

DISCUSSION PAPER SERIES

DP19047
(v. 2)

ARTIFICIAL BUGS FOR BUG BOUNTY

Hans Gersbach, Fikri Pitsuwan and Pio Blieske

**INDUSTRIAL ORGANIZATION AND
ORGANIZATIONAL ECONOMICS**

CEPR

ARTIFICIAL BUGS FOR BUG BOUNTY

Hans Gersbach, Fikri Pitsuwan and Pio Blieske

Discussion Paper DP19047
First Published 06 May 2024
This Revision 07 May 2024

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Industrial Organization
- Organizational Economics

Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Hans Gersbach, Fikri Pitsuwan and Pio Blieske

ARTIFICIAL BUGS FOR BUG BOUNTY

Abstract

Bug bounty programs, where external agents are invited to search and report vulnerabilities (bugs) in exchange for rewards (bounty), have become a major tool for companies to improve their systems. We suggest augmenting such programs by inserting artificial bugs to increase the incentives to search for real (organic) bugs. Using a model of crowdsearch, we identify the efficiency gains by artificial bugs, and we show that for this, it is sufficient to insert only one artificial bug. Artificial bugs are particularly beneficial, for instance, if the designer places high valuations on finding organic bugs or if the budget for bounty is not sufficiently high. We discuss how to implement artificial bugs and outline their further benefits.

JEL Classification: C72, D82, M52

Keywords: Crowdsearch, Bug Bounty, Artificial Bug, Cybersecurity

Hans Gersbach - hgersbach@ethz.ch
KOF Swiss Economic Institute, ETH Zurich and CEPR

Fikri Pitsuwan - fpitsuwan@ethz.ch
KOF Swiss Economic Institute, ETH Zurich

Pio Blieske - pblieske@ethz.ch
KOF Swiss Economic Institute, ETH Zurich

Acknowledgements

This research was supported by the Zurich Information Security and Privacy Center (ZISC). We thank Rainer Böhme, Kari Kostiaainen, Ueli Maurer, Martina Schwab, Marcel Zumbühl, representatives of Zurich Insurance and Swiss Post, as well as seminar participants at the Workshop on Real-life Impacts of Security Vulnerabilities (2024) at ETH Zurich for valuable comments and discussions. All errors are our own.

Artificial Bugs for Bug Bounty*

Hans Gersbach

KOF Swiss Economic Institute,
ETH Zurich, and CEPR

Leonhardstrasse 21
8092 Zurich, Switzerland

hgersbach@ethz.ch

Fikri Pitsuwan

KOF Swiss Economic Institute,
ETH Zurich

Leonhardstrasse 21
8092 Zurich, Switzerland

fpitsuwan@ethz.ch

Pio Blieske

KOF Swiss Economic Institute,
ETH Zurich

Leonhardstrasse 21
8092 Zurich, Switzerland

pblieske@ethz.ch

Last updated: May 6, 2024

Abstract

Bug bounty programs, where external agents are invited to search and report vulnerabilities (bugs) in exchange for rewards (bounty), have become a major tool for companies to improve their systems. We suggest augmenting such programs by inserting artificial bugs to increase the incentives to search for real (organic) bugs. Using a model of crowdsearch, we identify the efficiency gains by artificial bugs, and we show that for this, it is sufficient to insert only one artificial bug. Artificial bugs are particularly beneficial, for instance, if the designer places high valuations on finding organic bugs or if the budget for bounty is not sufficiently high. We discuss how to implement artificial bugs and outline their further benefits.

Keywords: Crowdsearch, Bug Bounty, Artificial Bug, Cybersecurity

JEL Classification: C72, D82, M52

*This research was supported by the Zurich Information Security and Privacy Center (ZISC). We thank Rainer Böhme, Kari Kostianen, Ueli Maurer, Martina Schwab, Marcel Zumbühl, representatives of Zurich Insurance and Swiss Post, as well as seminar participants at the Workshop on Real-life Impacts of Security Vulnerabilities (2024) at ETH Zurich for valuable comments and discussions. All errors are our own.

1 Introduction

How should we design a public intrusion test where external *security researchers* are allowed to probe the software and report any vulnerabilities (bug) in exchange for rewards (bounty)? This type of program, often called *bug bounty* or *crowdsourced security*, has become a major tool for detecting vulnerabilities in software used by governments, tech companies, and blockchains.¹ Bug bounty is used by companies even if they have strong internal security teams (e.g. Google and Meta).² It is also critical for blockchain infrastructure providers since such projects do not have dedicated security teams testing software upgrades. Once the software is deployed, there is no turning back and no legal mechanism defending against system exploitation, at least until the next hard fork—a major change in the blockchain protocol (Breidenbach et al., 2018; Böhme et al., 2020).

There are comprehensive accounts on the rules of engagement of bug bounty programs (Laszka et al., 2018), on the effectiveness and best practices of such programs (Walshe and Simpson, 2020; Malladi and Subramanian, 2020), and on the incentives of researchers to participate in bug bounty programs (Maillart et al., 2017).³ In this paper, we suggest augmenting such bug bounty programs by inserting artificial bugs to increase the incentives to search for real (organic) bugs.

We first extend the simple model of *crowdsearch*, wherein researchers with different abilities decide on whether or not to exert costly effort to search for unknown objects that are valuable to the organization. The model builds on the literature on binary contests (Dubey (2013); Ghosh and Kleinberg (2016); Sarne and Lepioshkin (2017); Levy et al. (2017); Gersbach et al. (2023a)) and contests with entry (Levin and Smith (1994); Moreno and Wooders (2011); Liu and Lu (2019)). We then propose that inserting artificial bugs and appropriately setting their parameters may be beneficial to the organization. In the

¹Their success in the past few years has led the Swiss authorities to systematically adopt bug bounty programs as a main measure in government cybersecurity. In a recent press release, the [Federal Department of Finance of Switzerland \(2022\)](#) states that “standardised security tests are no longer sufficient to uncover hidden loopholes. Therefore, in the future, it is intended that ethical hackers will search through the Federal Administration’s productive IT systems and applications for vulnerabilities as part of so-called bug bounty programmes.”

²See [Malladi and Subramanian \(2020\)](#).

³For extensive literature on bug bounty programs, see [Böhme \(2006\)](#); [Zrahia et al. \(2022\)](#); [Akgul et al. \(2023\)](#).

language of contests, this is akin to inserting artificial milestones that have no inherent value to the contest designer but may incentivize increased efforts, which indirectly benefits the designer.

Our main contribution is the proposal to insert artificial bugs in bug bounty programs. In particular, we examine how inserting artificial bugs can be useful to the organization by lowering its financial commitment. We show that the ability to adjust the complexities of artificial bugs allows the organization to motivate more participants to search for organic bugs. We establish that it is sufficient to insert one artificial bug to reap all possible efficiency gains. Moreover, we identify that an artificial bug is particularly beneficial if the designer has high valuations for finding organic bugs, if organic bugs are likely to exist, or if the designer’s budget is low. Finally, we outline different engineering approaches to implement artificial bugs in practice and identify some other benefits of artificial bugs.

The paper is organized as follows: [Section 2](#) sets up the model. [Section 3](#) analyzes a private bug bounty program in which a finite number of agents search for bugs. [Section 4](#) considers a public program with a large crowd. Numerical examples are shown in [Section 5](#). [Section 6](#) outlines three approaches to implement artificial bugs and [Section 7](#) discusses other benefits of artificial bugs. [Section 8](#) concludes. Proofs are relegated to the appendix.

2 The Model

We provide a foundation for artificial bugs in the simplest possible model of crowdsearch. There are potentially L organic bugs in a system, indexed by l , and each exists with probability $\boldsymbol{\mu} = (\mu^1, \dots, \mu^L) \in (0, 1]^L$. We assume that the occurrence of bugs l and l' ($l \neq l'$) are stochastically independent, and neither the organization nor potential participants in a bug bounty program have any knowledge of whether these bugs exist and where to find them. An organization (henceforth, *designer*) values finding the bugs at $\boldsymbol{w} = (w^1, \dots, w^L) \in [0, \infty)^L$. The values of w can be interpreted as utility or monetary values. The designer invites a set of risk-neutral agents N to search for them in exchange

for rewards $\mathbf{v} = (v^1, \dots, v^L) \in [0, \infty)^L$. If agent i decides to search, s/he finds bug l with probability $q^l \in (0, 1]$, which represents its complexity, and receives a prize v^l uniformly randomly rewarded to one of the agents who found it. The events finding bug l and finding bug l' ($l \neq l'$) are stochastically independent. Searching, however, entails a fixed cost c_i to agent $i \in N$, which is private information and drawn from a distribution F with support on $[\underline{c}, \bar{c}]$, where $-\infty \leq \underline{c} < \bar{c} \leq \infty$ and $\bar{c} > 0$.⁴ We impose that F is continuous, has full support, has a finite density function f , and that F/f is non-decreasing.⁵

The designer has a budget of $\bar{v} > 0$ and can set prizes \mathbf{v} to incentivize the agents. In addition to setting the prizes for the organic bugs, the designer can insert K artificial bugs, indexed by k , into the system, offering rewards $\mathbf{v}_a = (v_a^1, \dots, v_a^K) \in [0, \infty)^K$ and setting their complexities $\mathbf{q}_a = (q_a^1, \dots, q_a^K) \in [0, 1]^K$, where complexity q_a^k is the probability that an agent finds the artificial bug. The complexities and reward system are communicated to the participants. Note that while artificial bugs have no inherent value for the designer, they may nonetheless be useful for the designer in motivating agents to participate.

The game has two stages. In the first stage, the designer sets \mathbf{v} , \mathbf{v}_a , and \mathbf{q}_a . In the second stage, agents decide whether or not to exert costly effort to search for bugs and are rewarded accordingly if they find any. The designer receives their value for the bugs found and pays out the prizes accordingly. We analyze the game by first solving for the (Bayes-Nash) equilibrium behavior of the agents in the second stage, then for the designer's optimal prizes and complexities of the artificial bugs.

We analyze the model in two environments. First, [Section 3](#) considers a *private program*, where the set of agents is finite, $N = \{1, \dots, n\}$. This represents a private bug bounty program, where the designer invites n agents to search for bugs. Second, [Section 4](#) considers a *public program*, where we perform an asymptotic analysis of the game in which the number of agents tends to infinity, $n \rightarrow \infty$. This captures a public bug bounty program, where the public can freely choose to participate.

⁴We allow the possibility that the lower support \underline{c} is negative to capture agents with intrinsic gain from search.

⁵This holds for many common distributions such as the exponential distribution, the beta distribution with shape parameter $\beta \geq 1$, and $F(c) = \left(\frac{c-\underline{c}}{\bar{c}-\underline{c}}\right)^\alpha$ on $[\underline{c}, \bar{c}]$ with $\alpha > 0$, which includes the uniform distribution for $\alpha = 1$.

2.1 Discussion of the Model

For the purpose of exposition and tractability, we impose some information assumptions and simplifications in the model.

First, we assume that there are potentially L organic bugs in the system, which may or may not exist. Each of these bugs is to be interpreted as a *type* of bug with different severity levels. Thus, our informational assumption that the designer has an understanding of the value w^l of finding each type of bug could be interpreted as the *expected* gain for the designer if such a type of bug is found. Moreover, given a system, experienced researchers have quite an accurate idea of the complexity, i.e., the probability of finding bugs.

Second, we assume that the search decision is binary since once a researcher decides to enter a particular search, s/he devotes full resources to it at least for some amount of time. Thus, the cost to participate in a specific bug bounty contest is, to a large extent, fixed, and this cost includes time spent reading and understanding the update, writing a program, designing an attack environment, and the opportunity cost of *not* participating in other bug bounty programs.

Third, we abstract away from the time dimension as it is not pertinent for our main goal. The model, however, could be interpreted with the timing aspect as follows. Suppose that when agent i searches, i finds bug l with probability $q^l \in (0, 1]$ at a random time t_i^l that is uniformly distributed over a time period $[0, T]$, where T is the duration of the bug bounty program. Agent i then receives the prize for bug l if s/he is the first agent to find it. Now, if these arrival times t_i^l 's are identically distributed and stochastically independent across agents and bugs, then the resulting outcome is the prize-sharing scheme presented in the main model, where the prize is given uniformly randomly among those that found the bug.⁶

Fourth, the assumption that all agents find a particular bug with the same probability is only for ease of exposition. The model can straightforwardly be extended to allow for heterogeneous bug-finding probabilities, with q_i^l denoting agent i 's probability of finding

⁶The model also extends to the case where search times are distributed differently similarly to [Gersbach et al. \(2023b\)](#).

bug l . This heterogeneity captures the idea that each security researcher has a particular skill set that may favor finding some bugs rather than others. Maillart et al. (2017)’s empirical investigation of bug bounty programs supports this assumption.

3 Private Program

This section considers a private program for bug bounty, where the designer invites n agents to search for bugs. First, we characterize the equilibrium behavior of the agents in the second stage. Then, we solve for the designer’s optimal choices and determine that it is sufficient to insert only one artificial bug. We end this section by identifying the condition under which inserting an artificial bug is useful.

3.1 Equilibrium Characterization

Given \mathbf{v} , \mathbf{v}_a , and \mathbf{q}_a set by the designer, the agents choose whether or not to search for the bugs. Since participants can only take binary choices, it is without loss of generality to consider only threshold strategies of the agents. A threshold strategy stipulates that an agent searches if and only if his/her private cost is below a certain threshold. We also focus on the symmetric equilibrium, where all agents use the same threshold c^* . Such an equilibrium always exists.

To characterize the equilibrium threshold, denote by $\Phi(\hat{c}; q)$ the probability that an agent wins the prize associated with a bug with complexity q , conditioning on it existing, when the other $n - 1$ agents are using a threshold strategy \hat{c} .⁷ Then, agent i ’s expected benefit of searching is

$$\Psi(\hat{c}; \mathbf{v}, \mathbf{v}_a, \mathbf{q}_a) \equiv \sum_l v^l \mu^l \Phi(\hat{c}; q^l) + \sum_k v_a^k \Phi(\hat{c}; q_a^k),$$

where the first term sums up the expected benefit from finding organic bugs and the second term from finding artificial bugs. Now, suppose that the other $n - 1$ agents use the equilibrium threshold c^* , then agent i will search if and only if their expected benefit

⁷This event occurs if the agent finds the bug and is chosen among those who also found it.

of searching exceeds their private cost: $c_i \leq \Psi(c^*)$. The following result follows.

Lemma 1. *If $\Psi(\underline{c}) \leq \underline{c}$, then $c^* = \underline{c}$. If $\Psi(\bar{c}) \geq \bar{c}$, then $c^* = \bar{c}$. Otherwise, the symmetric equilibrium threshold is $c^* = c^*(\mathbf{v}, \mathbf{v}_a, \mathbf{q}_a)$ is the unique solution to*

$$\hat{c} = \Psi(\hat{c}; \mathbf{v}, \mathbf{v}_a, \mathbf{q}_a). \quad (1)$$

The equilibrium threshold c^* is the unique fixed point of Ψ and, therefore, the comparative statics properties are derived from the properties of Ψ , which in turn follows the properties of Φ . The function $\Phi(\hat{c}; q)$ is strictly decreasing in \hat{c} and strictly increasing in q . The comparative statics results are as follows. Increasing v^l , μ^l , and v_a^k , for some l or k increases c^* , as agents are more incentivized to search if the prizes and the probabilities that bugs exist are high. Moreover, decreasing the complexities—increasing q^l 's—also increases c^* .

3.2 Optimal Prizes and Number of Artificial Bugs

We now consider the designer's problem of choosing the optimal prizes (for organic and artificial bugs) and complexities (for artificial bugs). Let the probability that a bug with complexity q is found be denoted by

$$P(\hat{c}; q) \equiv 1 - (1 - qF(\hat{c}))^n$$

when agents follow the threshold \hat{c} to decide whether to search or not. The designer chooses \mathbf{v} , \mathbf{v}_a , and \mathbf{q}_a to maximize their objective function consisting of two components. The first is $\sum_l w^l \mu^l P(c^*; q^l)$, which is the sum of the values gained from any organic bugs found. The second component captures the rewards paid out to the agents for any organic *and* artificial bugs found: $-\sum_l v^l \mu^l P(c^*; q^l) - \sum_k v_a^k P(c^*; q_a^k)$.

Therefore, the designer with a prize budget of \bar{v} solves

$$\begin{aligned} & \max_{\mathbf{v}, \mathbf{v}_a, \mathbf{q}_a} \sum_l (w^l - v^l) \mu^l P(c^*; q^l) - \sum_k v_a^k P(c^*; q_a^k) \\ & \text{subject to } \sum_l v^l + \sum_k v_a^k \leq \bar{v}, v^l \geq 0, v_a^k \geq 0, q_a^k \in [0, 1], \end{aligned} \quad (2)$$

where $c^* = c^*(\mathbf{v}, \mathbf{v}_a, \mathbf{q}_a)$ is the symmetric equilibrium threshold that uniquely solves Equation (1). Let \mathbf{v}^* , \mathbf{v}_a^* , and \mathbf{q}_a^* denote a solution to the designer's problem. The next result states that it is without loss of generality to insert only one artificial bug, if any.

Lemma 2. *If $(\mathbf{v}^*, \mathbf{v}_a^*, \mathbf{q}_a^*)$ is a solution, then there exists another solution $(\mathbf{v}^{**}, \mathbf{v}_a^{**}, \mathbf{q}_a^{**})$ such that $\mathbf{v}_a^{**} = (v_a^{1**}, 0, \dots, 0)$.*

Lemma 2 holds because it is inefficient to have multiple artificial bugs of different complexities, since artificial bugs are only useful indirectly, namely by incentivizing the agents to participate. For this purpose, it is always most efficient when the budget is spent on the prize corresponding to the artificial bug with the lowest complexity. In light of this, the designer's problem boils down to choosing \mathbf{v} , v_a , and q_a .

The key to solving the designer's problem is to note that $P(\hat{c}; q)$ can be decomposed in terms of each agent's probability of winning the corresponding prize.

Lemma 3. *For a bug with complexity q , we have $P(\hat{c}; q) = nF(\hat{c})\Phi(\hat{c}; q)$.*

Intuitively, if a bug is found, one of the agents has to win the prize. To win the prize, the agent has to have a cost below \hat{c} and to find the bug and be chosen among all agents that search and also find the bug. Ex-ante, the former happens with probability $F(\hat{c})$ and the latter with probability $\Phi(\hat{c}; q)$. From the perspective of the designer, all individuals are ex-ante identical; thus, multiplying by n yields the probability that the bug is found.⁸

With Lemma 3, the designer's objective function simplifies to $\sum_l w^l \mu^l P(\hat{c}; q^l) - nF(\hat{c})\hat{c}$. Therefore, the designer's problem is to induce an optimal equilibrium threshold, denoted

⁸The relation continues to hold if it is common knowledge that agents are heterogeneous in other aspects. For instance, agents may have different probabilities of finding a certain bug, or their costs may be drawn from different distributions. In such cases, the equilibrium thresholds differ across agents and we have the relation: $P((\hat{c}_1, \dots, \hat{c}_n); (q_1, \dots, q_n)) = \sum_i F_i(\hat{c}_i)\Phi_i(\hat{c}_i; (q_i, q_{-i}))$.

\hat{c}^* , that maximizes the objective function, subject to the equilibrium threshold being *achievable* given the budget \bar{v} .⁹ Define the set achievable equilibrium thresholds as

$$C(\bar{v}) \equiv \left\{ \hat{c} : \hat{c} = \Psi(\hat{c}; \mathbf{v}, v_a, q_a), \sum_l v^l + v_a \leq \bar{v}, v^l \geq 0, v_a \geq 0, q_a \in [0, 1] \right\}.$$

The set of achievable thresholds is an interval. Given that $\Phi(\hat{c}; q)$ is strictly increasing in q , the highest achievable threshold is when the entire budget goes to the artificial bug that is found with certainty.

Lemma 4. *Let $c_a(\bar{v})$ be the unique fixed point of $\bar{v}\Phi(\hat{c}; 1)$. Then, $c_a(\bar{v})$ increases in \bar{v} and $C(\bar{v}) = [0, c_a(\bar{v})]$.*

Taken together, the previous lemmata show that in effect, the designer's problem is to maximize

$$W(\hat{c}) \equiv \sum_l w^l \mu^l P(\hat{c}; q^l) - nF(\hat{c})\hat{c}, \quad (3)$$

subject to $\hat{c} \in [0, c_a(\bar{v})]$.

Our first main result characterizes the solution to the designer's problem. Define

$$\Omega(\hat{c}) \equiv \sum_l w^l \mu^l q^l (1 - q^l F(\hat{c}))^{n-1} - \frac{F(\hat{c})}{f(\hat{c})} \quad (4)$$

and let \tilde{c} be its unique fixed point, which is increasing in w^l and μ^l . We have

Theorem 1. *The optimal equilibrium threshold is $\hat{c}^* = \min\{\tilde{c}, c_a(\bar{v})\}$. A set of prizes \mathbf{v}^* , v_a^* and complexity q_a^* that solve $\hat{c}^* = \Psi(\hat{c}^*; \mathbf{v}, v_a, q_a)$ is optimal.*

3.3 Usefulness of Artificial Bugs

Next, we identify when inserting an artificial bug is useful. Note that it is feasible for the designer not to add any artificial bug into the system, as $v_a = 0$ or $q_a = 0$ are feasible. In other words, the set of achievable equilibrium thresholds without any artificial bug is

⁹Achievability means that the available budget is sufficient to pay all prizes that have been promised.

defined as

$$C_0(\bar{v}) \equiv \left\{ \hat{c} : \hat{c} = \Psi(\hat{c}; \mathbf{v}, 0, 0), \sum_l v^l \leq \bar{v}, v^l \geq 0 \right\} \subseteq C(\bar{v}).$$

Analogously to [Lemma 4](#), let $c^l(\bar{v})$ be the unique fixed point of $\bar{v}\mu^l\Phi(\hat{c}; q^l)$. Then, $C_0(\bar{v}) = [0, c_0(\bar{v})]$, where $c_0(\bar{v}) = \max_l c^l(\bar{v})$. This leads to the next main result.

Theorem 2. *Inserting an artificial bug is beneficial to the designer if and only if $\tilde{c} > c_0(\bar{v})$.*

The condition in [Theorem 2](#) holds if w^l and μ^l are high, or \bar{v} is not too high. These cases will be illustrated in the numerical examples in [Section 5](#). In other words, inserting an artificial bug is beneficial if the designer has high valuations for finding organic bugs, if organic bugs are likely to exist, or if the designer’s budget is low.

4 Public Program

In this section, we examine the public bug bounty program and, in particular, the asymptotic behavior of the game as $n \rightarrow \infty$. This is of special interest since public bug bounty programs make use of the knowledge of a large group of experts. Throughout the section, we denote the equilibrium threshold and the equilibrium success probability when there are n agents in the public program by $c_n = c^*(n)$ and $P_n(q) = P(c^*(n); q)$, respectively. We assume throughout the section that $\underline{c} > 0$. This is reasonable in most circumstances since even high-ability agents have to exert effort to find the bugs and communicate with the designer.¹⁰ For ease of exposition, we impose two further assumptions, which exclude pathological cases in which a bug bounty program can not generate any benefits. First, we assume that $\bar{v}\Psi(\underline{c}) \geq \underline{c}$, i.e., that the budget constraint is greater than the smallest possible cost for an agent to search. Otherwise, no agent will participate, and there is no point in designing a bug bounty program. Second, we assume that $\sum_l w^l \mu^l q^l \geq \underline{c}$ since otherwise, the designer can never compensate an agent for searching. Again, the budget constraint \bar{v} and the bugs described by \mathbf{w} and \mathbf{q} are given for the designer.

¹⁰For $\underline{c} = 0$, it holds that $P_n(q) \rightarrow 1$ for any prizes and complexities (see [Gersbach et al. \(2023a\)](#)). This case seems to be less plausible.

We proceed as follows: First, we obtain results for the equilibrium behavior of the agents in the second stage as $n \rightarrow \infty$. Then, given the equilibrium behavior of the agents, we solve for the designer's optimal choices. Next, we verify that this is equivalent to solving first for the designer's optimal choices (i.e., the full game) and then performing an asymptotic analysis on such choices. Lastly, we analyze the usefulness of inserting an artificial bug in the public program.

4.1 Asymptotic Behavior

Given \mathbf{v} , v_q , and q_a , the following lemma summarizes the asymptotic results for the second stage game:

Lemma 5. *Let κ^* be the unique fixed point of*

$$\Psi_\infty(\hat{\kappa}; \mathbf{v}, v_a, q_a) \equiv \sum_l v^l \mu^l \frac{1 - e^{-q^l \hat{\kappa}}}{\underline{c}} + v_a \frac{1 - e^{-q_a \hat{\kappa}}}{\underline{c}}.$$

Then, for $n \rightarrow \infty$

- (i) $c_n \rightarrow \underline{c}$,
- (ii) $nF(c_n) \rightarrow \kappa^*$,
- (iii) $P_n(q) \rightarrow P_\infty(q) = 1 - e^{-q\kappa^*}$.

Note that the asymptotic behavior is distribution-free insofar that it does not depend on the distribution of F , but only on \underline{c} . The reason is that only agents with the highest ability (lowest cost) participate in a public program. The quantity κ^* is important as it represents the *asymptotic participation*, i.e., the asymptotic expected number of participants in the crowdsearch.

4.2 Optimal Prizes

Denote by $W_n(\hat{c})$ the designer's objective function (3) for a particular value of n . With Lemma 5, we obtain the designer's objective function at the equilibrium threshold, c_n , as $n \rightarrow \infty$.

Corollary 1. *The designer's objective function at the equilibrium threshold c_n converges:*

$$W_n(c_n) \rightarrow W_\infty(\kappa^*) \equiv \sum_l w^l \mu^l \left(1 - e^{-q^l \kappa^*}\right) - \kappa^* \underline{c}, \quad \text{as } n \rightarrow \infty.$$

Thus, the designer's payoff only depends on the asymptotic participation. As in the private program, we define the set of achievable κ^* 's given a budget constraint \bar{v} as

$$K(\bar{v}) \equiv \left\{ \hat{\kappa} : \hat{\kappa} = \Psi_\infty(\hat{\kappa}; \mathbf{v}, v_a, q_a), \sum_l v^l + v_a \leq \bar{v}, v^l \geq 0, v_a \geq 0, q_a \in [0, 1] \right\}.$$

The following lemma characterizes this set.

Lemma 6. *Denote by $\kappa_a(\bar{v})$ the unique fixed point of $\bar{v} \frac{1-e^{-\hat{\kappa}}}{\underline{c}}$. Then, $\kappa_a(\bar{v})$ is increasing in \bar{v} and $K(\bar{v}) = [0, \kappa_a(\bar{v})]$.*

Taking the previous lemmata together, the designer's problem becomes

$$\max_{\hat{\kappa} \in [0, \kappa_a(\bar{v})]} \sum_l w^l \mu^l \left(1 - e^{-q^l \hat{\kappa}}\right) - \hat{\kappa} \underline{c}.$$

Analogously to the private program, define

$$\Omega_\infty(\hat{\kappa}) \equiv \sum_l w^l \mu^l q^l e^{-q^l \hat{\kappa}} - \underline{c}.$$

We observe that $\Omega_\infty(\hat{\kappa})$ is strictly decreasing and thus has a unique root since $\sum_l w^l \mu^l q^l \geq \underline{c}$. Let $\tilde{\kappa}$ be its unique root, which is decreasing in w^l and μ^l . We obtain the following characterization.

Theorem 3. *The optimal asymptotic participation is $\hat{\kappa}^* = \min\{\tilde{\kappa}, \kappa_a(\bar{v})\}$ and any set of prizes \mathbf{v}^∞ , v_a^∞ and complexity q_a^∞ that solve $\hat{\kappa}^* = \Psi_\infty(\hat{\kappa}^*; \mathbf{v}, v_a, q_a)$ is optimal.*

4.3 Verification

We now verify that the designer's optimal prizes and complexities in a game with finite agents indeed converge to the solution obtained in [Theorem 3](#). Denote the set of optimal

prizes and complexities in a private program with n agents by

$$M_n = \{(\mathbf{v}, v_a, q_a) : \hat{c}^* = \Psi(\hat{c}^*; \mathbf{v}, v_a, q_a)\}$$

and the set of optimal prizes and complexities in a public program by

$$M_\infty = \{(\mathbf{v}, v_a, q_a) : \hat{\kappa}^* = \Psi_\infty(\hat{\kappa}^*; \mathbf{v}, v_a, q_a)\}.$$

To show that the two sets asymptotically coincide, we introduce the Hausdorff distance which intuitively describes the maximum distance from points in any two sets.

Definition 1. Let $A, B \subseteq \mathbb{R}^k$ be two non-empty subsets. The Hausdorff distance induced by the Euclidean norm $\|\cdot\|_2$ is given by

$$d(A, B) = \max \left\{ \sup_{x \in A} \inf_{y \in B} \|x - y\|_2, \sup_{x \in B} \inf_{y \in A} \|x - y\|_2 \right\}.$$

Equipped with this, we state our next theorem.

Theorem 4. *It holds that $d(M_n, M_\infty) \rightarrow 0$ as $n \rightarrow \infty$.*

Hence, as n increases, the maximal distance for any solution of the private program comes arbitrarily close to a solution of the public program.

4.4 Usefulness of Artificial Bugs

Analogously to [Theorem 2](#), we identify when inserting an artificial bug is useful in a public program. Define $\kappa_0(\bar{v}) = \max_l \kappa_l(\bar{v})$, where $\kappa_l(\bar{v})$ the fixed point of $v^l \mu^l \frac{1-e^{-q^l \kappa}}{\underline{c}}$.

Theorem 5. *In the public program, inserting an artificial bug is beneficial to the designer if and only if $\tilde{\kappa} > \kappa_0(\bar{v})$. This is the case if and only if there exists $N \in \mathbb{N}$ such that for all $n \geq N$, it is beneficial to insert an artificial bug in the private program with n agents.*

Before ending this section, a remark on the cost distribution is in order. Throughout this section, we kept the distribution F of the cost of the agents fixed. However, the

distribution can plausibly vary, depending on the size of the crowd, F_n . Because with increasing n , less selection of the agents can occur until the whole population is open to participating, it is safe to assume that F_n converges to some distribution. By [Lemma 5](#), the limit case does not depend on the distribution, and all the results still hold, where the distribution F_n converges to some limit distribution F_∞ .

5 Numerical Examples

We next illustrate the results with numerical examples and visualize the effects of an artificial bug and the dependence on the budget constraint.

5.1 Private Program

We start with the private program. Consider the uniform distribution on $[0, 1]$ and assume that there is one type of organic bug that exists with probability $\mu = \frac{1}{2}$, has complexity $q = \frac{1}{2}$, and that finding the bug derives a utility of $w = 2$ for the designer. For simplicity, we assume $n = 2$ agents. Then, the optimal threshold given by the fixed point of [\(4\)](#) is $\tilde{c} = \frac{2}{9}$. The critical value depending on the budget constraint \bar{v} is $c_0(\bar{v}) = \left(\frac{4}{\bar{v}} + \frac{1}{4}\right)^{-1}$. Thus, the condition $\tilde{c} > c_0(\bar{v})$ from [Theorem 2](#) holds if and only if $\bar{v} < \frac{16}{17}$, i.e. inserting an artificial bug is beneficial if and only if the budget is constrained by at least $\frac{16}{17}$.

For example, if $\bar{v} = \frac{1}{2}$ and if we do not allow the insertion of any artificial bug, the maximum is attained at $c_0(\bar{v}) = \frac{4}{33}$, leading to a utility of $\frac{32}{363} \approx 0.088$ for the designer. With an artificial bug, we can attain the optimal threshold of $\tilde{c} = \frac{2}{9}$ by choosing, for example, $v = \frac{4}{14}$, $v_a = \frac{3}{14}$ and $q_a = 1$, so that we obtain a utility of $\frac{1}{9} \approx 0.111$.

Next, we illustrate graphically how inserting an artificial bug lowers the designer's financial commitment, in other words allows spending less on the prizes v and v_a . Consider one organic and one artificial bug. In [Figure 1](#), we fix μ , q , and plot the optimal prize sets $\{(v, v_a) : \tilde{c} = v\mu\Phi(\tilde{c}; q) + v_a\Phi(\tilde{c}; q_a)\}$ for different q_a 's on which the maximal utility is achieved, i.e., at the optimal equilibrium threshold, \tilde{c} . The black line labeled $v + v_a = \bar{v}$ represents the budget constraint for $\bar{v} = \frac{2}{3}$. Therefore, the optimal equilibrium threshold

is achievable if the prizes (v, v_a) lie within the budget set. Observe that having a less complex artificial bug, i.e., higher q_a , leads to a larger portion of the optimal prize set being within the budget set. In contrast, if the artificial bug is too complex to find, for example, if $q_a = \frac{1}{5}$, we cannot achieve the optimal threshold.

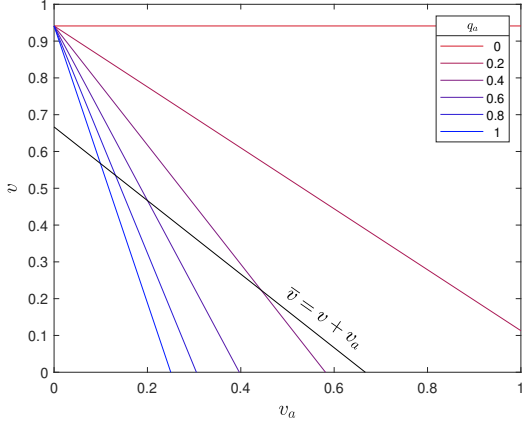


Figure 1: Optimal (v, v_a) given q_a

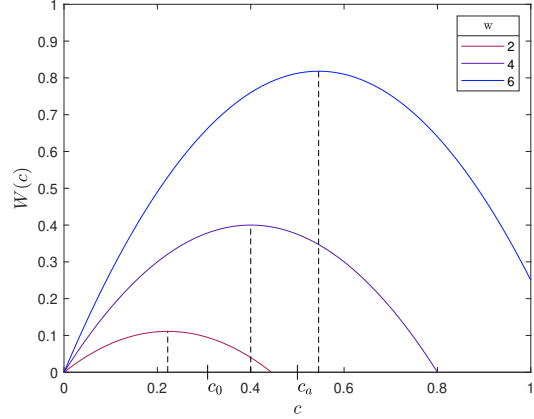


Figure 2: Dependence on w

In [Figure 2](#), we plot the utility $W(\hat{c})$ of the designer against the thresholds \hat{c} for different w 's. Observe that if we do not allow any artificial bugs, we can achieve only thresholds in $[0, c_0(\bar{v})] \approx [0, 0.308]$, whereas an artificial bug helps to achieve thresholds up to $c_a(\bar{v}) = \frac{1}{2}$. For low w (take, for example, $w = 2$), inserting an artificial bug does not help to achieve the maximum. If we increase w , an artificial bug is needed to achieve the maximal utility (say $w = 4$). Increasing w further ($w = 6$) can lead to the case that even when inserting an artificial bug, we are unable to achieve the maximal utility. Nevertheless, inserting an artificial bug leads to a greater utility than without one. Note that we obtain the same observations if we vary the probability μ that a bug exists instead of w .

5.2 Public Program

To illustrate results for the public program, consider the uniform distribution on $[1, 2]$. As before, we assume that there is one type of bug that exists with probability $\mu = \frac{1}{2}$, has complexity $q = \frac{1}{2}$, and that finding it derives a utility of $w = 10$ for the designer. The budget constraint is set to $\bar{v} = 5$. [Figure 3](#) is a visualization of the convergence of

the designer’s objective function $W_n(\hat{c})$ as $n \rightarrow \infty$ as seen in [Corollary 1](#). Observe that the larger the number of agents, the smaller the optimal threshold c_n as $c_n \rightarrow \underline{c} = 1$ from [Lemma 5](#). [Figure 4](#) depicts the same plot, but with the x -axis scaled by $nF(\hat{c})$, to visualize the convergence of $nF(c_n) \rightarrow \kappa^*$. Note that we can also see the convergence of the designer’s objective function as $n \rightarrow \infty$.

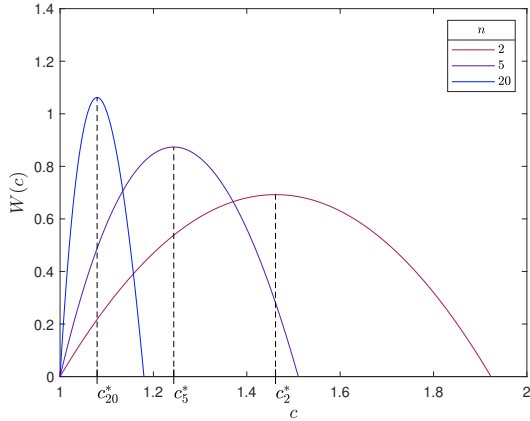


Figure 3: Convergence of $W_n(\hat{c})$ as n grows

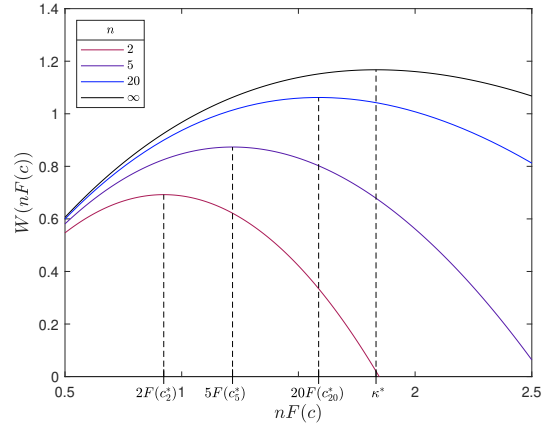


Figure 4: Scaled version

The effect of the budget constraint \bar{v} and the artificial bug q_a in the public program is analogous to those in the private program. We therefore omit its visualization and refer to [Figure 1](#) and [Figure 2](#) in the previous section.

Instead, we illustrate the convergence of the set of solutions for the private program with n agents to the solutions of the public program from [Theorem 4](#). In [Figure 5](#), we show projections of the sets M_n and M_∞ for a fixed $q_a \in \{\frac{1}{3}, \frac{1}{2}, 1\}$. We observe that the Hausdorff distance decreases as n increases. Another interesting observation is that with artificial bugs that are less complex to find, the designer gains more flexibility in setting the prizes in the sense that the set of optimal prizes enlarges and allows less spending.

6 Implementation

The results in the previous sections show that artificial bugs are a welcome instrument for the designer of bug bounty programs, especially when finding organic bugs is important for the designer. Using artificial bugs in the design of bug bounty programs does not pose any particular complexity from a technical perspective, as one can easily introduce

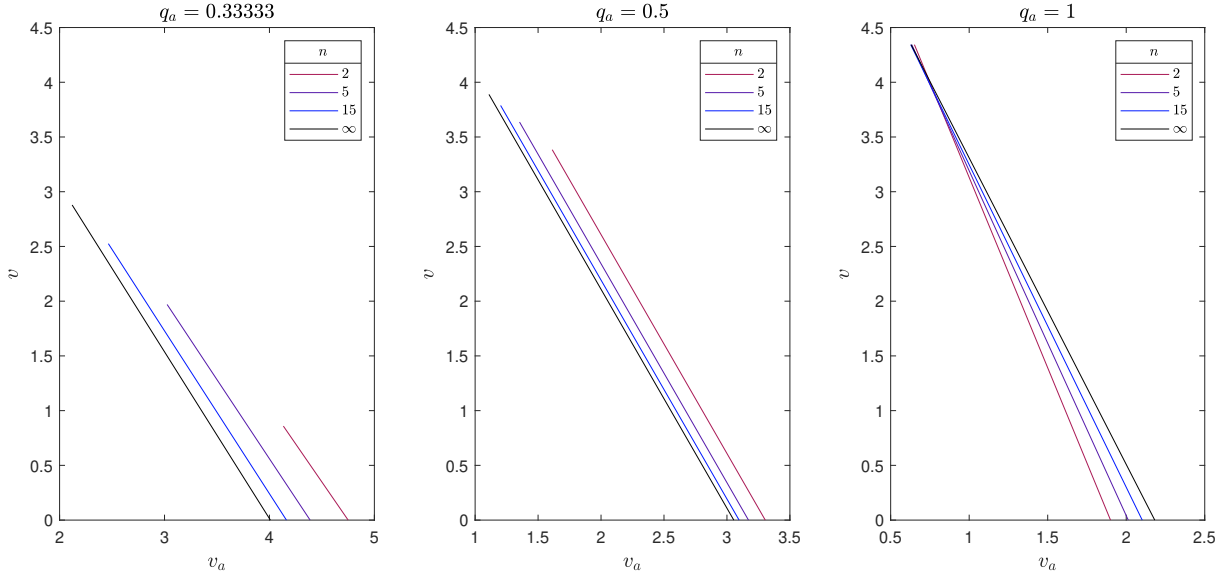


Figure 5: Convergence of the sets M_n to M_∞

simpler or more sophisticated bugs in software packages. Yet, as prizes paid for finding artificial and organic bugs may optimally differ, the designer may want to prove to the finders of the artificial bug, or even to all participants, that an artificial bug found was indeed inserted on purpose and was artificially designed by the designer at the start of the bug bounty program. Even more importantly, if the artificial bug is not found during the crowdsearch, it is important that the designer can prove that an artificial bug has been inserted before the crowdsearch started. This would ensure, or reaffirm, the credibility of the bug bounty program with artificial bugs. We outline three approaches that could be used to achieve this objective: encryption, commitment scheme, and zero-knowledge (ZK) proofs.¹¹

Let us start with asymmetric encryption (see, e.g., [Stallings \(2020\)](#)). The original block of the code and the modified code block with the artificial bug could be encrypted before the crowdsearch starts, and participants can decrypt it once the crowdsearch has ended. More specifically, the designer and participants are connected to a trusted entity that generates a public key and a private key for each participant. The public key information is stored in a directory. Before the crowdsearch starts, the designer encrypts both

¹¹Of course, the credibility of inserting artificial bugs could also be ensured by traditional means. For instance, the designer could invite a notary, who confirms in a written statement that an artificial bug has been inserted.

the original and the modified block of code with the artificial bug using the participant’s public key and sends the encrypted message to all participants. Once the crowdsearch competition has ended, the participants receive their private keys to decrypt the message and verify the existence of the artificial bug.

This approach also allows the artificial bug to only exist with a particular probability, which is common knowledge to the designer and all participants before the crowdsearch starts. The trusted entity flips a biased coin and, based on the outcome of the coin, sends a message to the designer about whether to insert a bug or not. In case no bug should be introduced, the private keys will be empty and not be usable to decrypt any message. This shows to the participants that no artificial bug has been inserted. We note that such an approach is incentive-compatible for the bug bounty designer. If no artificial bug should be inserted, there is also no benefit in inserting such a bug secretly, as such a bug would be treated as an organic bug. If detected, the designer has to pay, and thus, inserting a bug would only be worse off, as the incentive to search for bugs is unaffected.

While the bug bounty platform is a natural candidate for the trusted entity, an alternative approach that does not rely on having a third party is for the designer to use a commitment scheme (Goldreich, 2001). Specifically, before the crowdsearch starts, the designer computes a commitment over the modified block of code with the artificial bug. The designer then publishes the commitment without revealing any further information about the artificial bug (*hiding* property). Once the competition has ended, the designer then opens the commitment and proves to the participants that the designer knows that the modified code block contains an artificial bug (*binding* property).

Inserting an artificial bug with a certain probability can also be implemented with the commitment scheme but needs a public source of randomness and thus ultimately also a trusted third party (see Bonneau and Nikolaenko (2022) for an overview). To implement this, the designer commits to a private key and flips a coin based on it and a public source of randomness. At the end of the crowdsearch, the designer opens the commitment and allows participants to verify the result of the coin flips and that the protocol was followed.

A third approach to reveal an artificial bug convincingly to the participants—irrespective

of whether it has been found or not—is to use zero-knowledge (ZK) proofs. The idea is that before the start of a crowdsearch, the designer demonstrates the existence of an artificial bug in the software, without revealing the location and underlying technique of the bug. Hence, the designer can convince the participants of the artificial bug’s existence without giving any hint on how to find it.

While this concept is attractive in theory, proving bugs in ZK in practical software requires solving a variety of problems and engineering challenges. As pointed out by [Cuéllar et al. \(2023\)](#), ZK frameworks must be able to compile proofs of bugs that require many steps of execution. Moreover, it is necessary to efficiently create understandable statements. Yet, recent advances in the development of proof-statement compilers show promising progress as to how the existence of a bug can be proved without revealing the details or the inputs that could be used to exploit the bug.¹² For our problem, a ZK proof is applicable, as the bug can be chosen to be particularly suitable to the applications of the proof-statement compiler and supporting tools described in [Cuéllar et al. \(2023\)](#) and [Green et al. \(2023\)](#). Hence, the inserted bug can be chosen such that a ZK proof is available. Despite its requirements for sophisticated compilers, a key benefit of ZK proofs over the previous two approaches is that participants can verify *at the start* of the crowdsearch that an artificial bug exists rather than *at the end*.

While the three presented approaches can provide an implementation of artificial bugs as envisioned, which approach is most useful will depend on the specific applications, as well as on the reputation and engineering capabilities of the bug bounty designer.

7 Other Benefits of Artificial Bugs

Our paper suggests that inserting artificial bugs can be useful to the designer of bug bounty programs. We have examined one such benefit, showing that artificial bugs allow the designer to incentivize participants at a lower cost. We conclude the paper by discussing other potential benefits of artificial bugs.

¹²See [Cuéllar et al. \(2023\)](#); [Green et al. \(2023\)](#) and the survey on the applicability of ZKs in various cases, [Ernstberger et al. \(2024\)](#).

First, artificial bugs can help screen invalid submissions, which saves the organization costly resources on triage and verification. Bug bounty programs typically attract large numbers of invalid submissions as security researchers incur relatively small marginal costs to submit a report. [Zhao et al. \(2017\)](#) report that across major bug bounty platforms, less than 25% of submissions are valid, creating tremendous inefficiency in the verification efforts. Organizations can significantly reduce this inefficiency by prioritizing submissions from researchers who also reported artificial bugs. In other words, artificial bugs can serve as a badge for quality submissions.

Second, artificial bugs can be used to gauge participation in bug bounty programs. Organizations receive submissions of vulnerabilities but often cannot determine the number of researchers currently engaging in their bug bounty program. To remedy this uncertainty, artificial bugs can be used as entry checks or milestones in determining the number of active researchers, as well as the extent to which they have probed the system.

Third, artificial bugs can be used to renew interest in the bug bounty program. An empirical analysis of the HackerOne platform shows that bug bounty programs receive less engagement—as proxied by the number of submissions—over time, since researchers are attracted to newly published programs ([Sridhar and Ng, 2021](#)). [Maillart et al. \(2017\)](#) show that this front-loading effect results in the probability of finding bugs decaying at a rate $1/t^{0.4}$, where t is the time since the program was launched. Inserting artificial bugs may help enhance interest in an otherwise unattractive program.

8 Conclusion

We have developed a simple device to improve the effectiveness of bug bounty systems. The additional benefits outlined in the last section strengthen the proposal and establish promising directions for future research into modeling, utilizing, and implementing the insertion of artificial bugs into a bug bounty system.

References

- Akgul, O., Eghtesad, T., Elazari, A., Gnawali, O., Grossklags, J., Mazurek, M. L., Votipka, D., and Laszka, A. (2023). Bug Hunters’ perspectives on the challenges and benefits of the bug bounty ecosystem. In *32nd USENIX Security Symposium (USENIX Security 23)*, pages 2275–2291, Anaheim, CA. USENIX Association.
- Böhme, R. (2006). A Comparison of Market Approaches to Software Vulnerability Disclosure. In Müller, G., editor, *Emerging Trends in Information and Communication Security*, Lecture Notes in Computer Science, pages 298–311, Berlin. Springer.
- Böhme, R., Eckey, L., Moore, T., Narula, N., Ruffing, T., and Zohar, A. (2020). Responsible Vulnerability Disclosure in Cryptocurrencies. *Communications of the ACM*, 63(10):62–71.
- Bonneau, J. and Nikolaenko, V. (2022). Public Randomness and Randomness Beacons. <https://a16zcrypto.com/posts/article/public-randomness-and-randomness-beacons/>.
- Breidenbach, L., Daian, P., Tramèr, F., and Juels, A. (2018). Enter the Hydra: Towards Principled Bug Bounties and Exploit-Resistant Smart Contracts. In *27th USENIX Security Symposium (USENIX Security 18)*, pages 1335–1352, Baltimore, MD. USENIX Association.
- Cuéllar, S., Harris, B., Parker, J., Pernsteiner, S., and Tromer, E. (2023). Cheesecloth: Zero-Knowledge proofs of real world vulnerabilities. In *32nd USENIX Security Symposium (USENIX Security 23)*, pages 6525–6540, Anaheim, CA. USENIX Association.
- Dubey, P. (2013). The role of information in contests. *Economics Letters*, 120(2):160–163.
- Ernstberger, J., Chaliasos, S., Zhou, L., Jovanovic, P., and Gervais, A. (2024). Do you need a zero knowledge proof? Cryptology ePrint Archive, Paper 2024/050. <https://eprint.iacr.org/2024/050>.
- Federal Department of Finance of Switzerland (2022). Federal administration procures platform for bug bounty programmes [Press release]. <https://www.admin.ch/gov/en/start/documentation/media-releases.msg-id-89868.html>.
- Gersbach, H., Mamageishvili, A., and Pitsuwan, F. (2023a). Crowdsearch. *CEPR Discussion Paper No. 18529*.
- Gersbach, H., Mamageishvili, A., and Pitsuwan, F. (2023b). Decentralized Attack Search and the Design of Bug Bounty Schemes.
- Ghosh, A. and Kleinberg, R. (2016). Optimal Contest Design for Simple Agents. *ACM Transactions on Economics and Computation*, 4(4):22:1–22:41.
- Goldreich, O. (2001). *Foundations of Cryptography*. Cambridge University Press.
- Green, M., Hall-Andersen, M., Hennenfent, E., Kaptchuk, G., Perez, B., and Laer, G. V. (2023). Efficient Proofs of Software Exploitability for Real-world Processors. In *Proceedings on Privacy Enhancing Technologies*, pages 627–640.

- Laszka, A., Zhao, M., Malbari, A., and Grossklags, J. (2018). The Rules of Engagement for Bug Bounty Programs. In Meiklejohn, S. and Sako, K., editors, *Financial Cryptography and Data Security*, Lecture Notes in Computer Science, pages 138–159, Berlin, Heidelberg. Springer.
- Levin, D. and Smith, J. L. (1994). Equilibrium in Auctions with Entry. *The American Economic Review*, 84(3):585–599.
- Levy, P., Sarne, D., and Rochlin, I. (2017). Contest Design with Uncertain Performance and Costly Participation. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*, IJCAI 2017, pages 302–309.
- Liu, B. and Lu, J. (2019). The optimal allocation of prizes in contests with costly entry. *International Journal of Industrial Organization*, 66:137–161.
- Maillart, T., Zhao, M., Grossklags, J., and Chuang, J. (2017). Given enough eyeballs, all bugs are shallow? Revisiting Eric Raymond with bug bounty programs. *Journal of Cybersecurity*, 3(2):81–90.
- Malladi, S. S. and Subramanian, H. C. (2020). Bug Bounty Programs for Cybersecurity: Practices, Issues, and Recommendations. *IEEE Software*, 37(1):31–39.
- Moreno, D. and Wooders, J. (2011). Auctions with heterogeneous entry costs. *The RAND Journal of Economics*, 42(2):313–336.
- Sarne, D. and Lepioshkin, M. (2017). Effective Prize Structure for Simple Crowdsourcing Contests with Participation Costs. In *Fifth AAAI Conference on Human Computation and Crowdsourcing*.
- Sridhar, K. and Ng, M. (2021). Hacking for good: Leveraging HackerOne data to develop an economic model of Bug Bounties. *Journal of Cybersecurity*, 7(1):1–9.
- Stallings, W. (2020). *Cryptography and Network Security: Principles and Practice*. Pearson, 8th edition.
- Walshe, T. and Simpson, A. (2020). An Empirical Study of Bug Bounty Programs. In *2020 IEEE 2nd International Workshop on Intelligent Bug Fixing (IBF)*, pages 35–44.
- Zhao, M., Laszka, A., and Grossklags, J. (2017). Devising Effective Policies for Bug-Bounty Platforms and Security Vulnerability Discovery. *Journal of Information Policy*, 7:372–418.
- Zrahia, A., Gandal, N., Markovich, S., and Riordan, M. H. (2022). The Simple Economics of an External Shock on a Crowdsourced ‘Bug Bounty Platform’. Available at SSRN: <https://ssrn.com/abstract=4154516>.

A Proofs

Proof of Lemma 1. First, we show that every equilibrium strategy is a threshold strategy. Let $\sigma_i : [\underline{c}, \bar{c}] \rightarrow \{0, 1\}$ denote agent i 's strategy and let $\boldsymbol{\sigma}_{-i}$ denote the strategy profile of agents other than i . Given a strategy profile, the payoff of agent i is

$$u_i(\sigma_i, \boldsymbol{\sigma}_{-i}, c_i) = \sigma_i \left(\sum_l v^l \mu^l p(\boldsymbol{\sigma}_{-i}; q^l) + \sum_k v_a^k p(\boldsymbol{\sigma}_{-i}; q_a^k) - c_i \right),$$

where $p(\boldsymbol{\sigma}_{-i}; q)$ is the probability that agent i wins the prize corresponding to the bug with complexity q , given the strategy $\boldsymbol{\sigma}_{-i}$ of others.

Let σ_i^* be an equilibrium strategy for agent i . We prove by contradiction that σ_i^* is non-increasing. Suppose that there exists a pair of costs $c < c'$ such that $\sigma_i^*(c) = 0$ and $\sigma_i^*(c') = 1$. Given the private cost vector $\mathbf{c} = (c_1, \dots, c_n)$,

$$\begin{aligned} 0 &= \mathbb{E}[u_i(0, \boldsymbol{\sigma}_{-i}^*(c_{-i}), c_i) | c_i = c] \geq \mathbb{E}[u_i(1, \boldsymbol{\sigma}_{-i}^*(c_{-i}), c_i) | c_i = c] \\ &> \mathbb{E}[u_i(1, \boldsymbol{\sigma}_{-i}^*(c_{-i}), c_i) | c_i = c'] \geq \mathbb{E}[u_i(0, \boldsymbol{\sigma}_{-i}^*(c_{-i}), c_i) | c_i = c'] = 0 \end{aligned}$$

where the first and the last inequalities follow from the definition of the equilibrium. The strict inequality, which follows from the fact that u_i is strictly decreasing in c leads to a contradiction. Thus, any equilibrium strategy is a threshold strategy.

Furthermore, since we consider symmetric equilibrium, all agents use the same threshold strategy denoted by c^* . In equilibrium, the agent with search cost at the threshold must be indifferent between searching and not searching. Therefore, if the equilibrium threshold lies in the interior of $[\underline{c}, \bar{c}]$, it must hold that

$$c^* = \sum_l v^l q^l \Phi(c^*; q^l) + \sum_k v_a^k \Phi(c^*; q_a^k),$$

where $\Phi(\hat{c}; q)$ is the expectation of $p(\boldsymbol{\sigma}_{-i}(\mathbf{c}_{-i}); q)$, given the cost distribution of the other agents when they all use the threshold.

To conclude, note that $\Phi(\hat{c}; q)$ is decreasing in \hat{c} for every $q \in (0, 1]$ because a higher threshold adopted by the other agents increases their probability to search and find the

bug, which, in turn, must lower agent i 's probability of being rewarded for that particular bug. It follows that Ψ is also decreasing in \hat{c} . Thus, if $\Psi(\underline{c}) \leq \underline{c}$, every agent can only expect to incur a loss by participating in the search regardless of his/her skills, and hence $c^* = \underline{c}$. If $\Psi(\bar{c}) \geq \bar{c}$, every agent can only profit by participating, and therefore $c^* = \bar{c}$. Otherwise, we have that $\Psi(\underline{c}) < \underline{c}$ and $\bar{c} < \Psi(\bar{c})$. By continuity and monotonicity of Ψ , there exists a unique fixed point. This completes the proof. \square

Proof of Lemma 2. Suppose that $(\mathbf{v}^*, v_a^*, q_a^*)$ solves the designer's problem (2) and assume w.l.o.g. that $q_a^* \neq 0$ (else there is nothing to prove). Recall that the equilibrium threshold is given by

$$c^* = \sum_l v^{l*} \mu^l \Phi(c^*; q^l) + \sum_k v_a^{k*} \Phi(c^*; q_a^{k*}). \quad (5)$$

Introduce $\tilde{k} \equiv \arg \max_k \{q_a^{k*}\}$ and choose the new complexities $\mathbf{q}_a^{**} = (q_a^{1**}, 0, \dots, 0)$ with $q_a^{1**} = q_a^{\tilde{k}*}$. Moreover let $\mathbf{v}_a^{**} = (v_a^{1**}, 0, \dots, 0)$ where the first prize is given by

$$v_a^{1**} = v_a^{\tilde{k}*} + \frac{\sum_{k \neq \tilde{k}} v_a^{k*} P(c^*; q_a^{k*})}{P(c^*; q_a^{1**})}.$$

Observe that with this choice, the second sum in equation (5) does not change, i.e.

$$\sum_k v_a^{k*} \Phi(c^*; q_a^{k*}) = \sum_k v_a^{**k} \Phi(c^*; q_a^{**k}) = v_a^{1**} \Phi(c^*; q_a^{1**}). \quad (6)$$

Thus if we take $\mathbf{v}^{**} = \mathbf{v}^*$ we obtain the same equilibrium threshold and therefore the designer receives the same utility because of equation (6). Note that the boundary conditions of (2) are satisfied since $P(c^*; q)$ is increasing in q . Hence, it is sufficient to introduce only one artificial bug. \square

Proof of Lemma 3. Recall that $\Phi(\hat{c}; q)$ is the probability that an agent wins the prize for the bug with complexity q when participating. Since agents find a given bug

independently of other agents, we can write

$$\begin{aligned}\Phi(\hat{c}; q) &= q \sum_{k=0}^{n-1} \left\{ \binom{n}{k} F(\hat{c})^k (1 - F(\hat{c}))^{n-1-k} \left[\sum_{t=0}^k \binom{k}{t} \frac{q^t (1-q)^{k-t}}{t+1} \right] \right\} \\ &= \frac{1 - (1 - qF(\hat{c}))^n}{nF(\hat{c})} = \frac{P(\hat{c}; q)}{nF(\hat{c})},\end{aligned}$$

where we used a generalization of the binomial theorem twice. For details, see the proof of Proposition 3 and Fact 1 in [Gersbach et al. \(2023a\)](#). \square

Proof of Lemma 4. Observe that the set of achievable equilibrium thresholds $C(\bar{v})$ has to be an interval since $\sum_l v^l \mu^l \Phi(\hat{c}; q^l) + v_a \Phi(\hat{c}; q_a)$ is continuous. The smallest possible value for \hat{c} is 0, which can be achieved by choosing $v_a = 0$, $q_a = 0$ and $v^l = 0 \quad \forall l$. Moreover for any l we know from [Lemma 3](#) that

$$\Phi(\hat{c}; q^l) = \frac{P(\hat{c}; q^l)}{nF(\hat{c})} = \frac{1 - (1 - q^l F(\hat{c}))^n}{nF(\hat{c})} \leq \frac{1 - (1 - F(\hat{c}))^n}{nF(\hat{c})} = \Phi(\hat{c}; 1).$$

This, together with the fact that $\Phi(\hat{c}, 1)$ is decreasing in \hat{c} (see Proof of [Lemma 3](#)), tells us that the highest $\hat{c} \in C(\bar{v})$ is achieved by allocating all of the reward \bar{v} to the finding of an artificial bug with complexity $q_a = 1$, i.e. $\max C(\bar{v}) = c_a(\bar{v})$ with $c_a(\bar{v})$ the fixed point of $\bar{v}\Phi(\hat{c}, 1)$. Thus $C(\bar{v}) = [0, c_a(\bar{v})]$. \square

Proof of Theorem 1. Recall that we can write the designer's problem as

$$\max_{\hat{c} \in [0, c_a(\bar{v})]} \sum_l w^l \mu^l P(\hat{c}; q^l) - nF(\hat{c})\hat{c}.$$

Taking the derivative with respect to \hat{c} leads to the first-order condition

$$\begin{aligned}\frac{d}{d\hat{c}} \left(\sum_l w^l \mu^l P(\hat{c}; q^l) - nF(\hat{c})\hat{c} \right) &= \sum_l \frac{d}{d\hat{c}} (w^l \mu^l (1 - (1 - q^l F(\hat{c}))^n) - nF(\hat{c})\hat{c}) \\ &= \sum_l w^l \mu^l n (1 - q^l F(\hat{c}))^{n-1} q^l f(\hat{c}) - n f(\hat{c})\hat{c} - nF(\hat{c}) \\ &= 0,\end{aligned}$$

where we used that $P(\hat{c}; q^l) = 1 - (1 - q^l F(\hat{c}))^n$. By rearranging, we see that

$$\hat{c} = \sum_l w^l \mu^l q^l (1 - q^l F(\hat{c}))^{n-1} - \frac{F(\hat{c})}{f(\hat{c})} = \Omega(\hat{c}),$$

i.e. \hat{c} is a fixed point of Ω . Note that $(1 - q^l F(\hat{c}))^{n-1}$ is decreasing in \hat{c} and recall that we assumed F/f is non-decreasing, and therefore $\Omega(\hat{c})$ is decreasing in \hat{c} . Hence, $\Omega(\hat{c})$ has a unique fixed point \tilde{c} .

This fixed point has to be a maximum since the derivative is $\frac{dW(\hat{c})}{d\hat{c}} = n f(\hat{c})(\Omega(\hat{c}) - \hat{c})$, and therefore $\frac{dW}{d\hat{c}}(\hat{c}) > 0$ for $\hat{c} < \tilde{c}$ and $\frac{dW}{d\hat{c}}(\hat{c}) < 0$ for $\tilde{c} < \hat{c}$. By [Lemma 4](#), we know that the set of achievable thresholds is $C(\bar{v}) = [0, c_a(\bar{v})]$ and since the utility is increasing on $[0, \tilde{c}]$, the optimal achievable equilibrium threshold is $\hat{c}^* = \min\{\tilde{c}, c_a(\bar{v})\}$. Using [Lemma 1](#), we can conclude that any prizes v^*, v_a^* and complexity q_a^* that solve $\hat{c}^* = \sum_l v^l \mu^l \Phi(\hat{c}^*; q^l) + v_a \Phi(\hat{c}^*; q_a)$ are optimal. \square

Proof of [Theorem 2](#). This theorem follows from the observation that not inserting an artificial bug reduces the set of achievable equilibrium thresholds to $C_0(\bar{v}) = [0, c_0(\bar{v})] \subseteq C(\bar{v})$. Since by [Theorem 1](#), the optimal equilibrium threshold is given by $\hat{c}^* = \min\{\tilde{c}, c_a(\bar{v})\}$ and $c_0(\bar{v}) \leq c_a(\bar{v})$, this affects the optimal solution if and only if $\tilde{c} > c_0(\bar{v})$. \square

Proof of [Lemma 5](#).

- (i) Recall from [Lemma 1](#) that for any bug with complexity q we have that $P_n(\hat{c}; q) = nF(\hat{c})\Phi_n(\hat{c}; q)$. Thus the equilibrium condition can be written as

$$c_n nF(c_n) = \sum_l v^l \mu^l P_n(c_n; q^l) + v_a P_n(c_n; q_a). \quad (7)$$

Assume by contradiction that $\lim_{n \rightarrow \infty} c_n$ is not equal to \underline{c} . By using a generalized binomial theorem, one can prove that $\Psi_n(c)$ is decreasing in n . Therefore, c_n has to be decreasing in n . Hence, there has to exist some value $c' > \underline{c}$ such that $c_n > c'$ for any $n \in \mathbb{N}$. In this case, $c_n nF(c_n) > c' nF(c')$. This holds because F is strictly increasing. Note that $F(c') > 0$, since $c' > \underline{c}$. Therefore, $\lim_{n \rightarrow \infty} c_n nF(c_n) \geq$

$\lim_{n \rightarrow \infty} c' n F(c') = \infty$, which cannot be equal to

$$\lim_{n \rightarrow \infty} \sum_l v^l \mu^l P_n(c_n; q^l) + v_a P_n(c_n; q^a) \leq \sum_l v^l + v_a,$$

i.e., we obtain a contradiction.

- (ii) For the second part, we assume that $nF(c_n)$ converges as $n \rightarrow \infty$ and denote the limit by $\tilde{\kappa}$.¹³ Next define $\kappa^* \in (0, \infty)$ as the unique fixed point of $\Psi_\infty(\hat{\kappa}) = \sum_l v^l \mu^l \frac{1 - e^{-q^l \hat{\kappa}}}{\underline{c}} + v_a \frac{1 - e^{-q_a \hat{\kappa}}}{\underline{c}}$. The fixed point exists since $\Psi_\infty(0) = 0$ and $\Psi_\infty(\hat{\kappa})$ is bounded by $\sum_l v^l + v_a$. Uniqueness follows from the fact that $\Psi_\infty(\hat{\kappa})$ is strictly increasing in $\hat{\kappa}$. Thus, we only need to prove that $\kappa^* = \tilde{\kappa}$. Note that for any complexity $q \in [0, 1]$ it holds that

$$\lim_{n \rightarrow \infty} (1 - qF(c_n))^n = \lim_{n \rightarrow \infty} \left[\left(1 - \frac{q}{1/F(c_n)} \right)^{\frac{1}{F(c_n)}} \right]^{nF(c_n)} = e^{-q\tilde{\kappa}}, \quad (8)$$

by the fact that $F(c_n) \rightarrow F(\underline{c}) = 0$, the definition of e^x , and the continuity of the function. Using this observation together with equation (7) from part (i) leads to

$$\begin{aligned} \underline{c}\tilde{\kappa} &= \lim_{n \rightarrow \infty} c_n n F(c_n) = \lim_{n \rightarrow \infty} \sum_l v^l \mu^l P_n(c_n; q^l) + v_a P_n(c_n; q_a) \\ &= \sum_l v^l \mu^l \lim_{n \rightarrow \infty} (1 - (1 - q^l F(c_n))^n) + v_a \lim_{n \rightarrow \infty} (1 - (1 - q_a F(c_n))^n) \\ &= \sum_l v^l \mu^l (1 - e^{-q^l \tilde{\kappa}}) + v_a (1 - e^{-q_a \tilde{\kappa}}). \end{aligned}$$

In particular, $\tilde{\kappa}$ is a fixed point of $\Psi_\infty(\hat{\kappa})$ and by uniqueness we can conclude that $\tilde{\kappa} = \kappa^*$.

- (iii) The proof of $\lim_{n \rightarrow \infty} P_n(q) = 1 - e^{-q\kappa^*}$ follows directly from equation (8). □

Proof of Corollary 1. The corollary follows directly from Lemma 5. □

¹³It can be proven that, under the maintained assumptions, $nF(c_n)$ converges by using the arguments from Proposition 9 in Gersbach et al. (2023a).

Proof of Lemma 6. Observe that

$$\Psi_\infty(\hat{\kappa}; \mathbf{v}, v_a, q_a) = \sum_l v^l \mu^l \frac{1 - e^{-q^l \hat{\kappa}}}{\underline{c}} + v_a \frac{1 - e^{-q_a \hat{\kappa}}}{\underline{c}}$$

is increasing in $\hat{\kappa}$ and q_a . Thus, the lemma can be proven analogously to Lemma 4. \square

Proof of Theorem 3. As in the proof of Theorem 1, we use the first-order condition

$$\frac{d}{d\hat{\kappa}} \left(\sum_l w^l \mu^l (1 - e^{-q^l \hat{\kappa}}) - \hat{\kappa} \underline{c} \right) = \sum_l w^l \mu^l q^l e^{-q^l \hat{\kappa}} - \underline{c} = \Omega_\infty(\hat{\kappa}).$$

By assumption, we have $\Omega_\infty(0) \geq 0$ and $\Omega_\infty(\hat{\kappa}) \rightarrow -\underline{c}$ as $\hat{\kappa} \rightarrow \infty$, we obtain an unique $\tilde{\kappa}$. The rest of the proof uses the same arguments as the proof of Theorem 1. \square

The proof of Theorem 4 requires the following lemma.

Lemma A.1. *It holds that $nF(\hat{c}_n^*) \rightarrow \hat{\kappa}^*$ as $n \rightarrow \infty$.*

Proof of Lemma A.1. Note that $\Omega(c)$ approaches $-F(c)/f(c)$ for $c > \underline{c}$ as n goes to infinity, while $\Omega(0) = \sum_l w^l q^l \mu^l$ for all n . Hence, $\hat{c}_n^* \rightarrow \underline{c}$ as $n \rightarrow \infty$.

From here, we can apply the same arguments as in the proof of Proposition 9 in Gersbach et al. (2023a) by replacing $V \frac{1 - (1 - qF(c_n))^n}{c_n}$ with $\sum_l v^l \mu^l \frac{1 - (1 - q^l F(\hat{c}_n^*))^n}{\hat{c}_n^*} + v_a \frac{1 - (1 - q_a F(\hat{c}_n^*))^n}{\hat{c}_n^*}$ and V/\underline{c} with \bar{v}/\underline{c} . In step 3, the argument still holds, due to the fact that the sum of strictly monotone and continuous functions is also strictly monotone and continuous. \square

Proof of Theorem 4. For the proof of the theorem, we remove first the budget constraint for the prize for the first bug. Later, the constraint is reintroduced by taking the intersection of M_n , respectively M_∞ , with a closed subset of \mathbb{R}^{L+2} , and therefore, the convergence will not be affected. Take an arbitrary $(\mathbf{v}_n^*, v_{a,n}^*, q_{a,n}^*) \in M_n$. Define $(\mathbf{v}^*, v_a^*, q_a^*) \in M_\infty$ by setting $v_a = v_{a,n}$, $q_a^* = q_{a,n}^*$, $(v^2, \dots, v^L) = (v_n^2, \dots, v_n^L)$ and

$$v^1 = \frac{\underline{c}}{\mu^1 (1 - e^{-q^1 \hat{\kappa}^*})} \left(\hat{\kappa}^* \underline{c} - \sum_{l \neq 1} v^l \mu^l (1 - e^{-q^l \hat{\kappa}^*}) - v_a (1 - e^{-q_a \hat{\kappa}^*}) \right).$$

Recall that v_n^1 of the public program is given by

$$v_n^1 = \frac{\hat{c}_n^*}{\mu^1 P(\hat{c}_n^*; q_1)} \left(nF(\hat{c}_n^*)\hat{c}_n^* - \sum_{l \neq 1} v^l \mu^l P(\hat{c}_n^*; q^l) - v_a P(\hat{c}_n^*; q_a) \right)$$

and thus, by taking the limit for every summand separately and [Lemma A.1](#), we obtain

$$\lim_{n \rightarrow \infty} v^1 - v_n^1 = \frac{\underline{c}}{\mu^1 (1 - e^{-q^1 \hat{\kappa}^*})} \lim_{n \rightarrow \infty} v_a \left((1 - e^{-q_a \hat{\kappa}^*}) - P(\hat{c}_n^*; q_a) \right).$$

Since the sequence $e^{-nx} - (1-x)^n$ converges uniformly on bounded intervals and $q_a \in [0, 1]$, the right-hand side converges uniformly, and therefore

$$\sup_{m_n \in M_n} \inf_{m \in M_\infty} \|m_n - m\|_2 \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

In the same way, one can prove that

$$\sup_{m \in M_\infty} \inf_{m_n \in M_n} \|m_n - m\|_2 \rightarrow 0, \quad \text{as } n \rightarrow \infty,$$

and therefore, $d(M_n, M_\infty) \rightarrow 0$ in the Hausdorff distance. \square

Proof of [Theorem 5](#). The proof of the first equivalence relation is analogous to the proof of [Theorem 2](#). For the second equivalence, note that $\kappa^l(\bar{v}) = \lim_{n \rightarrow \infty} nF(c_n^l(\bar{v}))$ for all $l \in \{1, \dots, L\}$ by [Lemma 5](#). Assume that $\tilde{\kappa} > \kappa_0(\bar{v})$, which is equivalent to

$$\lim_{n \rightarrow \infty} \min\{nF(c_n), nF(c_a(\bar{v}))\} > \lim_{n \rightarrow \infty} \max_l \{nF(c_n^l(\bar{v}))\}.$$

This is the case exactly when there exists N , such that for all $n \geq N$, we have

$$\min\{nF(c_n), nF(c_a(\bar{v}))\} > \max_l nF(c_n^l(\bar{v})).$$

But is equivalent to $c_{0,n}(\bar{v}) > \tilde{c}_n$ for $n \geq N$, since $F(c)$ is strictly increasing in c as a cumulative distribution function. \square