with maximal ransom insurance the ransom request decreases with the probability of an attack α; that is, $\frac{dr_m}{d\alpha} < 0$. This is because the more likely an attack is, the more costly it is to buy one unit of insurance in a competitive insurance market. Hence agents will have a lower final budget and this will moderate the attacker's ransom request.

We next consider the difference in the agent's payoff between the cases of maximal insurance coverage and no insurance and show that from an ex-ante point of view the agent is better off with insurance if and only if the value of the data is above a threshold.

Proposition 7. *Suppose $v_d \leq \frac{\bar{b}}{\alpha} \leq v$.*

1. *If $b < \alpha r_m$, then the agent doesn't recover the data and his expected payoff is the same with maximal insurance coverage and with no insurance.*
2. *If $\alpha r_m \leq b < r_0$, then the agent's expected payoff is higher with maximal insurance coverage than with no insurance.*
3. *If $r_0 \leq b$, then the agent's expected payoff is higher with no insurance than with maximal insurance coverage.*
4. *There exists v^* such that the agent's ex-ante expected payoff is higher with maximal insurance coverage than with no insurance if and only if $v > v^*$.*

6.2 Uncertainty about Attackers' Valuations

We have assumed that the darknet value v_d is common knowledge and therefore there is no uncertainty about the ransom request of the attacker. As we have argued, it is sufficient that the insurance companies know this value; the victim does not need to know it. We believe this to be a reasonable assumption in many settings, but in this section we explore what would change if v_d were private information of the attacker and hence the ransom request uncertain.

We argue that it remains the case that in order to reduce the ransom request of the attacker the designer would want to offer insurance policy menus that bunch the final budget of agents on an atom (using single-cap policy menus), but it may be the case that it is optimal to offer a policy menu with more than one cap that induces attackers of different types to ask for different ransoms.

Suppose there are two possible darknet values $\{\underline{v}_d, \overline{v}_d\}$, with $\underline{b} \leq \underline{v}_d < \overline{v}_d \leq \overline{b}$ and q being the probability of $\underline{v}_d$. Let $r_0(\underline{v}_d)$ and $r_0(\overline{v}_d)$ be the ransom requests of the two attacker's types if insurance is not available and let $\underline{r} = r_1(\underline{v}_d)$ and $\overline{r} = r_1(\overline{v}_d)$ be the caps in the single-cap policy menu that would yield the attacker the same profit as under no insurance if his type were known, as derived in Section 4.

If the designer offers a menu with these two caps, then agent types $b \geq \alpha\overline{r}$ select the policy with the higher cap $\overline{r}$ and types $b \in [\alpha\underline{r}, \alpha\overline{r})$ select the policy with the lower cap $\underline{r}$. It is often not the case, however, that the attacker will want to select the ransom that it's meant for him. Attacker type $\underline{v}_d$ may prefer to ask for the higher ransom $\overline{r}$, rather than $\underline{r}$, for example. If this is the case, then one or both caps must be raised and the attacker must obtain an information rent above the profit with no insurance. Depending on the model parameters, it could also be the case that the optimal menu for the insurance designer is to offer a single-cap, as shown in the following example.

Example 2. *Suppose the agent's type is uniformly distributed in the unit interval and that $\underline{v}_d = \frac{1}{3}$ and $\overline{v}_d = \frac{2}{3}$. It is immediate to see that $r_0(\underline{v}_d) = \frac{2}{3}$ and $r_0(\overline{v}_d) = \frac{5}{6}$; without insurance, the profit of attacker type $\underline{v}_d$ is $\underline{\pi}_0 = \frac{4}{9}$ and of attacker type $\overline{v}_d$ is $\overline{\pi}_0 = \frac{25}{36}$. Suppose also that $\alpha = \frac{1}{2}$. Then the minimum cap for type $\overline{v}_d$ is the smallest solution of the equation*

$\left(\overline{r}-\frac{2}{3}\right)\left(1-\frac{\overline{r}}{2}\right)+\frac{2}{3}=\frac{25}{36}$; *that is,* $\overline{r}=\frac{8-\sqrt{14}}{6}$.

We claim that if the designer sets any cap r *below* $\overline{r}$ *meant to be the ransom chosen by attacker type* $\underline{v}_d$*, then attacker type* $\underline{v}_d$ *deviates and asks for ransom* $\overline{r}$ *instead. Hence, in this example, as long as it is sufficiently likely that the attacker's type is* $\overline{v}_d$*, the optimal insurance policy menu is to choose the single-cap* $\overline{r}$ *and have both attacker's type charge the ransom* $\overline{r}$.

To prove our claim, we show that if $r<\overline{r}$*, then the payoff of attacker type* $\underline{v}_d$ *is increasing in the cap* r *designed for him for all* $r\leq\overline{r}$*. Hence type* $\underline{v}_d$ *would choose ransom* $\overline{r}$ *if agents are offered both cap* $r<\overline{r}$ *and cap* $\overline{r}$*. The profit of type* $\underline{v}_d$ *when the cap is* r*, and hence all agents types* $b\geq\frac{r}{2}$ *buy protection, is* $\left(r-\frac{1}{3}\right)\left(1-\frac{r}{2}\right)+\frac{1}{3}$*. The profit is a concave function of* r *with a maximum at* $r=\frac{7}{6}>\overline{r}$.

6.3 Risk Aversion

While it is natural to assume that the potential victim of a ransomware attack, typically a company, is risk neutral, it is important to check how our results would be affected if the agent were risk averse. Thus, suppose that the agent has a concave utility function u over the monetary payoff.

Under no insurance, the ransom request r_0 of the attacker is the same as if the agent were risk neutral, as the victim will pay the ransom r as long as $r\leq b$ as in the case of a risk neutral agent. To shorten the discussion, in this section we will assume that $v_d<\overline{b}$ so that $r_0<\overline{b}$. This implies that the optimal single-cap policy under risk neutrality is β_{r_*} with $r_*=r_1<r_0<\overline{b}$. We will argue that β_{r_1} remains the optimal policy when the agent is risk averse.

None of the insurance policy feasibility conditions change when the agent is risk averse. The set of implementable single-cap policy menus β_r also does not change; it must be $r\in[r_1,r_2]$ to induce the attacker to choose a ransom equal to the cap (and not higher than the value of the data v). What changes is the agent's expected payoff (8). Using the zero profit condition and the final budget definition, under a single-cap policy menu with cap $r<\overline{b}$ we have that if $b\geq\alpha r$, then it is $b-p(b)=\frac{b-\alpha r}{1-\alpha}$ and $b-p(b)+x(b)=r$. It follows that the agent's expected utility is:

$$\int_{\underline{b}}^{\alpha r}[(1-\alpha)u(v+b)+\alpha u(b)]dF(b)+\int_{\alpha r}^{r}\left[(1-\alpha)u\left(v+\frac{b-\alpha r}{1-\alpha}\right)+\alpha u(v)\right]dF(b) \quad (20)$$
$$+\int_{r}^{\bar{b}}\left[(1-\alpha)u\left(v+b\right)+\alpha u(v+b-r)\right]dF(b).$$

Proposition 8. *Assume the agent is risk averse with concave utility u over monetary payoffs and $v_d < \bar{b}$. The ransom request of the attacker with no insurance is the same as when the agent is risk neutral. The single-cap insurance policy menu that maximizes the insurance designer's expected utility in (20 subject to the feasibility constraints is the same that solves the maximization problem (8) when the agent is risk neutral.*

6.4 Bargaining about the Ransom

We have assumed that the attacker is able to make a take-it-or-leave-it ransom offer. In this section we discuss how our results change if the victim has some bargaining power.

The simplest way to model bargaining is to assume that the ability to make a take-it-or-leave-it offer is randomly assigned to the attacker with probability λ and to the victim with probability $(1-\lambda)$. The attacker will accept any offer at least as high as v_d and thus the victim optimal offer is exactly $r = v_d$.

Assuming again $v_d < \bar{b}$ to focus on the most interesting case, we have that the expected ransom payment with no insurance is $\widehat{r}_0 = \lambda r_0 + (1-\lambda)v_d$. The cap r_1^{λ} of the optimal single-cap insurance policy is lower than the cap r_1 that is optimal when the attacker has all the bargaining power. When the attacker is selected as the one making the take-it-or-leave-it offer, the optimal cap must yield the attacker the same profit π_0 as from charging r_0 (i.e., at least the same profit as when making the offer under no insurance). However, the expected outlay of the insurance company and hence the premium charged is lower when the agent has a higher bargaining power, allowing more agent types to insure and pay the ransom and thus increasing the attacker's profit above π_0 if the cap and ransom request is r_1.

Unsurprisingly, both with and without insurance an increase in bargaining power benefits the victim. It does so by decreasing the expected ransom and increasing the probability of the agent being able to pay. With insurance an increase in bargaining power also reduces the insurance premium, and thus increases the mass of agents able to insure.

One reason often given as a benefit of ransomware insurance is that the insurance company is often in charge of bargaining with the attacker and might have a stronger bargaining power than the agent. This benefit can be seen quite simply in our model. If it were the case that the bargaining power λ of the attacker is lower when the agent has insurance and can let the insurance company bargain for him, then it is immediate that the optimal single-cap insurance policy (indeed any policy with cap less than r_t) would be even more beneficial to the agent.

6.5 Attackers' Entry and Choosing Protection

One of the reason often given against ransomware insurance is that by facilitating the payment of a ransom it might encourage further attacks; that is, it might encourage entry in the ransomware market. This could be modeled by assuming that the probability α of an attack is an increasing function of the attacker's expected profit (e.g., because the expected profit must cover the entry cost in the ransomware market). Our model shows that the presence of a ransomware insurance market cannot reduce an attacker's profit, unless insurance companies are able to have stronger bargaining power in dealing with attackers. It thus lends some support to the concern that an active insurance market my encourage ransomware attacks. Note, however, that under the optimal insurance policy the attacker obtains the same profit as under no insurance. The optimal insurance policy thus raises efficiency without encouraging entry by attackers.

We have also assumed that the probability of a successful attack is given and does not depend on the protection effort of the agent. We now extend our model to include the choice of protection effort. Suppose that the probability of an attack is a function of the protection effort. Let e denote the monetary effort level and $\alpha(e)$ the probability of an attack, with the twice differentiable function α being decreasing and convex. For simplicity we assume that the protection effort does not affect the initial budget b available to pay the ransom. We show that if the insurance designer does not include any protection provisions in the insurance policy and agents privately choose their effort level, then protection and insurance are substitutes. The agents that are able to insure choose less protection than if insurance is not available.

Proposition 9. *Suppose $v_d < \bar{b}$. Let $r \in [r_1, r_0)$ be the cap in the single-cap insurance policy menu β_r. Suppose agents choose their protection level privately, with no requirements from the insurance policy.*

1. *With no insurance, low-budget types $b < r_0$ choose a protection effort e_0^L which is higher than the protection level e_0^H chosen by high-budget types $b \geq r_0$.*
2. *With insurance policy menu β_r, types $b < \alpha(e^H)r$ choose the same protection level e_0^L as with no insurance, while types $b \geq \alpha(e^H)r$ choose a lower protection level e^H than the protection level e_0^H chosen by the agents able to pay the ransom with no insurance.*

When agents privately choose their protection effort, they forgo the benefits from coordinating on a higher protection level; that is, they select a suboptimal level of protection effort. An increase in protection by all types that buy insurance reduces the probability of an attack of an insured agent. It thus reduces the price of insurance and expands the set of types who can insure and pay the ransom. In addition, it reduces the cap r_1 in the optimal single-cap policy menu β_{r_1}. This is because when more types are able to buy insurance, the attacker makes a higher expected profit for any given ransom request r. Thus, if protection effort is observable and is included as a requirement of the insurance policy β_{r_1}, then the insurance designer will be able to raise the protection effort, increase the set of types that get insurance coverage, and reduce the ransom request of the attackers.

6.6 Social Cost of Darknet Market Sales

We have assumed that the only cost of a darknet sale of data is borne by the victim, but what if there is also a social cost? For example, there might be economies of scale in the darknet market leading to more attacks after data disclosure, or the currently disclosed data may impose damage to other victims of data loss.

Let L be the additional social loss if the data is sold in the dark web. Subtracting the expected social loss $\alpha \int_{\underline{b}}^{\bar{b}} L \mathbb{1}_{r_* > \beta(b)} dF(b)$ from (15), the designer objective function can be written as

$$(1-\alpha)v + \mathbb{E}b + \alpha \int_{\underline{b}}^{\bar{b}} \big(v + L - v_d\big) \mathbb{1}_{r_* \leq \beta(b)} dF(b) - \alpha\big[\pi(r_*; \beta) + L - v_d\big].$$

It is thus clear that the optimal insurance policy does not change. The only effect is that now the designer may find it beneficial to subsidize the agent to pay a ransom above the private value v of the data and up to the total "social value" $V + L$.

7 Conclusions

In this paper, we have taken a mechanism design approach to examine the welfare implications of ransomware insurance. We have shown that the reimbursement limits, or caps, that we observe more and more frequently in practice are a fundamental ingredient of the competitive insurance policy that maximizes the expected utility of potential victims. Optimally designed reimbursement caps prevent attackers from strictly benefiting from the existence of insurance and thus do not induce entry of attackers.

Even if sub-optimal, capped ransomware insurance yields two complementary benefits. First, victims recover encrypted data more frequently, since insurance facilitates payment where recovery is optimal. Second, victims face lower equilibrium ransom demands—an effect driven by attackers' strategic response to the capped reimbursement scheme. Attackers' profits increase if the reimbursement cap is not set optimally, but the gains to victims are not offset by increased rents to attackers and thus social surplus increases if insurance becomes available.

Thus, in contrast to common concerns in public policy circles, our results show that appropriately designed ransomware insurance can increase the welfare of potential victim organizations without overly benefiting cyber criminals. More broadly, our findings highlight the central role of contract design in ensuring that cyber-insurance markets complement, rather than undermine, broader cybersecurity objectives.

References

Arrow, K. (1963). Uncertainty and the welfare economics of medical care. *American Economic Review 53*(4, Part 1), 941–973.

Baker, T. and A. Shortland (2023). Insurance and enterprise: cyber insurance for ransomware. *The Geneva Papers on Risk and Insurance - Issues and Practice 48*, 275–299.

Block, W. and P. Tinsley (2008). Should the law prohibit paying ransom to kidnappers? *American Review of Political Economy 6*, 40–45.

Carannante, M. and A. Mazzoccoli (2025). An analytical review of cyber risk management by insurance companies: A mathematical perspective. *Risks 13*(8), 144.

Cartwright, A., E. Cartwright, J. MacColl, G. Mott, S. Turner, J. Sullivan, and J. R. C. Nurse (2023). How cyber insurance influences the ransomware payment decision: Theory and evidence. *The Geneva Papers on Risk and Insurance - Issues and Practice 48*, 300–331.

Cartwright, E., J. Castro, and A. Cartwright (2019). To pay or not: Game theoretic models of ransomware. *Journal of Cybersecurity 5*.

Chade, H. and E. Schlee (2012). Optimal insurance with adverse selection. *Theoretical Economics 7*, 571–607.

Chainalysis Team (2024). Ransomware hit $1 billion in 2023. `https://www.chainalysis.com/blog/ransomware-2024/`.

Chainalysis Team (2025). Crypto ransomware 2025: 35% year-over-year decrease in ransomware payments. `https://www.chainalysis.com/blog/crypto-crime-ransomware-victim-extortion-2025/`.

Chen, D. L., M. Schonger, and C. Wickens (2016). oTree - an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance 9*, 88–97.

Control Risks (2022). Kidnap for ransom in 2022. https://www.controlrisks.com/our-thinking/insights/kidnap-for-ransom-in-2022.

CRO Forum (2023). Ransomware threats, countermeasures and trends within the insurance industry. https://www.thecroforum.org/wp-content/uploads/2023/03/Ransomware-Threats-Countermeasures-and-Trends-within-the-Insurance-Industry.pdf.

Fink, A. and M. Pingle (2014). Kidnap insurance and its impact on kidnapping outcomes. *Public Choice 160*(3/4), 481–499.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives 19*(4), 25–42.

Geneva Association (2022). Ransomware: An insurance market perspective. https://www.genevaassociation.org/sites/default/files/ransomware_report_online.pdf?

Güth, W. and M. Kocher (2014). More than thirty years of ultimatum bargaining experiments: Motives, variations, and a survey of the recent literature. *Journal of Economic Behavior & Organization 108*, 396–409.

Jones, T. (2025). Piracy on the rise across the world — maritime watchdog. https://www.dw.com/en/piracy-on-the-rise-across-the-world-maritime-watchdog/a-72249554.

Kharraz, A., W. Robertson, D. Balzarotti, L. Bilge, and E. Kirda (2015). Cutting the gordian knot: A look under the hood of ransomware attacks. In *Proceedings of the 12th International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment - Volume 9148*, DIMVA 2015, Berlin, Heidelberg, pp. 3–24. Springer-Verlag.

Konrad, K. A. and S. Skaperdas (1997). Credible threats in extortion. *Journal of Economic Behavior & Organization 33*(1), 23–39.

Lapan, H. and T. Sandler (1988a). To bargain or not to bargain: That is the question. *The American Economic Review 78*(2), 16–21.

Lapan, H. E. and T. Sandler (1988b). To bargain or not to bargain: That is the question. *American Economic Review, Papers and Proceedings 78*(2), 16–21.

Laszka, A., S. Farhang, and J. Grossklags (2017). On the economics of ransomware. In S. Rass, B. An, C. Kiekintveld, F. Fang, and S. Schauer (Eds.), *Decision and Game Theory for Security*, pp. 397–417. Springer International Publishing.

Mossin, J. (1968). Aspects of rational insurance purchasing. *Journal of Political Economy 76*(4, Part 1), 553–568.

Paquet-Clouston, M., B. Haslhofer, and B. Dupont (2019, 05). Ransomware payments in the bitcoin ecosystem. *Journal of Cybersecurity 5*(1), tyz003.

Risk & Insurance Editorial Team (2025). Global cyber insurance market reaches $16.6 billion in 2024. `https://riskandinsurance.com/global-cyber-insurance-market-reaches-16-6-billion-in-2024/`.

Rothschild, M. and J. E. Stiglitz (1976). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *Quarterly Journal of Economics 90*, 629–649.

Selten, R. (1988). A simple game model of kidnapping. In R. Selten (Ed.), *Models of Strategic Rationality*, pp. 77–93. Springer, Dordrecht.

Sophos (2025). The state of ransomware 2025. `https://news.sophos.com/en-us/2025/06/24/the-state-of-ransomware-2025/`.

Stiglitz, J. E. (1977). Monopoly, non-linear pricing and imperfect information: The insurance market. *Review of Economic Studies 40*, 407–430.

US Office of the Director of National Intelligence (2025). Worldwide ransomware, 2024: Increasing rate of attacks tempered by law enforcement disruptions. `https://www.odni.gov/files/CTIIC/documents/products/Worldwide_Ransomware_2024.pdf`.

A Proofs

Proof of Lemma 1.

(*i*) If $r_0 = v_d$, then $\Delta\pi(r, r_0) = \big(r - v_d\big)\big[1 - F(\alpha r)\big]$ and the solutions to $\Delta\pi(r, r_0) = 0$ are $r = v_d = r_0$ and $r \geq \frac{\bar{b}}{\alpha}$.

(*ii*) Suppose $v_d < r_0 < \bar{b}$. Note that the function $\Delta\pi(r, r_0)$ is continuous in r. That there exists a solution r_1 to $\Delta\pi(r, r_0) = 0$ with $v_d < r_1 < r_0$ follows from $\Delta\pi(v_d, r_0) < 0$ and $\Delta\pi(r_0, r_0) > 0$. Since $\Delta\pi\left(\frac{\bar{b}}{\alpha}, r_0\right) < 0$ there also exists a solution $\bar{r}$ with $r_0 < \bar{r} < \frac{\bar{b}}{\alpha}$.

It remains to show that there are no other solutions. Note that the derivative

$$\frac{\partial \Delta\pi(r, r_0)}{\partial r} = 1 - F(\alpha r) - \alpha(r - v_d) f(\alpha r) = [1 - F(\alpha r)]\,[1 - \alpha(r - v_d) h(\alpha r)]$$

is decreasing in r at any stationary point r_s of $\Delta\pi(r, r_0)$, because the hazard rate $h(\alpha r)$ is increasing in r. Hence the function $\Delta\pi(r, r_0)$ has a unique stationary point, a maximum at $r_m \in \left(r_0.\frac{\bar{b}}{\alpha}\right)$; $\Delta\pi(r, r_0)$ is positive in $(r_1, \bar{r})$ and negative outside the interval $[r_1, \bar{r}]$. There are no other solutions to $\Delta\pi(r, r_0) = 0$ besides r_1 and $\bar{r}$ $\square$

Proof of Proposition 3. Suppose that the optimal policy menu β (not necessarily single-cap) that solves the maximization problem (15) induces the choice of r_* by the attacker. We will argue that there is a single-cap policy menu that achieves at least the same payoff for the designer. There are three cases.

First, if there are types b such that $b < \alpha r_*$ (i.e., if $r_* > \frac{b}{\alpha}$), then these agent types cannot pay the ransom r_* under any feasible insurance policy β, because it must be $\beta(b) \leq \frac{b}{\alpha} < r_*$. Consider a policy menu β_* that for all such types b sets $\beta_*(b) = b$ and leaves the final budget of the other types the same as under policy menu β. Such a policy variation does not change the payoff of any agent type, as there is no change in the types b that are able to pay the ransom r_*. It follows that r_* remains optimal for the attacker and the attacker's profit also doesn't change. Hence if β was optimal, then β_* remains optimal for the designer.

Second, if in the policy menu β and hence in β_* there are agent types b with $b \geq r_*$ (i.e., if $r_* \leq \bar{b}$), then these types are able to pay the ransom under any feasible insurance policy menu β because it must be $\beta(b) \geq b$. Because increasing the final budget of agents that have enough initial budget to pay the ransom might induce the attacker to increase his ransom request, varying the policy menu from β_* to β_{**} by setting $\beta_{**}(b) = b$ for all types in this set and leaving the final budget unchanged for all other types remains an optimal insurance policy menu for the designer, because it remains optimal for the attacker to set the ransom equal to r_*.

Third, if in the policy menu β and hence in β_* and β_{**} there are agent types b such that $\alpha r_* \leq b < r_*$ (i.e., if $\underline{b} \leq r_* \leq \frac{\bar{b}}{\alpha}$), then varying the policy menu from β_{**} to β_{r_*} by setting $\beta_{r_*}(b) = r_*$ for all types in the interval $[\alpha r_*, r_*]$ and leaving the final budget the same as in the menu β_{**} for all other types, increases the payoff of any agent b in this interval for which $\beta_{**}(b) < r_*$ and leaves the payoff unchanged of any type for which $\beta_{**}(b) \geq r_*$. To see this note that under the policy menu β_{r_*} all types in this interval are able to pay the ransom r_*, and if it was optimal for the attacker to choose r_* when the policy menu chosen was β_{**}, then it remains optimal for the attacker to set the ransom equal to r_* with the menu β_{r_*}. Thus, if β_{**} is an optimal menu it must coincide with β_{r_*} for all $b \in [\alpha r_*, r_*]$.

We have established that if the policy menu β is a solution of the maximization problem (15) with r_* being the ransom chosen by the attacker, then the single-cap policy menu β_{r_*}, with $\beta_{r_*}(b) = r_*$ if $b \in [\alpha r_*, r_*]$ and $\beta_{r_*}(b) = b$ otherwise, is also a solution. It then follows from Lemma 1 that the policy menu β_{r_1} with cap r_1 equal to the lowest solution of $\Delta\pi(r, r_0) = 0$ solves the designer's maximization problem (15). □

Proof of Proposition 4.

1. If $r < r_0$, then the ransom request is lower with the insurance policy β_r than with no insurance. In addition, types $b \geq \alpha r$ are able to buy protection and hence pay the ransom. Since $\alpha r < r_0$, more types than with no insurance can recover the data.
2. If $\alpha r < r_0 < r$, then the ransom request is higher with the insurance policy β_r than with no insurance. Types $b \geq r_0$ are able to pay the ransom and recover the data in both cases and are thus strictly worse off with the insurance policy β_r. Types $b \in [\alpha r, r_0)$, on the other hand can pay the ransom and recover the data only under the insurance policy. Hence they are strictly better off with the insurance policy.
3. If $r_0 < \alpha r$, then the ransom request is higher with the insurance policy β_r than with no insurance. In addition, types $b \in [r_0, \alpha r)$ are not able to buy protection and hence

pay the ransom with policy β_r, while they would be able to pay the ransom with no insurance.

4. The ex-ante expected payoff of the agent with insurance is $(1-\alpha)v + \mathbb{E}[b] + (v-r)[1-F(\alpha r)]$ while the ex-ante expected payoff with no insurance is $(1-\alpha)v + \mathbb{E}[b] + (v-r_0)[1-F(r_0)]$. Let $r_t \geq r_0$ be the ransom at which the two are equal; that is, the unique solution (or the smallest solution if $r_0 = v_d \geq \bar{b}$) of the equation

$$(v-r)[1-F(\alpha r)] = (v-r_0)[1-F(r_0)].$$

As the lhs of the equation is decreasing in r, it follows that the ex-ante expected payoff under insurance is higher if and only if $r < r_t$. As an immediate corollary, note that the agent's ex-ante expected payoff is higher under the optimal insurance policy than under no insurance, as $r_* = r_1 < r_0 \leq r_t$. □

Proof of Proposition 6.

Note that if $\alpha = 1$, then $r_m = r_0$. We can thus compare the ransom requests in the two settings by applying the implicit function theorem to equation (19) defining r_m as a function of α. Letting $h'(\cdot) > 0$ be the derivative of the hazard function, we obtain:

$$\left[1 + \frac{\alpha^2 h'(\alpha r_m)}{(\alpha h(\alpha r_m))^2}\right] dr_m + \left[\frac{h(\alpha r_m) + \alpha r_m h'(\alpha r_m)}{(\alpha h(\alpha r_m))^2}\right] d\alpha = 0$$

Since both expressions in square brackets are positive, it follows that $\frac{dr_m}{d\alpha} < 0$ and hence $r_m > r_0$.

By (2) and (19) we have

$$r_0 - \frac{1}{h(r_0)} = v_d \qquad \text{and} \qquad \alpha r_m - \frac{1}{h(\alpha r_m)} = \alpha v_d.$$

Since $\alpha v_d < v_d$ and $r - \frac{1}{h(r)}$ is increasing in r, we have $\alpha r_m < r_0$.

Let $\pi_m(\alpha)$ be the value function of the attacker's profit maximization problem. The expected payoff of the attacker is $\alpha \pi_m(\alpha)$ in the setting with maximal insurance coverage, and $\alpha \pi_m(1)$ when there is no insurance. Since $\pi_m(\alpha)$ is differentiable in α, we can apply the envelope theorem to derive from (18) that $\frac{d\pi_m(\alpha)}{d\alpha} = -r_m(r_m - v_d) f(\alpha r_m) < 0$. It then follows that the attacker's expected payoff is higher under a maximal insurance coverage policy than with no insurance. □

Proof of Proposition 7.

1. If $b < \alpha r_m$, then $\beta(b) = \frac{b}{\alpha} < r_m$; with maximal insurance coverage, the agent's final budget is below the ransom request and so it will not be able to pay the ransom. Since $b < r_0$ the agent is not able to pay the ransom in the case of no insurance either. Hence in both settings the agent obtains the same expected payoff $(1-\alpha)v + b$.
2. If $\alpha r_m \leq b < r_0$, then the agent's final budget is such that he is able to pay the ransom under maximal insurance coverage, but not without insurance, since $\beta(b) = \frac{b}{\alpha} \geq r_m$, but $b < r_0$. It follows that the agent is better off with maximal insurance coverage; his expected payoff is $(1-\alpha)v + b + \alpha(v - r_m)$, while with no insurance his expected payoff is $(1-\alpha)v + b$.
3. If $r_0 \leq b$, then the agent's final budget is sufficient to pay the ransom request in both settings, since since $\beta(b) = \frac{b}{\alpha} \geq r_m$ and $b \geq r_0$. Since the ransom request is lower with no insurance, $r_0 \leq r_m$, it follows that the agent has a higher expected payoff without insurance, $(1-\alpha)v + b + \alpha(v - r_0)$, than with maximal insurance coverage, $(1-\alpha)v + b + \alpha(v - r_m)$.
4. The ex-ante expected gain from having maximal insurance coverage over having no insurance is given by:

$$G_m = \alpha\big[1 - F(\alpha r_m)\big](v - r_m) - \alpha\big[1 - F(r_0)\big](v - r_0).$$

   Since by Proposition 6 we have $\alpha r_m < r_0$, the result follows from G_m being linearly increasing in v. $\square$

Proof of Proposition 8.

Differentiating the expected utility in (20) with respect to r yields:

$$\begin{aligned}\Big[\alpha(1-\alpha)u(v+\alpha r) + \alpha^2 u(\alpha r)\Big] f(\alpha r) + \Big[(1-\alpha)u(v+r) + \alpha u(v)\Big] f(r)\\ -\Big[\alpha(1-\alpha)u(v) + \alpha^2 u(v)\Big] f(\alpha r) - \int_{\alpha r}^{r} \alpha u' \left(v + \frac{b - \alpha r}{1-\alpha}\right) dF(b)\\ -\Big[(1-\alpha)u(v+r) + \alpha u(v)\Big] f(r) - \alpha \int_{r}^{\bar{b}} u'\left(v + b - r\right) dF(b)\end{aligned}$$

which is equal to:

$$\Big[\big[(1-\alpha)u(v+\alpha r)+\alpha u(\alpha r)\big]-u(v)\Big]\alpha f(\alpha r)$$
$$-\int_{\alpha r}^{r}\alpha u'\left(v+\frac{b-\alpha r}{1-\alpha}\right)dF(b)-\alpha\int_{r}^{\bar{b}}u'\left(v+b-r\right)dF(b)<0$$

where the inequality follows from the fact that $u'>0$, $v>r$, and, by the concavity of u:

$$(1-\alpha)u(v+\alpha r)+\alpha u(\alpha r)<u\big((1-\alpha)(v+\alpha r)+\alpha\alpha r\big)=u\big((1-\alpha)v+\alpha r\big)<u(v)$$

Thus, the agent's expected utility is decreasing in the cap r and it is thus optimal for the insurance designer to choose the same single-cap insurance policy β_{r_1} as when the agent is risk neutral. □

Proof of Proposition 9.

1. With no insurance, agent types $b<r_0$ choose protection effort to maximize $[1-\alpha(e)]v+b-e$. The optimal choice e_0^L satisfies the first order condition $\alpha'(e_0^L)=-\frac{1}{v}$.
   Agent types $b\geq r_0$ choose protection effort to maximize $v+b-e-\alpha(e)r_0$. The optimal choice e_0^H satisfies the first order condition $\alpha'(e_0^H)=-\frac{1}{r_0}$.
   It follows from the convexity of the function α that types with a low budget $b<r_0$ choose higher protection effort than types with a high budget, $e_0^L>e_0^H$.
2. Now suppose a single-cap insurance policy menu β_r is available. Let e^L and e^H be the effort choices of those that are not and those that are able to raise the budget available to pay the ransom under policy β_r. Note that the price of a final budget equal to r is $\alpha(e^H)$.
   Agent types $b<\alpha(e^H)r$ choose protection effort to maximize $[1-\alpha(e)]v+b-e$. The optimal choice is $e^L=e_0^L$, the same effort level as with no insurance, as the first order condition is $\alpha'(e^L)=-\frac{1}{v}$.
   Agent types $b\geq\alpha(e^H)r$ choose protection effort to maximize $v+b-e-\alpha(e)r$. The optimal choice e^H satisfies the first order condition $\alpha'(e^H)=-\frac{1}{r}$.
   It follows from the convexity of the function α that types with a low budget $b<\alpha(e^H)r$ choose higher protection effort than types with a high budget, $e^L>e^H$. It also follow that if $r<r_0$, then types with $b<\alpha(e^H)r$ choose the same level of protection as with

no insurance, while types with $b \geq \alpha(e^H)r$ choose a lower level of protection than with no insurance. $\square$

B Additional Empirical Results

Table B.1: Summary Statistics

	Mean	Std. Dev.
CRT Score	1.45	1.20
Female	0.58	0.50
Age	22.83	4.82
English	0.24	0.43
Economics	0.37	0.48
Subjects	179.00	

Notes: CRT Score is the number of correct answers on a Cognitive Reflection Test, ranging from 0 to 3. Female, English, and Economics are equal to one if the subjects report being female, speaking English as a first language, and majoring in Economics, respectively.

Table B.2: Insurance Purchase Decision

	(1) Purchased Insurance	(2) Purchased Insurance
Soph. Treat	0.025	0.40
	(0.052)	(0.38)
Initial Budget		0.039**
		(0.016)
Initial Budget × Soph. Treat		-0.0048
		(0.019)
Constant	0.76***	-0.050
	(0.041)	(0.34)
Observations	416	416

Notes: Linear regression with Target random effects and standard errors clustered at the Target level. The data for this regression are from only those rounds in which insurance was available and Targets had an initial budget that allowed them to purchase insurance. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1.

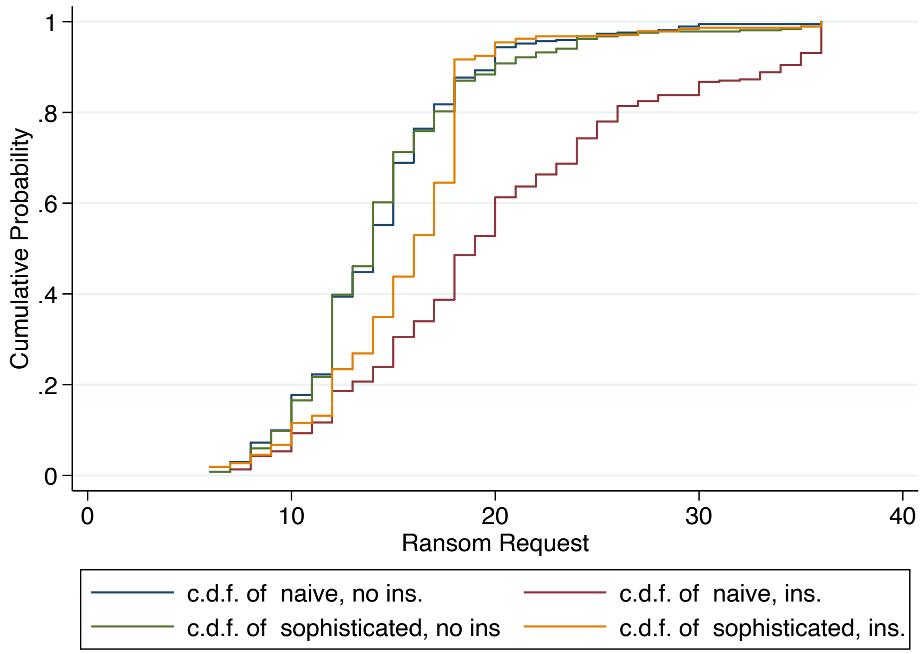


Figure B.1: CDFs of Ransom Requests. This figure shows the CDF of ransom requests split across treatments for all rounds (i.e., not restricting to rounds 6-25). The figure does not differ substantially from Figure 2.

Table B.3: Primary Outcomes: Ransom Requests and Surplus

	(1) Ransom Request	(2) Total Surplus	(3) Target Surplus
Soph. Treat	0.24	-0.86	-0.62
	(0.53)	(0.91)	(0.80)
Insurance Available	6.14***	-3.15***	-3.69***
	(0.75)	(0.77)	(0.69)
Soph. Insurance Available	-4.69***	7.37***	5.78***
	(0.85)	(1.18)	(1.01)
Constant	14.2***	1.64	0.37
	(0.78)	(1.26)	(1.03)
Initial Budget FE	No	Yes	Yes
Observations	1192	1192	1192

Notes: Linear regression with Attacker random effects, round fixed effects, and standard errors clustered at the Attacker level. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1. This table reproduces Table 1 including the data from rounds in which the subject that left the experiment would have taken part. The results do not change substantively.

Table B.4: Ransom Requests

	(1) Ransom Request	(2) Ransom Request	(3) Ransom Request
Soph. Treat	0.25	0.11	0.26
	(0.53)	(0.54)	(0.53)
Insurance Available	6.15***	6.12***	6.14***
	(0.75)	(0.71)	(0.75)
Soph. Insurance Available	-4.69***	-4.73***	-4.70***
	(0.85)	(0.81)	(0.86)
Constant	14.0***	15.5***	14.2***
	(0.36)	(0.80)	(0.76)
Round FE	No	Yes	Yes
Last 20 Rounds	Yes	No	Yes
Observations	1180	1475	1180

Notes: Linear regression with subject random effects and standard errors clustered at the attacker level. Significance indicated by: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.
This table replicates Column (1) of Table 1 both without using budget fixed effects and using the data from all 25 rounds. The results do not change substantively.

Table B.5: Total Surplus

	(1) Total Surplus	(2) Total Surplus	(3) Total Surplus
Soph. Treat	-0.34	-0.013	-0.72
	(1.30)	(0.86)	(0.92)
Insurance Available	-2.90***	-3.17***	-3.17***
	(0.98)	(0.72)	(0.77)
Soph. Insurance Available	6.63***	6.67***	7.22***
	(1.70)	(1.15)	(1.18)
Constant	13.2***	0.13	0.95
	(0.82)	(1.57)	(1.45)
Round & Initial Budget FE	No	Yes	Yes
Last 20 Rounds	Yes	No	Yes
Observations	1180	1475	1180

Notes: Linear regression with subject random effects and standard errors clustered at the attacker level. Significance indicated by: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.
This table replicates Column (2) of Table 1 both without using budget fixed effects and using the data from all 25 rounds. The results do not change substantively.

Table B.6: Target Surplus

	(1) Target Surplus	(2) Target Surplus	(3) Target Surplus
Soph. Treat	-0.13	-0.012	-0.50
	(1.05)	(0.74)	(0.80)
Insurance Available	-3.44***	-3.71***	-3.71***
	(0.79)	(0.61)	(0.69)
Soph. Insurance Available	5.13***	5.38***	5.68***
	(1.33)	(0.94)	(1.01)
Constant	10.2***	0.31	1.37
	(0.69)	(1.19)	(1.19)
Round & Initial Budget FE	No	Yes	Yes
Last 20 Rounds	Yes	No	Yes
Observations	1180	1475	1180

Notes: Linear regression with subject random effects and standard errors clustered at the attacker level. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1.
This table replicates Column (3) of Table 1 both without using budget fixed effects and using the data from all 25 rounds. The results do not change substantively.

C Experiment Screenshots

Below, we include screenshots of the Naive treatment of the experiment. The sophisticated treatment included only minor differences in the explanations and displays.

Introduction

PLEASE READ CAREFULLY AND DO NOT PRESS NEXT UNTIL INSTRUCTED TO DO SO.

Thank you for participating in this study. This study is about decision-making. It should take about 2 hours, and you will be paid based on your earnings from the experiment. The money you earn will be paid either in cash at the end of the study or electronically within a few days of the end of the study.

Please do not use any electronic devices or talk with other participants during this study.

There will be no deception in this study. Every decision you make will be carried out exactly as it is described in the instructions. Anything else would violate the human ethics protocol under which we run the study (UQ Human Research Ethics Approval 2023/HE000101).

In the study you will make decisions that will affect the amount of money you earn. The study will consist of games that you will play with other randomly selected players. The players that you are paired with in a round are selected independently of whom you play with in any other match.

Please pay close attention to the instructions on the next page and the examples on the page after that. After you read the examples, there will be a short quiz on the instructions. You will receive $0.5 (i.e., 50 cents) for each question you answer correctly.

If you have questions at any point, please raise your hand and we will answer your questions privately.

Next

Figure C.1: Introduction

Instructions

PLEASE READ CAREFULLY AND DO NOT PRESS NEXT UNTIL INSTRUCTED TO DO SO.

In this study, you will play 25 rounds of a game with monetary payoffs. In each round, there will be two other players; they will be selected independently of who you play with in any other round.

Before playing the game, you will be asked to answer a quiz with multiple choice questions about your understanding of the game. Each correct answer in the quiz is worth 50 cents. Your payment for participating in this study will be based on the outcomes of the game. At the end of the study, a round of play will be chosen randomly, and your payoffs from that round will be the ones that determine your payment. In addition to the payoff from the randomly chosen round and the payoff from the quiz, you will receive a baseline payment of $25.

In each round, each player will be assigned a role, which determines the player's available choices and payoffs. One player will be the ransomware "Attacker" and the other two players will be the potential "Targets" of a ransomware attack. Each Target owns valuable data that might be stolen by the Attacker. One of the two Targets becomes the Victim of a successful attack and the other does not. The Attacker may give the stolen data back to the Victim after payment of a ransom. Without insurance, the Victim may or may not have a sufficient budget to cover the ransom requested by the Attacker. Before knowing if they are attacked, Targets may be able to purchase insurance that will supplement their available budget in case of an attack.

In each round, each Target may have the opportunity to buy insurance, while the Attacker will choose what ransom (price) to demand to return the data. In each round, each Target has an initial budget that they control. The initial budget may take any value between $1 and $24; each value is equally likely. The budget can be used to purchase insurance and to pay the ransom.

Each round has two stages:

Target Stage 1: Insurance Choice With a roughly 50-50 chance, insurance will be available to the Targets. If their initial budget is at least $18, Targets can buy full insurance. If they buy insurance and they are attacked, then their final budget is $36. If they buy insurance and they are not attacked, then their final budget is twice their initial budget minus $36. If they don't buy insurance, then their final budget is the same as their initial budget, regardless of whether they are attacked.

Attacker Stage 1: Ransom Choice The Attacker chooses a ransom to return the stolen data. The ransom must be between $6 and $36. Before choosing the ransom, the Attacker is told whether full insurance is available to the Targets. The Attacker does not know the Targets' budgets or whether they purchased insurance.

Figure C.2: Instructions

Targets Stage 2: Purchase Choice One of the two Targets becomes the Victim of a successful attack, and the other does not. If the final budget of the Victim of an attack is at least as high as the ransom demanded by the Attacker, then the Victim chooses whether to pay the ransom to have the data returned. Otherwise, the Victim cannot pay the ransom. The Target who has not been attacked has no choice to make.

Target Payoffs: Targets receive their final budget, minus the ransom if they pay it. They also receive a value for the use of the data of $36 if either they are not the Victim (their data is not stolen), or if they paid the ransom and the data was returned to them.

Attacker Payoffs: If the Victim pays the ransom, then the Attacker receives the ransom. Otherwise, the Attacker receives the value of owning the stolen data, which is $6.

Below is a demonstration of the payoffs achievable by the players in a game. By moving the two sliders below you can select different levels of the Target's initial budget and the ransom chosen by the Attacker. The table below the sliders gives the players' payoffs. The first row gives the Attacker's payoff, the second row gives the Victim's payoff and the third row gives the payoff of the Target that was not attacked. The columns specify the Target's choices of whether to buy the insurance and to pay the ransom. For example, the cell in the second row and the third column contains the Victim's payoff if the Victim bought insurance, but did not pay the ransom. Note that a cell's payoff could be "Not Possible". This occurs if the initial budget and the ransom are incompatible with the players' choices associated with the cell.

1		24
	Example Budget: **21**	
6		36
	Example Ransom Demand: **17**	

	Did not buy insurance, did not pay the ransom	**Did not buy insurance, paid the ransom**	**Bought insurance, did not pay the ransom**	**Bought insurance, paid the ransom**
Attacker	$6	$17	$6	$17
Victim	$21	$40	$36	$55
Target, not attacked	$57	Not Possible	$42	Not Possible

The next page contains a few examples to help you understand the game better. Read through the examples at your own pace, then continue to complete the quiz.

Figure C.3: Instructions (continued)

Examples

Example 1: Allan is assigned the role of the Attacker; Betty and Carol are assigned the roles of Targets. Betty draws an initial budget of $10, and Carol draws an initial budget of $20.

In stage 1, Full insurance is available. Betty cannot buy insurance as her budget is below $18. Carol chooses to buy insurance. Allan knows that insurance is available and demands a ransom of $26 to return the stolen data.

In stage 2, Betty is not attacked. Carol is the Victim of an attack. Carol decides to pay the ransom.

The payoffs are as follows: Betty's final budget is $10 because she could not buy insurance. Since she was not attacked, her payoff is $10 + $36. Carol's final budget is $36 because she bought insurance and she was the Victim of an attack. Since she pays the ransom, her payoff is $36 - $26 + $36. Allan's received the ransom payment and thus his payoff is $26.

Example 2: Danielle is assigned the role of the Attacker. Ephraim and Francisco are assigned the role of Targets. Ephraim draws an initial budget of $16, and Francisco draws an initial budget of $12.

In stage 1, Ephraim and Francisco find out that there is no option to buy insurance in this match. Their final budgets are the same as their initial budgets and there is no insurance choice for them to make. Danielle knows that insurance is not available and chooses a ransom demand of $14 to return the stolen data.

In stage 2, Ephraim is not attacked. Francisco is the Victim of an attack. Francisco cannot pay the ransom and have the stolen data returned because the ransom demand chosen by Danielle is higher than his budget.

The payoffs are as follows: Because he was the Victim of an attack but he could not pay the ransom, Francisco receives only his budget of $12. Because he was not attacked, Ephraim receives his budget plus the value of the data $16 + $36. Because Danielle's ransom demand was (automatically) rejected, she receives the value of owning the stolen data, which is $6.

Quiz: When you are ready, you may click next and begin the quiz.

Next

Figure C.4: Examples

Quiz

Please answer the following questions. You will earn 50 cents for each question you answer correctly.

Suppose: (i) Insurance is NOT available; (ii) The Attacker demands a ransom of \$26 to return the stolen data; (iii) The Victim has an initial budget of \$18. What are the payoffs of the Attacker and of the Victim?

- ( ) The Attacker's payoff is \$26; The Victim's payoff is \$10.
- ( ) The Attacker's payoff is \$6; The Victim's payoff is \$18.
- ( ) The Attacker's payoff is \$26; The Victim's payoff is \$18.
- (•) Cannot tell. It depends on whether the Victim pays the ransom.

Suppose: (i) Insurance is NOT available; (ii) The Attacker demands a ransom of \$16 to return the stolen data; (iii) The Victim has an initial budget of \$18; (iv) The Victim pays the ransom. What are the payoffs of the Attacker and the Victim?

- ( ) The Attacker's payoff is \$6; The Victim's payoff is \$18.
- ( ) The Attacker's payoff is \$16; The Victim's payoff is \$38.
- ( ) The Attacker's payoff is \$16; The Victim's payoff is \$2.
- (•) The Attacker's payoff is \$6; The Victim's payoff is \$20.

Suppose: (i) Insurance is available; (ii) The Attacker demands a ransom of \$30 to return the stolen data; (iii) The Victim bought insurance. How do the payoffs of the Attacker depend on whether the Victim pays the ransom?

- ( ) If the Victim pays the ransom, the Attacker's payoff is \$6; If the Victim doesn't pay the ransom, the Attacker's payoff is \$30.
- ( ) If the Victim pays the ransom, the Attacker's payoff is \$6; If the Victim doesn't pay the ransom, the Attacker's payoff is \$20.
- ( ) If the Victim pays the ransom, the Attacker's payoff is \$20; If the Victim doesn't pay the ransom, the Attacker's payoff is \$6.
- (•) If the Victim pays the ransom, the Attacker's payoff is \$30; If the Victim doesn't pay the ransom, the Attacker's payoff is \$6.

Suppose: (i) Insurance is available; (ii) The Attacker demands a ransom of \$32 to return the stolen data; (iii) The Victim bought insurance. How does the payoff of the Victim depend on whether or not the Victim pays the ransom?

- ( ) If the Victim pays the ransom, the Victim receives \$4; If the Victim doesn't pay the ransom, the Victim receives \$0.

Figure C.5: Quiz (continued)

- ( ) If the Victim pays the ransom, the Victim receives \$6; If the Victim doesn't pay the ransom, the Victim receives \$6.
- ( ) If the Victim pays the ransom, the Victim receives \$40; If the Victim doesn't pay the ransom, the Victim receives \$36.
- (•) Cannot tell. It depends on the Victim's initial budget.

Suppose: (i) Insurance is NOT available; (ii) The Attacker demands a ransom of \$16 to return the stolen data. What is the chance that the Victim has a final budget that is at least as high as the demanded ransom?

- ( ) The chance that the final budget is at least as high as the ransom is 2 out of 24.
- ( ) The chance that the final budget is at least as high as the ransom is 6 out of 24.
- ( ) The chance that the final budget is at least as high as the ransom is 9 out of 24.
- (•) The chance that the final budget is at least as high as the ransom is 14 out of 24.

Suppose: (i) Insurance is available; (ii) The Attacker has demanded a ransom of \$30; (iii) The Victim buys insurance whenever they can. What is the chance that the Victim has a final budget that is at least as high as the demanded ransom?

- ( ) The chance that the final budget is at least as high as the ransom is 0 out of 24.
- ( ) The chance that the final budget is at least as high as the ransom is 3 out of 24.
- ( ) The chance that the final budget is at least as high as the ransom is 5 out of 24.
- (•) The chance that the final budget is at least as high as the ransom is 7 out of 24.

Next

Figure C.6: Quiz

Quiz Answers

The answers for the quiz are given below. Please review the answers and note any mistakes you have made.

Question 1: Suppose: (i) Insurance is NOT available; (ii) The Attacker demands a ransom of $26 to return the stolen data; (iii) The Victim has an initial budget of $18. What are the payoffs of the Attacker and of the Victim?
Correct Answer: The Attacker's payoff is $6; The Victim's payoff is $18.
Your Answer: Cannot tell. It depends on whether the Victim pays the ransom.

Question 2: Suppose: (i) Insurance is NOT available; (ii) The Attacker demands a ransom of $16 to return the stolen data; (iii) The Victim has an initial budget of $18; (iv) The Victim pays the ransom. What are the payoffs of the Attacker and the Victim?
Correct Answer: The Attacker's payoff is $16; The Victim's payoff is $38.
Your Answer: The Attacker's payoff is $6; The Victim's payoff is $20.

Question 3: Suppose: (i) Insurance is available; (ii) The Attacker demands a ransom of $30 to return the stolen data; (iii) The Victim bought insurance. How do the payoffs of the Attacker depend on whether the Victim pays the ransom?
Correct Answer: If the Victim pays the ransom, the Attacker's payoff is $30; If the Victim doesn't pay the ransom, the Attacker's payoff is $6.
Your Answer: If the Victim pays the ransom, the Attacker's payoff is $30; If the Victim doesn't pay the ransom, the Attacker's payoff is $6.

Question 4: Suppose: (i) Insurance is available; (ii) The Attacker demands a ransom of $32 to return the stolen data; (iii) The Victim bought insurance. How does the payoff of the Victim depend on whether or not the Victim pays the ransom?
Correct Answer: If the Victim pays the ransom, the Victim receives $40; If the Victim doesn't pay the ransom, the Victim receives $36.
Your Answer: Cannot tell. It depends on the Victim's initial budget.

Question 5: Suppose: (i) Insurance is NOT available; (ii) The Attacker demands a ransom of $16 to return the stolen data. What is the chance that the Victim has a final budget that is at least as high as the demanded ransom?
Correct Answer: The chance that the final budget is at least as high as the ransom is 9 out of 24.
Your Answer: The chance that the final budget is at least as high as the ransom is 14 out of 24.

Question 6: Suppose: (i) Insurance is available; (ii) The Attacker has demanded a ransom of $30; (iii) The Victim buys insurance whenever they can. What is the chance that the Victim has a final budget that is at least as high as the demanded ransom?

Figure C.7: Quiz Answers

Correct Answer: The chance that the final budget is at least as high as the ransom is 7 out of 24.
Your Answer: The chance that the final budget is at least as high as the ransom is 7 out of 24.

You earned $1.0 from your correct answers. Please review any questions you answered incorrectly. When you are ready to begin the first round, click the next button.

Next

Figure C.8: Quiz Answers

Insurance Choice: Round 1

In this round, you have been randomly assigned the role of Target. Your budget is $21.

Insurance is available this round for a price of $15. If you choose to purchase, your budget will be $36 if you are the victim of an attack and $6 if you are not the victim of an attack.

Do you want to buy insurance?

- (•) Yes
- ( ) No

Next

Below, you will find an interactive feature to help you make decisions. Targets have access to this feature when they are purchasing insurance, and Attackers have access when they are choosing the ransom to demand. Move both of the sliders to activate the feature to see what the payoffs to each player are of various choices. Remember that Attackers will not know the budget or insurance decision when choosing their ransom demand.

1		24
	Example Budget: **8**	
6		36
	Example Ransom Demand: **20**	

	Did not buy insurance, did not pay the ransom	**Did not buy insurance, paid the ransom**	**Bought insurance, did not pay the ransom**	**Bought insurance, paid the ransom**
Attacker	$6	Not Possible	Not Possible	Not Possible
Victim	$8	Not Possible	Not Possible	Not Possible
Target, not attacked	$44	Not Possible	Not Possible	Not Possible

Figure C.9: Insurance Choice (sufficient budget)

Insurance Choice: Round 5

In this round, you have been randomly assigned the role of Target. Your budget is $16.

Insurance is available this round, but its price is $20 so your budget is not large enough to purchase it. Please click next.

Next

Below, you will find an interactive feature to help you make decisions. Targets have access to this feature when they are purchasing insurance, and Attackers have access when they are choosing the ransom to demand. Move both of the sliders to activate the feature to see what the payoffs to each player are of various choices. Remember that Attackers will not know the budget or insurance decision when choosing their ransom demand.

1		24
	Example Budget: **16**	
6		36
	Example Ransom Demand: **10**	

	Did not buy insurance, did not pay the ransom	**Did not buy insurance, paid the ransom**	**Bought insurance, did not pay the ransom**	**Bought insurance, paid the ransom**
Attacker	$6	$10	Not Possible	Not Possible
Victim	$16	$42	Not Possible	Not Possible
Target, not attacked	$52	Not Possible	Not Possible	Not Possible

Figure C.10: Insurance Choice (insufficient budget)

Insurance Choice: Round 2

In this round, you have been randomly assigned the role of Target. Your budget is $19.

Insurance is not available in this round. Please click next.

Next

Below, you will find an interactive feature to help you make decisions. Targets have access to this feature when they are purchasing insurance, and Attackers have access when they are choosing the ransom to demand. Move both of the sliders to activate the feature to see what the payoffs to each player are of various choices. Remember that Attackers will not know the budget or insurance decision when choosing their ransom demand.

1		24
6		36

MOVE BOTH SLIDERS TWICE TO ACTIVATE PAYOFF TABLE

	Did not buy insurance, did not pay the ransom	**Did not buy insurance, paid the ransom**	**Bought insurance, did not pay the ransom**	**Bought insurance, paid the ransom**
Attacker				
Victim				
Target, not attacked				

Figure C.11: Insurance Choice (no insurance available)

Ransom Choice: Round 1

In this round, you have been randomly assigned the role of Attacker.

Recall that the Target's initial budget is random, and is equally likely to take all values between $1 and $24.

In this round, **insurance was available**, but a Target could only purchase it if their initial budget was at least 18. Furthermore, you cannot know whether Targets chose to purchase insurance.

If the Target purchased insurance, the final budget will be 36. Otherwise, the final budget is the same as the initial budget.

Move the slider below to select the ransom you will demand the Victim. When you are happy with your choice, click next.

6 36

Your Ransom Demand: **27**

Next

Below, you will find an interactive feature to help you make decisions. Targets have access to this feature when they are purchasing insurance, and Attackers have access when they are choosing their ransom. Move both of the sliders to activate the feature to see what the payoffs to each player are of various choices. Remember that Attackers will not know the budget or insurance decision when choosing their ransom demand.

1 24

Example Budget: **24**

6 36

Example Ransom Demand: **36**

	Did not buy insurance, did not pay the ransom	**Did not buy insurance, paid the ransom**	**Bought insurance, did not pay the ransom**	**Bought insurance, paid the ransom**
Attacker	$6	Not Possible	$6	$36

Figure C.12: Ransom choice (insurance round)

Victim	$24	Not Possible	$36	$36
Target, not attacked	$60	Not Possible	$48	Not Possible

Figure C.13: Ransom choice (insurance round, continued)

Ransom Choice: Round 2

In this round, you have been randomly assigned the role of Attacker.

Recall that the Target's initial budget is random, and is equally likely to take all values between $1 and $24.

In this round, **insurance was not available**, so the final budget is the same as the initial budget.

Move the slider below to select the ransom you will demand the Victim. When you are happy with your choice, click next.

6		36

Next

Below, you will find an interactive feature to help you make decisions. Targets have access to this feature when they are purchasing insurance, and Attackers have access when they are choosing their ransom. Move both of the sliders to activate the feature to see what the payoffs to each player are of various choices. Remember that Attackers will not know the budget or insurance decision when choosing their ransom demand.

1		24
6		36

MOVE BOTH SLIDERS TWICE TO ACTIVATE PAYOFF TABLE

	Did not buy insurance, did not pay the ransom	**Did not buy insurance, paid the ransom**	**Bought insurance, did not pay the ransom**	**Bought insurance, paid the ransom**
Attacker				
Victim				

Figure C.14: Ransom choice (no insurance available)

Target, not attacked				

Figure C.15: Ransom choice (no insurance available, continued)

Ransom Payment Choice: Round 1

In this round, you are the Victim.

Your randomly drawn budget is $21 and you bought insurance, so your total budget is $36. The ransom that the Attacker demanded is $27.

If you pay the ransom, you will receive $36 - $27 + $36 = $45 and the Attacker will receive $27.

If you do not pay the ransom, you will receive $36 and the Attacker will receive $6.

Will you pay the ransom?

- ○ Yes
- ○ No

Next

Figure C.16: Ransom payment choice

Ransom Payment Choice: Round 4

In this round, you are the Victim.

Your randomly drawn budget is $9 and you did not buy insurance, so your total budget is $9. The ransom that the Attacker demanded is $22, which is more than your budget. So, you do not have the option to pay the ransom and have your data returned. You will receive $9 and the Attacker will receive $6.

Next

Figure C.17: Ransom payment choice (insufficient budget)

Results: Round 1

In this round, you were the Victim.

Your randomly drawn initial budget was $21, and you bought insurance for a price of $15. So, your final budget was $36.

You paid the ransom demand of $27, so you paid the Attacker that amount and received the value of using the data of $36. So if this round is selected to be the one that counts, your final payoff will be $36 - $27 + $36 = $45.

Next

Figure C.18: Victim Results

Results: Round 1

In this round, you were the Target who was not attacked.

Your randomly drawn initial budget was $19, and you bought insurance for a price of $17. So, your final budget was $2.

Because you were not attacked, you automatically receive the value of using the data of $36. So if this round is selected to be the one that counts, your final payoff will be $2 + $36 = $38.

Next

Figure C.19: Target Results

Results: Round 1

In this round, you were the Attacker.

The Victim had a randomly drawn initial budget of $21 and chose to buy insurance. So, their final budget was $36.

The Victim paid your ransom demand of $27. So if this round is selected to be the one that counts, your final payoff will be $27.

Next

Figure C.20: Attacker Results

Final Results

The computer has randomized over all rounds, and the round that will determine payoffs for all participants is Round 22. In that round, you received a payoff of $56. Combined with your quiz payment of $1.0 and the base payment of $25, your total payoff from the study will be $82.00.

Please click next and complete a short survey while we prepare your payments. After the survey is completed, stay in your seat. We will deliver your payments to you.

Next

Figure C.21: Final Results

Survey

Please answer the following questions.

What is your age?

What gender do you identify with the most?

- Female
- Male
- Other/Prefer Not to Say

Is English your first language?

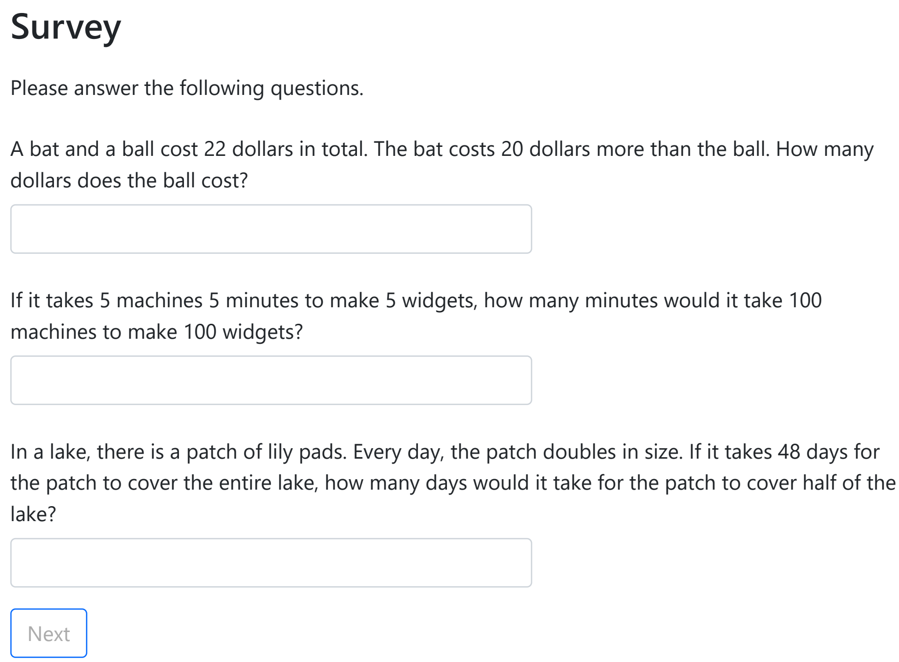
- Yes
- No

Are you completing or have you completed an economics degree?

- Yes
- No

Next

Figure C.22: Demographic Survey



Survey

Please answer the following questions.

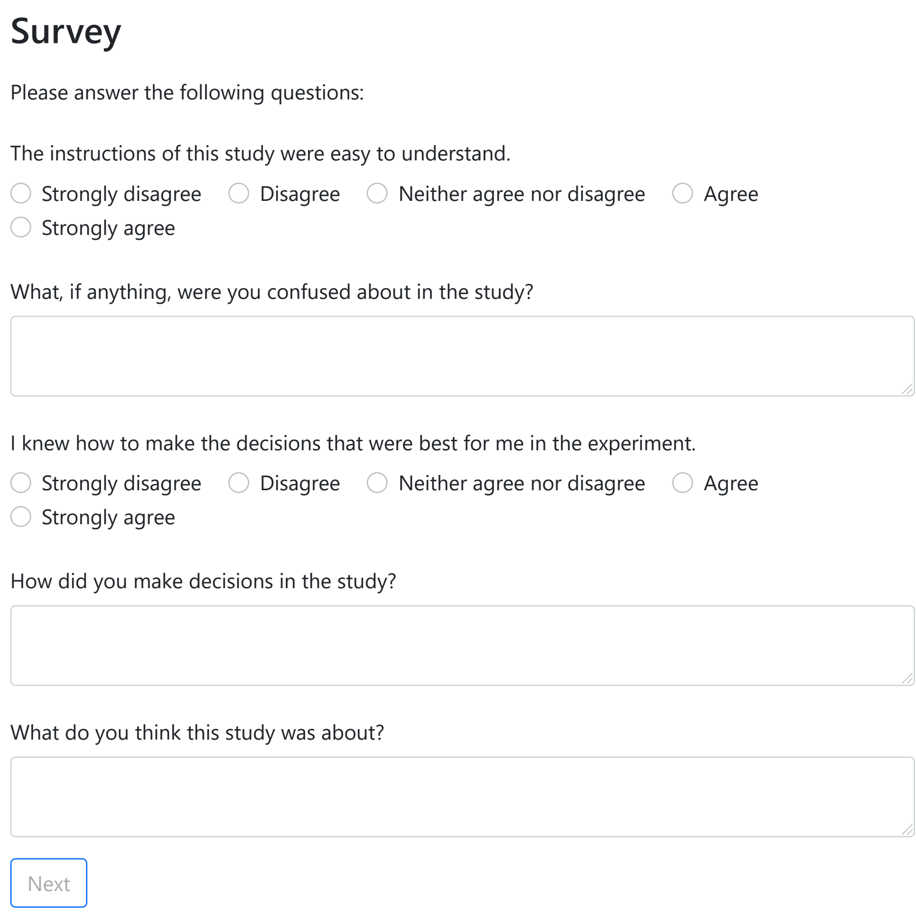
A bat and a ball cost 22 dollars in total. The bat costs 20 dollars more than the ball. How many dollars does the ball cost?

If it takes 5 machines 5 minutes to make 5 widgets, how many minutes would it take 100 machines to make 100 widgets?

In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how many days would it take for the patch to cover half of the lake?

Next

Figure C.23: CRT


Survey

Please answer the following questions:

The instructions of this study were easy to understand.

◯ Strongly disagree ◯ Disagree ◯ Neither agree nor disagree ◯ Agree
◯ Strongly agree

What, if anything, were you confused about in the study?

I knew how to make the decisions that were best for me in the experiment.

◯ Strongly disagree ◯ Disagree ◯ Neither agree nor disagree ◯ Agree
◯ Strongly agree

How did you make decisions in the study?

What do you think this study was about?

Next

Figure C.24: Experiment Feedback