JOURNAL OF CYBERSECURITY

Journal of Cybersecurity, 2021, 1–9
doi: 10.1093/cybsec/tyab007
Research Paper

Research Paper

Hacking for good: Leveraging HackerOne data to develop an economic model of Bug Bounties

Kiran Sridhar [1,*] and Ming Ng[2]

[1]Stanford University; Operations and Technology Management, University of Cambridge, CA 94103 and [2]Department of Data Science, HackerOne, San Francisco, USA

*Correspondence address. Judge Business School, Trumpington Street, Cambridge, CB2 1AG, UK.
E-mail: kirsrid@gmail.com

Received 1 March 2020; revised 10 November 2020; accepted 22 January 2021

Abstract

We ran a study of bug bounties, programs where gig economy security researchers are compensated for pinpointing and explaining vulnerabilities in company code bases. Bug bounty advocates have argued that they are a cost-effective means for companies of all types to shore up their security posture. Our research—which analyzes a large, proprietary dataset and which leverages instrumental variables to eliminate potential sources of endogeneity—provides empirical support for this assertion. Security researchers have a price elasticity of supply of between 0.1 and 0.2 at the median, indicating that they are largely motivated by non-pecuniary factors; a company is still able to derive utility from bug bounties even if they have a limited ability to pay security researchers. Moreover, a company's revenue and brand profile do not have an economically significant impact on the number of valid security vulnerabilities reports its program receives. However, we found that companies in the finance, retail, and healthcare sectors are notified of fewer valid vulnerabilities, ceteris paribus, than companies in other sectors, though these estimates are not statistically significant at the 5% level. We also found no evidence that new companies joining the HackerOne platform dampen the number of reports that firms receive. Finally, we find that programs receive fewer valid reports as they grow older and bugs become harder to find. This negative age effect may be dampened if the program increases the code base available for hacking.

Key words: bug bounties; cybersecurity; crowdsourced cybersecurity; IT management; HackerOne

Introduction

Many firms are skeptical about receiving reports of cybersecurity vulnerabilities from third-party researchers. According to HackerOne, 93% of companies in the Forbes Global 2000 lack vulnerability disclosure policies (VDPs), which stipulate how security researchers can submit bugs to organizations without fear of being sued. This has a chilling effect, preventing companies from learning about the blind spots in their cybersecurity posture. A survey of 1698 of HackerOne's top security researchers found that nearly 25% had withheld submitting a vulnerability out of liability fears because the company had not established a VDP [1].

However, attitudes are rapidly shifting. Speaking at the Global Cybersecurity Summit in 2017, then-US Deputy Attorney General Rod Rosenstein advocated that "all companies consider promulgating a vulnerability disclosure policy" [2]. And in late 2019, the US Department of Homeland Security, the federal government agency tasked with cybersecurity, instructed all government agencies to create a VDP [3]. Major cyber attacks have resulted in significant financial losses and have dramatically increased public consciousness about technology security. Investors, officers, and directors understand that companies must do everything in their power to learn about their cybersecurity weaknesses and prevent attacks; to behave otherwise amounts to corporate malpractice.

But many companies are going even further, participating in bug bounty markets, where freelance security researchers (also called hackers) are tasked with finding bugs in corporate IT systems and code bases and cogently explaining them to the companies.[1] Companies provide hackers with monetary rewards, called bounties, based upon their performance. Major bug bounty platforms, including HackerOne and Bugcrowd, have already facilitated the payment of hundreds of millions of dollars; in May 2020, HackerOne reached the threshold of $100 million in bug bounty payments. And the industry is rapidly growing. Half of HackerOne's bounties were paid in the last year and Gartner projects that by 2022, 50% of enterprises will employ crowdsourced cybersecurity [5].

As we will elaborate in our literature review, there are several powerful rationales for bug bounty programs. First, they help companies identify bugs in their code base that they may not be aware of; an axiom in programming is that when more eyeballs examine a software product, more vulnerabilities are remediated. Second, they provide companies that lack the cachet to recruit top-tier talent with an outlet to engage freelance hackers.

However, there is a big lacuna in the literature: there has been limited empirical study of this increasingly important industry. Past bug bounty research has been hindered by limited publicly available data and has struggled to establish causality. As a result, researchers have yet to definitively establish the effects that a company's revenue, industry, and brand profile have on the number of reports that their programs receive. Similarly, there is an ongoing debate about whether new programs dampen reports to existing programs. Moreover, the supply elasticity of hackers has heretofore not been calculated.

Our paper fills much of this void. We explore the organizations that benefit the most from bug bounties by evaluating the factors that determine the number of valid reports that each program i receives in month j. We leverage HackerOne's database to examine panel program data from August 2014 to January 2020. Our data set, composing of over 3800 observations, is far more comprehensive than those of previous papers. To better establish causality, we employ a two-stage least squares (2SLS) regression identification strategy. Much like Ramey [6], we leverage narrative instruments to address endogeneity. We draw upon a comprehensive database of major public breaches, which could induce more companies to embrace bug bounties. We also address endogeneity through lagged instrumental variables [7, 8] and fixed effect regressions.

While this article makes significant strides toward establishing causality, there may still be omitted variable bias. We were unable to include measures of report severity and scope—the number of assets for which a hacker can find vulnerabilities. These omitted variables may be correlated with several of the independent variables we included in our regressions, thereby biasing our parameter estimates.

This article has a number of significant findings. First, it finds that hackers are price insensitive—with an elasticity at the median of between 0.1 and 0.2—indicating that companies with limited resources can still derive value from bug bounties. This is the first time that hacker price elasticity has been estimated in the academic literature. Second, it finds that a company's size and profile do not have an economically significant impact on the number of reports it receives, reinforcing the value of bug bounties for smaller, less prestigious companies. Third, it finds that finance, retail, and healthcare companies receive fewer reports, all else being equal, than companies in other industries, though researchers should amass more data to generate industry coefficient estimates with greater statistical significance.Fourth, it finds that new programs have a statistically insignificant impact on the number of reports which companies receive. If these results hold in the future, then companies will continue to derive benefits from bug bounties even as the number of new programs multiplies. Fifth, we find that programs receive fewer reports as they grow older. This age effect may be ameliorated if a program expands its scope—the attack surface which bug bounty security researchers are eligible to hack. Sixth, we underscore how much research on bug bounties still required: our regression only accounts for 40% of the variation we observe in the data.

The layout of our article is as follows. In our background section, we discuss the theoretical benefits of bug bounties; the factors that have been posited to impact ethical hacker supply (many of which we will include in our regression); and past empirical studies of bug bounty markets. In our methodology section, we introduce our data set, present summary statistics on it and explain our identification strategy. We then share the results from our empirical work and the implications of our work. Finally, we discuss potential avenues for future research. The article draws upon both literature and interviews with HackerOne employees, security researchers, and a former chief security officer who managed a bug bounty program.

Background

The logic of bug bounties

The literature proposes two rationales for enterprises to seek out freelance hackers to find bugs—one practical and one theoretical. First, there is an acute worldwide shortage of 4 million cybersecurity professionals according to (ISC)[2], an IT industry association [9]. Companies with less cachet, particularly small and medium enterprises (SMEs), find it difficult to recruit workers in this competitive environment. Their problems are exacerbated as major companies, recognizing the immense operational and reputational costs of cyber attacks, recruit and retain the most talented cybersecurity professionals. JP Morgan Chase, for instance, spends around $600 million on cybersecurity [10]. This is perverse, as it is SMEs that need to recruit talent the most. A study by Verizon found that 60% of small businesses shut down within 6 months of suffering a major breach [11]. Bug bounty programs allow firms that might struggle to employ talent to bring in freelance security researchers and better protect against a top downside risk. In this way, bug bounties are a component of the gig economy. The vast majority of security researchers, who are paid a median of $800 per bounty, work part time and 27% of them are full-time students, who appreciate the flexibility afforded by freelance security research [12].[2]

Second, academics have suggested that bug bounties enable companies of all sizes to discover vulnerabilities that they would otherwise overlook. This can best be explained by Linus's Law, a dictum first posited by software developer Eric Raymond: "Given enough

1 Bug bounties are distinct from penetration testing because they rely on gig workers as opposed to professional security researchers, though the top bug bounty hackers may also work for penetration testing companies [4]

2 Of course, standard economic theory would dictate that the most experienced hackers would want to spend their time searching for vulnerabilities in the companies with the greatest ability to pay, leaving smaller companies with less experienced hackers. This is why in our article, we measure the price elasticity of hackers and the effects of revenue on hacker supply.

eyeballs, all bugs are shallow" [13]. Different programmers and hackers have different skill sets, use varying testing methods, and are better positioned to identify unique sets of bugs [14]. This suggests that firms should employ a host of methods and diverse groups of people to find the greatest number of bugs. Through bounty programs, firms are able to engage hackers from around the world, complementing the work undertaken by their internal security teams.

Our article's goal is to provide empirical evidence to substantiate these claims. Do bug bounties improve the cybersecurity of all organizations—not just the biggest, best resourced, most pedigreed ones? Do HackerOne hackers help firms find the vulnerabilities that their internal technical teams missed? What motivates HackerOne's security researchers to submit bugs to particular organizations? We answer these questions by examining the factors that impact the number of valid vulnerabilities bug bounty programs receive.

Factors potentially impacting security researcher supply

The computer science, economics, and cybersecurity literatures have posited that multiple factors influence the number of security researcher reports companies receive on HackerOne and other bug bounty platforms. We incorporate these factors in our regression model to answer our questions of interest.

Program age

Some code base vulnerabilities can be discovered by scanners [15] that are either custom-made or off-the-shelf (like Metasploit, an increasingly popular tool). There is almost zero marginal cost associated with finding and reporting these bugs. However, identifying other vulnerabilities require a lot of time and reconnaissance: a hacker must spend hours mapping out computers in the network, the administrative privilege structure, and the most vulnerable, and valuable targets [16]. Over time, low-hanging fruit gets picked and hackers must expend more time and effort to find vulnerabilities.[3] Indeed, HackerOne has found that bounty payments increase as a program ages, though it has not established that this uptick is sufficient to prevent hacker attrition [12]. HackerOne's success as a business rests upon its ability to convince customers and investors that bug bounty programs still deliver value as they age.

Industry

There are a host of industry effects that might have an impact on the number of bugs that are reported. In certain industries, bugs might be endemic, because there is greater complexity. In addition, certain industries have trouble recruiting technical talent. For instance, the federal government has trouble hiring IT talent because of its low pay and restrictive drug policies. Consequently, it relies on outmoded technology more prone to cyber attacks [17]. Moreover, in certain industries, companies will internalize all of the costs of a vulnerability. In contrast, software products will likely be leveraged by other enterprises who will bear some of the costs of a vulnerability. This in turn may impact how responsive a company is to cybersecurity flaws.

Another potential industry effect is the ease with which criminals can monetize a vulnerability in a particular sector. The security researchers on HackerOne are frequently characterized as ethical hackers, who would never engage in criminal activity [18]. However, Alex Stamos, the former Chief Security Officer of Facebook, who paid out over $10 million in bounties during his tenure, suggests that this assumption might not be true [19]. Hackers may submit their reports to bug bounty programs when they believe they will enjoy rewards from vulnerable companies . For instance, social media company vulnerabilities most frequently lead to stolen login credentials, which are not valuable on the dark web. Social media companies, which care about safeguarding their reputation, will likely remunerate hackers more generously than the black market. In contrast, hacks in the financial industry often lead to stolen bank accounts; it is easy to maliciously monetize bugs. Social media companies thus may receive more bug bounty reports, ceteris paribus, than their financial counterparts.

Brand profile

Hacking well-known companies, including, on the HackerOne and BugCrowd platforms, GM, Google, Uber, Tesla, Starbucks, Goldman Sachs, and HP will generate more publicity. Hackers want publicity for a couple of reasons. First, the hacking community has its own distinctive culture and hierarchy, which members are striving to climb [20]. Second, finding bugs in prominent companies might impress potential full-time employers [21]. Third, a small subset of bug bounty hackers have won acclamation as security experts and have built lucrative careers consulting for companies and appearing in the media. The desire for publicity is evidenced by the fact that companies receive a dramatic uptick in hacker engagement when they start publicly reporting vulnerabilities that are found on HackerOne [12]. The bugs found from well-known companies will naturally receive outsized attention.

Bounty amount

In a perfectly competitive market, we would expect the prevailing market price to be the marginal cost of producing a good. Bug bounties deviate from perfectly competitive markets because, inter alia, it is not immediately obvious how exclusive a bug is [22]. Some programs may compensate hackers more generously than others. Hackers would devote more time to the programs which offer the most generous compensation.

Time to resolution

It is possible that two hackers come to the same vulnerability independently; this is known as a collision [23]. The code base is only updated once a vulnerability is resolved and, in the event of a collision, only the first hacker to discover a bug is compensated. If a company has slower resolution time, it is more likely that a hacker expends immense effort only to be uncompensated because he or she submitted a duplicate. Fast resolution times are important because hackers operate largely autonomously; there is little—or no—coordination among hackers to avoid duplication. It is likely that quicker resolution times will become more important as the community of security researchers participating in bug bounties grows and the chances of a collision increase.

Revenue

The literature is equivocal on the impact of company revenue on the number of bugs that are found in a given month. On the one hand, bigger companies tend to have the most complicated technology assets [19]; complexity increases the risk of cyber attacks. On the

3 This is mitigated somewhat by the fact that there are constant software updates which lead to new vulnerabilities.

other hand, companies with more resources can hire more talented software engineers, implement more cybersecurity controls, and subject their products to more robust quality assurance processes than smaller companies [24]. Additionally, there is a strain of thought, most passionately championed by security researcher Katie Moussouris, that bug bounties are only useful for large companies. Her rationale is that bugs are emblematic of larger flaws in process; only large companies have the resources to diagnose these flaws [25]. If her assertion is accurate, we would expect to see that revenue has a negative relationship with bugs reported, because every bug reported to a bigger company catalyzes a broader improvement in cybersecurity that does not occur for smaller companies. Finally, an important distinction is that smaller companies are more reliant on third-party software products, while larger companies build more software in-house [26].

Scope

The computer science literature is consistent in its conclusion that bugs increase as the number of lines of code increase, though their exact models differ (see, e.g., [27–30]). As scope increases, the number of vulnerabilities found should as well.

New programs

Bug bounties have increased in popularity and, as the Gartner report suggests, will become much more common going forward; 50% of companies will employ crowdsourced cybersecurity by 2022 [5]. There are two potential effects of new programs. Traditional economic theory suggests that competition will cannibalize hacker reports to existing programs. Alternatively, it is possible that new programs will generate positive network effects [31, 32], increasing the quantity of hackers and the amount of time hackers spend on the platform; this would cause an uptick in the number of reports all programs receive. Determining an answer will be important to assessing bug bounty markets as a whole. The top 7% of hackers accounting for nearly 40% of valid vulnerability submissions [33]. If new programs dramatically increase competition, it may trigger an arms race to lure adept security researchers. Deep-pocketed companies may lure the best talent, leaving SMEs in the lurch.

Private vs. Public programs

Public bug bounty programs enable any HackerOne user to submit vulnerabilities. Private bug bounty programs require an invitation; hackers are selected to participate if they possess pertinent skills for a particular company, or if they meet certain other criteria (e.g., in order to hack a Pentagon program, a security researcher must be based in the USA). On average, public programs engage 3.5 times more hackers than private programs [2]. However, public programs receive an order of magnitude more invalid reports. Sifting through the torrent of public program reports is costly. When figures were last reported, 79% of programs were private and 21% were public; the number of public programs has been steadily rising [12].

Past empirical work

A limited number of studies have examined how much bug bounty programs cost and how effectively they shore up security.

Maillart *et al.* [21] analyzed HackerOne's public program data. They find that the number of bugs a program receives is super-linearly related to the number of hackers they have enrolled. This supports Linus's Law's argument that more security professionals help firms unearth more bugs. They also run an ordinary least squares (OLS) regression to investigate the impact that new programs have on previous programs; they find that new programs reduce the number of reports that older programs receive. However, this regression suffers from potential bias in the form of reverse causality: new programs could sign up because existing programs were receiving a lot of reports. Moreover, private programs may be more impervious to the effect of new programs than public programs because they have a more stable corps of hackers.

Zhao *et al.* [34] analyze public data from HackerOne and Wooyun, a now-defunct Chinese bug bounty platform. They find that many programs receive fewer vulnerability reports over time and that the size of monetary incentives is positively correlated with the number of reports a program receives.

Last year, an analysis of HackerOne and Bugcrowd public programs found that on average, the annual cost of bug bounty programs is $85 000, less than the cost of hiring two in-house software engineers in the UK [35]. Moreover, the average program uncovers 156 unique vulnerabilities. However, their price estimates did not include the costs of HackerOne and Bugcrowd subscriptions or program management; it also excluded the internal costs of sifting through vulnerability reports.

Our article is an important addition to the literature because it addresses many of the shortcomings of previous papers. Our data set is far more comprehensive than those used by other researchers: it includes both public and private program data. Our instrumental variable strategy enables us to more convincingly establish causality. Finally, for the first time, in academic literature, we calculate the price elasticity of hackers.

Methodology

Summary of data

We leverage HackerOne data from August of 2014 to January of 2020. Our dataset contains observations from all programs that:

- Started out as private (most firms initially are private and become public if they are satisfied with bug bounties and want to increase the number of vulnerability reports they receive).
- Started no later than May 2019.
- Subscribe to HackerOne program management.
- Offer cash bounties to hackers.

In total during this time period, these programs have collectively received over 50 000 valid reports. This is a comprehensive database of the largest bug bounty platform's clients, meaning that the conclusions of this report are externally valid for the current bug bounty market. (After all, the major bug bounty platforms compete with each other to sign clients and recruit hackers.) The sample companies are big, medium, and small and are in industries as diverse as aeronautics, financial services, and retail. We present summary statistics of the data in Table 1. For our narrative instruments, described in the subsequent subsection, we rely on the Privacy Rights Clearinghouse database, the most comprehensive publicly available record of data breaches. In generating our variable for

Table 1: Summary statistics of dataset

Variable	Median	Standard Deviation
Revenue (in thousands of dollars)	1 80 000	9 999 293
Time to Resolution (in days)	52.15	97.52
Bounty Amount (in USD)	366.70	1249.79
Twitter Followers	44 978	2 395 441

data breaches over the last 9 months, we used a subset of the database that comprised of 8769 breaches and over 10.1 billion pilfered records.

Identification strategy

We first posit the following OLS regression model:

$$\begin{aligned} Y_{ij} &= \beta_0 + \beta_1 * NewPrograms_j + \beta_2 * Finance_i + \beta_3 * Retail_i \\ &+ \beta_4 * Medicine_i + \beta_5 * Government_i \\ &+ \beta_6 * AverageTimetoResolution_{ij} + \beta_7 * Revenue_i \\ &+ \beta_8 * BrandProfile_i + \beta_9 * LogBounty + \beta_{10} * ProgramAge_{ij} \end{aligned} \quad (1)$$

Here, the dependent variable of interest is the number of valid vulnerabilities submitted to program i in month j. NewPrograms refers to the number of programs started across the HackerOne platform in month j. If the β^1 coefficient is positive, that would indicate that positive network effects are stronger than competition effects; if the β^1 coefficient is negative, this would suggest that competition effects outstrip positive network effects. We include a series of industry dummies intended to capture industry effects. All companies in our sample are categorized as being in the finance, medicine, retail, government, or other industries. We omit other industries from the regression to avoid perfect multicolinearity. The dummy variables take on the value 1 if a company is in an industry and 0 if it is not.

AverageTimetoResolution represents the mean resolution time for a program in days for all bugs reported in month j. Revenue measures a company's annual revenue, in thousands of dollars. HackerOne receives this data from DataFox, a business intelligence software owned by Oracle. DataFox receives revenue estimates from a third party. However, they provide a snapshot of revenues, not a time series; the revenue estimate used in the regression is from January 2020—the end of the time series.[4]

We estimate BrandProfile through a proxy. Data Fox crawls the internet to derive estimates of a company's web traffic and Twitter followers, both in thousands. Once again, these are snapshots and will be inaccurate for the entire panel duration of five-and-a-half years. We believe that Twitter followers is a more accurate proxy for brand profile than web traffic. Web traffic is a distorted measure because technology companies that sell goods solely online should have more website hits than comparable or potentially even higher-profile companies who conduct commerce offline. Consider the cases of Walmart and eBay. Walmart is a better-known brand, but because it conducts most of its business in brick-and-mortar stores, it has less web traffic than eBay, an online auction market. While there might still be shortcomings in using Twitter followers as a metric, we believe it is a better measure. In this paper, we initially rely upon Twitter followers for brand profile and then conduct a robustness check, using web traffic as a proxy for BrandProfile. Program age measures how many months program i has operated in month j.

The trouble with this model is that there is potential for endogeneity. Time to resolution suffers from simultaneity: if a company receives a lot of valid reports, it will overwhelm their internal security team and their average resolution time will naturally go up. To address this endogeneity, we regressed valid reports per month by time to resolution lagged by 3 months. The new accounts variable, as previously described, also suffers from potential simultaneity. If HackerOne has been very successful over the last several months, generating a high volume of valid reports and improving its clients' cybersecurity postures, then more accounts would sign on. The way to combat this endogeneity is by using a narrative instrument, the number of publicly reported breaches that had occurred in the USA from the period of month *j-9* to month *j*, as measured by the Privacy Rights Clearinghouse. If cybersecurity has dominated the news cycle, it may push executive teams and corporate boards to approve an expenditure in HackerOne. Of course, not all companies in the HackerOne dataset are from the USA; however, the bulk of them are and the Privacy Rights Clearinghouse database is the most complete publicly available source of its kind. We questioned whether the data base was endogenous: if some of these breaches were the result of vulnerabilities in third party programs that were also used by HackerOne clients, this would violate the exclusion condition. However, corporate software literature indicates that third-party software vulnerabilities stem from improper configuration, suggesting they are idiosyncratic [36]. Therefore, we adjudged this to be the best instrumental variable.[5] We transition from an OLS model to a two stage least squares regression (2SLS) model:

$$\begin{aligned} Y_{ij} &= \beta_0 + \beta_1 * NewPrograms_j + \beta_2 * Finance_i + \beta_3 * Retail_i \\ &+ \beta_4 * Medicine_i + \beta_5 * Government_i \\ &+ \beta_6 * AverageTimetoResolution_{ij} + \beta_7 * Revenue_i \\ &+ \beta_8 * BrandProfile_i + \beta_9 * LogBounty_{ij} + \beta_{10} * ProgramAge_{ij} \end{aligned} \quad (2)$$

$$NewPrograms_j = \beta_0 + \beta_1 * BreachesPast9Months \quad (3)$$

$$\begin{aligned} &AverageTimetoResoultion_{ij} \\ &= \beta_0 + \beta_1 * AverageTimetoResolution_{i,j-3} \end{aligned} \quad (4)$$

However, there still is likely lingering endogeneity. Programs that are older in the panel signed up on HackerOne when bug bounties were in their nascency. These companies are likely systemically different from companies that adopted HackerOne at a later date. Similarly, programs with quicker resolution times, even if they are lagged by 3 months, either have more skilled security teams who can quickly parse through reports, or are more attuned to cybersecurity and invest more in resolving reports. To control for these differences, we run a fixed effects regression via a least square dummy variable estimator (LSDV). A fixed effects regression drops all time-invariant variables which in this regression are the industry dummies, revenue, and brand profile.

We also ran a number of robustness checks beyond substituting in web traffic for Twitter followers. We ran the regression with median bounty as opposed to log median bounty; and included a private program dummy in the regression, which took on the value of 1 if a program was public and 0 if it was private.

This methodology removes many of the sources of endogeneity that have plagued previous empirical research on bug bounties. However, there are still potential sources of omitted variable bias;

4 If possible, we would have liked to use lagged variables of revenue from before the start of the time series—as these would be least likely to be endogenous. However, the DataFox estimates were all that were available to us.

5 Another potential narrative instrument with less of a fear of endogeneity would be the number of mentions of cybersecurity in *Agenda*, the most widely circulated trade magazine for corporate board members. This would be an excellent indication of the issues that are dominating discourse among the stewards of corporate governance. Unfortunately, we were not able to partner with them to get easy access to their data.

Table 2: Primary results: regression with Twitter followers

	Regression methods		
	OLS	2SLS	2SLS + FE
Constant	3.326**	5.402*	14.463*
	(1.50)	(2.93)	(8.58)
Finance	−1.083	−2.348***	.
	(1.31)	(0.70)	.
Retail	−0.203	−1.422*	.
	(1.32)	(0.81)	.
Medicine	−3.136*	−4.561	.
	(1.48)	(4.011)	.
Government	−1.148	−1.232	.
	(1.30)	(2.91)	.
TimetoResolution	0.003*	−0.002	−0.035
	(0.001)	(0.01)	(0.03)
NewPrograms	−0.038**	−0.067	−0.148
	(0.01)	(0.08)	(0.16)
LogBountyAmount	0.780***	0.790***	0.678***
	(0.09)	(0.10)	(0.09)
ProgramAge	−0.022	−0.021	−0.128***
	(0.02)	(0.02)	(0.05)
Revenue	5.35e-08*	5.56e-08**	.
	(2.88e-08)	(2.74e-08)	.
TwitterFollowers	3.75e-07***	3.75e-07***	.
	(1.42e-07)	(1.12e-07)	.
Adjusted R^2	0.0386	0.0369	0.393
Wald value	.	96.90	2572.55
Root MSE	12.560	12.551	9.956

*** $p < 0.01$.
** $p < 0.05$.
* $p < 0.1$.

Table 3: Robustness check: regression with median bounty, not log median bounty

	Regression methods		
	OLS	2SLS	2SLS + FE
Constant	7.30***	9.449***	18.07**
	(0.71)	(2.89)	(8.68)
Finance	−2.205***	−2.321***	.
	(0.70)	(0.71)	.
Retail	−1.039	−1.023	.
	(0.81)	(0.82)	.
Medicine	−4.935	−5.458	.
	(4.00)	(4.03)	.
Government	−1.055	−1.398	−2.817
	(2.91)	(2.93)	(15.97)
TimetoResolution	0.003	−0.002	−0.039
	(2.26e-03)	(0.007)	(0.03)
NewPrograms	−0.034*	−0.056	−0.111
	(0.02)	(0.08)	(0.05)
BountyAmount	9.36e-04***	9.57e-04***	0.0003
	(1.87e-04)	(1.90e-04)	(2.11e-04)
ProgramAge	−0.019***	−0.023	−0.111
	(0.02)	(0.02)	(0.05)
Revenue	3.05e-08**	2.92e-08	.
	(2.84e-08)	(2.91e-08)	.
TwitterFollowers	4.20e-07***	4.30e-07***	.
	(1.12e-07)	(1.12e-07)	.
Adjusted R^2	0.0252	0.0262	0.373
Wald value	.	61.49	2444.40
Root MSE	12.63	12.62	9.546

*** $p < 0.01$,
** $p < 0.05$,
* $p < 0.1$.

several pertinent variables are inconsistently and unreliably measured. First, our model does not incorporate the effect of scope. Second, this model does not measure bug severity: the potential monetary cost an enterprise would incur if a bug were maliciously exploited. Bug severity would likely be correlated with a number of variables in our model including bounty amount, time to resolution (companies resolve acute threats more quickly), industry (in certain industries, severe bugs are likely more prevalent), and program age (it often takes time for hackers to find severe bugs).

Results

Examining our primary results in Table 2, we see that the 2SLS regressions comfortably pass the Wald test, meeting the first requirement of instrumental variables: the relevancy condition. We also ran a Durban–Hausman–Wu test of exogeneity for both of the 2SLS regressions; while we were not able to reject the null hypothesis in both cases at the 5% significance level, we were able to at the 8% level. These findings are consistent with our ex-ante assumption that OLS is endogenous and 2SLS is exogenous, but since it is a group hypothesis test, it is not definitive. Exploring our robustness test results, conveyed in Tables 3-5, we conclude that Table 2 is the best specification for the model.

Key findings and discussion

Hacker supply is price inelastic

The log average bounties variable reveals a price elasticity of between 0.1 and 0.2 at the median for hackers. It is likely we underestimate hacker elasticity, because our regression does not control for the effects of bug severity. Still, our results indicate that hackers are largely motivated by non-monetary factors: some may want to gain experience and reputation, while others may have altruistic motivations. This is positive news for SMEs, who often lack the resources to offer generous bounties.

HackerOne program managers suggested that new, unseasoned hackers are most inelastic: eager to gain exposure, they show little price sensitivity. In contrast, the best hackers, who have significant opportunities, are more price elastic. If this is the case, Google's program, which pays established hackers a high retainer for regularly participating in their programs and then offer additional performance-based bonuses, may very well shave off the best hackers. Jack Cable, a top security researcher on the HackerOne platform, said that Google's model was attractive to him and more companies—including some companies on the HackerOne and BugCrowd platforms—are adopting it [37]. This could diminish the value of bug bounties for companies with less resources. But Mårten Mickos disagrees with this line of thinking. He notes that hackers take at least 2–3 years to earn a high-caliber reputation. Thus, companies with less resources are able to draw upon the talents of "up and comers" [38]. It would be instructive to conduct research on the effects of new compensation structures.

Brand profile and revenue have an economically insignificant impact on reports companies receive

We look to the 2SLS without program fixed effects to examine the impact that company revenue and brand profile have. Our estimate on the coefficient for revenue is positive and statistically significant,

Table 4: Robustness check: regression with web traffic, not twitter followers

	Regression methods		
	OLS	2SLS	
Constant	4.468***	5.568**	
	(0.86)	(2.93)	
Finance	−2.651***	−2.673***.	
	(0.70)	(0.70))	
Retail	−1.343*	−1.297	
	(0.80)	(0.81))	
Medicine	−4.463***	−4.723	
	(3.99)	(4.02))	
Government	−1.337	−1.418	
	(2.90)	(2.91))	
TimetoResolution	0.003	−6.42e-04	
	(0.002)	(7.31e-03)	
NewPrograms	−0.042**	−0.066	
	(0.02)	(0.08)	
LogBountyAmount	0.784***	0.790***	
	(0.10)	(0.10)	
LogBountyAmount	0.784***	0.790***	
	(0.10)	(0.10)	
Revenue	9.96e-08***	1.03e-07***	
	(2.67e-08)	(2.73e-08)	
ProgramAge	−0.019	−0.018	
	(0.02)	(0.02)	(0.05)
WebTraffic	1.44e-11	1.19e-11	
	(3.11e-11)	(3.14e-11)	
Adjusted R^2	0.0324	0.0338	
Wald value	.	81.55	
Root MSE	12.581	12.571	

***$p < 0.01$.
**$p < 0.05$.
*$p < 0.1$.

Table 5: Robustness check: tier regression with dummy private program variable

	Regression methods		
	OLS	2SLS	2SLS + FE
Constant	3.591***	4.400	1.113
	(0.90)	(3.24)	(9.64)
Finance	−2.091***	−2.12***	.
	(0.71)	(0.71)	.
Retail	−1.462**	−1.427*	.
	(0.80)	(0.80)	.
Medicine	−3.784	−4.005	.
	(3.99)	(4.04)	.
Government	−0.671	−0.765	.
	(2.90)	(2.91)	.
TimetoResolution	0.004	1.21e-03	0.007
	(0.020)	(7.81e-03)	(0.04)
NewPrograms	−0.065	−0.053	−0.035
	(0.04)	(0.08)	(0.16)
LogBountyAmount	0.784***	0.787***	0.549***
	(0.10)	(0.10)	(0.10)
Revenue	5.56e-08**	5.69e-08**	.
	(2.68e-08)	(2.72e-08)	.
TwitterFollowers	3.28e-07***	3.32e-07***	.
	(1.13e-07)	(1.13e-07)	.
ProgramAge	−0.037***	−0.027	−0.128***
	(0.02)	(0.02)	(0.05)
PrivateProgramDummy	1.157**	1.055*	5.569
	(0.46)	(0.57)	(7.18)
R^2	0.0370	0.0393	0.450
Wald value	.	100.10	2766.92
Root MSE	12.552	12.535	9.483

***$p < 0.01$.
**$p < 0.05$.
*$p < 0.1$.

but it is not economically significant. At the point estimate, a company in the 75th percentile of revenue in the data set would only receive about 0.05 more valid reports per month than a company in the 25th percentile, ceteris paribus. We found similar results for our proxy for brand profile, Twitter followers. While our coefficient estimate was once again positive and statistically significant, it was not economically significant. A program in the 75th percentile of Twitter followers receives about 0.09 more valid reports, ceteris paribus, than a program in the 25th percentile.

These results suggest that bug bounties are effective for companies of all sizes and all levels of prominence. This is particularly positive for SMEs, which often lack the cachet and resources to recruit in-demand cybersecurity professionals. Much like in other fields, this segment of the gig economy seems to democratize access to IT talent. Moreover, our findings do not provide support for Katie Moussouris's theory that bigger companies tend to learn more from bugs creating a virtuous cycle that lowers future bug flow, since we found a positive coefficient on revenue. However, these findings offer snapshots of revenue and brand profile; it is possible that a company's revenue and brand profile could have changed over the 5.5 year panel. In the future, it would be useful to conduct research with more accurate revenue and brand profile estimates.

Industry effects

To explore industry effects, we again examine the 2SLS without program fixed effects. Our findings suggest that bug bounties are effective for companies in a host of industries. However, we find that companies in the financial and retail industries receive 2.34 and 1.42 fewer valid reports per month than companies in the other category, ceteris paribus. These estimates are statistically significant at the 8.2% significance level. We also find that medical companies receive 4.6 fewer reports than companies in other industries; however, this negative coefficient is not statistically significant, likely because of multicolinearity. These measures are economically significant given that the median program receives four valid reports per month.

These numbers are consistent with Alex Stamos's theory that hackers are driven in part by the opportunity cost of reporting a vulnerability. Vulnerabilities in the finance industry can more easily be maliciously monetized, disincentivizing hackers from reporting them on HackerOne. Furthermore, according to a joint study by IBM and the Ponemon Institute, healthcare data records are the most monetarily valuable, because they contain personally intimate details [39]; security researchers may be more inclined to sell them on the black market. Finally, because these companies face greater costs in the event of an attack, they may more actively seek to preempt breaches. For example, the finance industry has implemented cybersecurity best practices working in conjunction with the New York Department of Finance. It may thus be harder for HackerOne gig workers to find bugs.

The number of new programs has a statistically insignificant effect on reports companies receive

We did not find evidence that new programs dramatically dampen the number of reports companies on the HackerOne platform receive. This suggests that as competition for hacker time has increased, HackerOne has been able to recruit more hackers and convince hackers to spend more time on the platform. This trend may augur well for HackerOne, which anticipates significant growth over the next several years. However, we should be cautious about extrapolating into the future. If it grows large enough, HackerOne and its rivals may exhaust the talent pool of security researchers and new programs may have a far more discernable effect on the reports all companies receive. HackerOne CEO Mårten Mickos noted that this is a challenge that the company is actively trying to confront by, inter alia, matching security researchers by skillsets to the programs where they are most likely to find vulnerabilities and working to improve the quality of gig workers by rolling out white hat hacking education modules [38].

Programs receive fewer reports over time

Programs at the 75th percentile in program age receive 2.56 fewer reports per month than programs in the 25th percentile in program age, ceteris paribus. This is significant because programs received a median of four reports per month. Furthermore, this estimate may be depressed because of omitted variable bias: we do not include a program's scope, which may increase over time and which would introduce new bugs for hackers to find.

This provides evidence that if programs do not increase their bounties as they grow older, they will receive significantly fewer reports. Still, old programs receive valid reports, even if vulnerabilities become harder to exhume.

Most Variation in program reports remains unexplained

The 2SLS regression without fixed effects has an R^2 value of 0.037. This means that it fails to explain more than 96% of the variation programs receive in valid reports each month. However, the 2SLS with program fixed effects has a much higher R^2 value of 0.393. This indicates that there are other unmeasured and constant differences between programs that explain a significant the variation in valid reports between programs. These factors seem to be uncorrelated with measures of revenue, brand profile, and industry; otherwise, they would appear in the standard 2SLS coefficient estimates through omitted variable bias.

Moreover, in our fixed effects regression, more than 60% of the variation in valid reports between programs remains unexplained. The effects of variables we were not able to include in our model—scope and bug severity—may explain some of this variation. It is also possible that some heretofore unidentified variables impact the flow of vulnerabilities.

Conclusion

Drawing upon a comprehensive data set of bug bounty programs, we were able to remove many of the sources of endogeneity plaguing research in the crowdsourced cybersecurity field and identify many of the factors which influence the quantity of valid vulnerability reports bug bounty programs receive.

Our research had six significant findings. For the first time in academic literature, we calculated an elasticity of hacker supply. Hackers are relatively price insensitive, with an elasticity of between 0.1 and 0.2 at the median. Second, we found that bug bounties are effective tools for companies of all sizes and levels of prominence. Third, we found that companies in certain industries received fewer reports, ceteris paribus, than companies in other industries. Fourth, we found that the number of new programs created in any given month has a marginal—and statistically insignificant—impact on the number of reports companies receive on the HackerOne platform in that month. Fifth, we found that programs receive fewer valid reports over time, all else remaining constant.

Finally, this article emphasized how little we know about the bug bounty markets. We failed to specifically identify most of the time-invariant variables which impacted hacker supply. And our final fixed effects regression model explained less than half of the variation we observed between data points. Future research should focus on identifying and measuring more of the variables which determine hacker supply. Subsequent research will elucidate how bug bounty markets work, sharpening our understanding of an increasingly important cybersecurity tool.

Acknowledgements

We would like to thank the Stanford University Department of Economics, which guided Kiran in the first version of this paper, which was his undergraduate thesis. Michael Boskin, Professor of Economics at Stanford, was integral as an advisor, pushing us to build a model that controlled for confounding variables. Douglas Bernheim, Chair of the Stanford Department of Economics, and Marcelo Clerici-Arias, Lecturer in Economics, also reviewed the article and Levi Boxell, PhD student in Economics, provided valuable tips on data analysis. We would also like to thank HackerOne for embracing this project and guiding it through completion. HackerOne CEO Mårten Mickos consistently supported the project and provided us with a host of introductions to experts in the field. HackerOne's former VP of Policy, Debby Chang, and CTO, Alex Rice, reviewed the article and offered helpful suggestions.

Funding

This work was supported by HackerOne. We had freedom to report all results which did not violate HackerOne's terms of service by revealing the identities of companies or hackers.

Conflict of interest statement. Kiran Sridhar was formerly employed by HackerOne. Ming Ng is currently employed by HackerOne. HackerOne provided us with full academic freedom to report the results of our investigation.

References

1. HackerOne. What is a vulnerability disclosure policy and why you need one. 2018.
2. HackerOne. Hacker Powered Security Report. 2018.
3. Department of Homeland Security. Binding operational directive 20-01: develop and publish a vulnerability disclosure policy. 2019.
4. Gardner D. Emerging technology analysis: Bug bounties and crowdsourced security testing. Gartner Research. 2018.
5. Pieters W, Hadžiosmanović D, Dechesne F. Security-by-experiment: lessons from responsible development in cyberspace. *Scie Eng Ethics* 2016; **22**:831–850. CrossRef] 10.1007/s11948-015-9648-y]
6. Ramey V. Identifying government spending shocks: it's all in the timing. *Quart J Econ* 2011;**126**:1–50.
7. Sargan JD. The estimation of economic relationships using instrumental variables. *Econometrica J Econ Soc* 1958;**26**:393–415.
8. Fair RC. The estimation of simultaneous equation models with lagged endogenous variables and first order serially correlated errors. *Econometrica* 1970;**38**:507–16.
9. ISC2. Cybersecurity workforce study. 2019.
10. Dimon J. CEO letter to shareholders. JP Morgan Chase. 2019.

11. Verizon Enterprise Solutions. Data breach investigations report. 2018.
12. HackerOne. The 2020 Hacker Report. 2020.
13. Raymond E. The cathedral and the bazaar. *Knowledge Technol Policy* 1999;**12**:23–49.
14. Brady R, Anderson R, Ball R. *Murphy's Law, the Fitness of Evolving Species, and the Limits of Software Reliability*. Computer Laboratory, University of Cambridge, 1999.
15. Fonseca J, Vieira M, Madeira H. The web attacker perspective-a field study. In 2010 IEEE 21st International Symposium on Software Reliability Engineering, 299-308. IEEE. 2010.
16. Bohme R, Moore T. The "iterated weakest link" model of adaptive security investment. *J Informat Security* 2016;**07**:81–103.
17. Corrigan J. The government's struggle to hire tech talent is worse than you thought. *NextGov*. 2017.
18. Harper A, Harris S, Ness J *et al. Gray Hat Hacking: The Ethical Hacker's Handbook*. McGraw-Hill Osborne Media, 2011.
19. Interview with Alex Stamos, former Chief Security Officer of Facebook. Conducted by Kiran Sridhar.
20. Coleman EG, Golub A. Hacker practice: moral genres and the cultural articulation of liberalism. *Anthropol Theory* 2008;**8**:255–277.
21. Maillart T, Zhao M, Grossklags J *et al.* Given enough eyeballs, all bugs are shallow? Revisiting Eric Raymond with bug bounty programs. *J Cybersecurity* 2017;**3**:81–90.
22. McKinney D. Vulnerability bazaar. *IEEE Security Privacy* 2007;**5**: 69–73.
23. Ablon L, Libicki M. Hacker's bazaar: the markets for cybercrime tools and stolen data. *Def Counsel J* 2015;**82**:143–52.
24. Jones C. *Software Engineering Best Practices: Lessons from Successful Projects in the Top Companies*. New York, NY: McGraw-Hill, 2010.
25. Markus ML, Tanis C. *The Enterprise Systems Experience—from Adoption to Success. Framing the Domains of IT Research: Glimpsing the Future through the Past*. 2000.
26. Shahzad B, Abdullatif AM, Ikram N *et al*. Build software or buy: a study on developing large scale software. *IEEE Access* 2017;**5**: 24262–74.
27. Ostrand TJ, Weyuker EJ, Bell RM. Predicting the location and number of faults in large software systems. *IEEE Trans Software Eng* 2005;**31**: 340–55.
28. Engler D, Chen DY, Hallem S *et al*. Bugs as deviant behavior: a general approach to inferring errors in systems code. *ACM SIGOPS Operating Systems Rev* 2001;**35**:57–72.
29. Bessey A, Block K, Chelf B *et al*. A few billion lines of code later: using static analysis to find bugs in the real world. *Commun ACM* 2010;**53**:66–75.
30. Concas G, Marchesi M, Murgia A *et al.* On the distribution of bugs in the eclipse system. *IEEE Trans Software Eng* 2011;**37**:872–77.
31. Baake P, Boom A. Vertical product differentiation, network externalities, and compatibility decisions. *Int J Industr Organ* 2001;**19**:267–84.
32. Kornish LJ. Technology choice and timing with positive network effects. *Euro J Operat Res* 2006;**173**:268–82.
33. Ellis R, Huang K, Siegel M *et al*. Fixing a hole: the labor market for bugs. In: Alex Pentland, Howard Shrobe, and David Shrier (eds), *New Solutions for Cybersecurity*. MIT Press, 2017.
34. Zhao M, Grossklags J, Liu P. An empirical study of web vulnerability discovery ecosystems. In: *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, 1105–1117. 2015.
35. Walshe T, Simpson A. An empirical study of Bug Bounty Programs. In: *2020 IEEE 2nd International Workshop on Intelligent Bug Fixing (IBF)*, 35–44. 2020.
36. Leon A. *Software Configuration Management Handbook*. Artech House. 2015.
37. Interview with Jack Cable, HackerOne hacker. Conducted by Kiran Sridhar.
38. Interview with Mårten Mickos, HackerOne CEO. Conducted by Kiran Sridhar.
39. IBM and Ponemon Institute. Cost of data breach study. 2018.