

# **Optimal design of Incentive Mechanisms: Badges and their role in open source Innovation**

Timothy Toth

Pace University

Master's Thesis: Economics

Dr. Mary Kaltenberg

November 25, 2024

## **Abstract:**

This paper examines the impact of non-monetary incentives, specifically badges, on open-source software innovation through an analysis of GitHub's Badges Program. Using a panel dataset of over 49,000 GitHub users from May 2022 to June 2023, I document two main effects. First, badge earners shifted their behavior post-launch, increasing focus on original repository creation but reducing collaborative activities like pull requests. Second, specific badges currently in use showed no significant impact on contributions across key activities. These findings suggest that while badges influence behavior, their design needs refinement to foster meaningful contributions and engagement in open-source communities.

## 1 Introduction

Numerous studies have examined the effectiveness of incentive programs in stimulating innovation, ranging from financial rewards to government contracts (Boudreau et al., 2011; Daim et al., 2022;). Prior research in the open-source setting has mostly concentrated on the financial incentives that drive contributions, paying less attention to the ways in which more recent incentive mechanisms, such achievements or community-based recognitions, affect contributor behavior. The goal of the GitHub Badges Program was designed to recognize and motivate contributors through non-monetary rewards. Badges celebrate and showcase your journey on GitHub (Brooks, 2022).

Open source software (OSS) has played a major role in modernized times for increasing innovation efforts throughout all industries (OpenLogic, 2023; GitHub 2023; Conti et al., 2021). Open Source Software has been fueling advancements in fields like machine learning, cloud computing, and artificial intelligence by providing freely accessible tools, frameworks, and collaborative platforms. The widespread use and progress of machine learning technologies heavily depend on open-source resources like python libraries, stack overflow, and GitHub. This growing dependence on open source for machine learning suggests a potential shift in the structural dynamics of innovation creation and sourcing, with GitHub emerging as a pivotal platform driving this growth and collaboration.

While the current understanding seems incomplete for innovation incentives in open-source environments, recent work has started to make itself more known. Existing work has emphasized the role of intrinsic motivations (Bitzer et al., 2004; Hausberg et al, 2020), alongside a recent movement in monetary rewards as seen through the GitHub Sponsorship Program and others (Conti et al., 2023; Krishnamurthy et al., 2013).

Incentive mechanisms are fundamental in shaping behavior, particularly in environments that rely on voluntary contributions like open-source software. Financial incentives such as monetary rewards or sponsorships directly compensate individuals for their efforts and are often tied to specific outputs or goals. These incentives are effective in aligning contributors' activities with organizational or project priorities but can crowd out intrinsic motivations, such as after receiving monetary compensation the desire to innovate or collaborate decreases in the long term(Conti, 2022).

Non-monetary incentives by contrast, target social and psychological drivers such as recognition, status, or skill development. Badges as implemented in the GitHub Badges program, represent a non-monetary incentive that leverages gamification and community recognition. Gamification introduces game-like elements, such as levels or rewards into non-game contexts to increase engagement. By earning badges contributors receive visible acknowledgment of their efforts which can signal expertise or commitment to peers or potential employers. This visibility helps establish reputation and status within the OSS community.

The use of non-financial incentives extends beyond OSS. Platforms like Stack Overflow use badges to encourage users to answer questions, edit posts, and contribute to community knowledge. Similarly, gamified education platforms, such as Duolingo award badges to learners for achieving milestones. Even when looking at fitness apps such as Strava they employ achievements and leaderboards to motivate users to track their progress and compete with peers.

To the best of my knowledge, this paper is among the first to evaluate the impact of community recognition-based incentives on innovative output in open source. Incentive mechanisms such as badges, have been shown to enhance user motivation by leveraging intrinsic motivators like recognition, mastery, and enjoyment. By incorporating these features, GitHub

can encourage oss contributions. I examine this by studying a program targeted at motivating users to develop open-source contributions: the GitHub Badges Program launched in July 2022. This program enabled users to display badges on their profile by unlocking them essentially playing to earn. GitHub kept the number of achievements and process of obtaining them discrete. It was up to the user to maintain their daily routines and unlock by creating through GitHub. Eventually through open source contributors, a public repository was created that lists all the achievements and their respective definitions to earn them. By using information from 4083 global users with achievements on their GitHub profiles and a corresponding group of 45218 randomly chosen users without achievements unlocked, I gathered comprehensive records of each user's monthly activity on the platform, including the repositories created and forked, issues opened, commits, and pull requests. The resulting panel dataset spans from May 2022 to June 2023 alongside the baseline June 2021 month year.

This allowed me to document two main relationships regarding the efficacy of the badges program in incentivizing open-source innovation. While documenting these relationships, it is important to recognize potential challenges associated with such incentive mechanisms. Overemphasis on achievements might inadvertently prioritize quantity over quality in contributions, that is why it is important for GitHub to optimize the design of badges to keep the user engaged and innovating not only in the short term but over the long term.

To do this I implemented inverse probability weighting and a difference-in-differences approach, I estimate the extent to which the launch of the GitHub Badges program that allows individuals to receive community recognition – i.e., to unlock achievements to put onto your GitHub profile – influenced their open-source innovation.

The remainder of this paper is structured as follows. In Section II related literature is discussed, in Section III I describe the context and data. Sections IV and V present the empirical specification and the results, respectively. In Section VI potential interpretations of the findings as well as contributions to the literature and conclude.

## 2 Literature Review

In economics, incentive mechanisms are essential because they motivate actions that lead to the desired results, innovation. Conventional economic theories, like game theory and principle-agent theory have long investigated how incentives might be set up to maximize productivity and creativity (Gibbons, 2010., Prendergast, 2022., J. Liu et al., 2022). One of the biggest challenges in incentive mechanisms is keeping the user's intrinsic motivation to finish the activity or task. Extrinsic benefits, such as cash prizes or rewards, can boost output and innovation, but they run the danger of weakening the intrinsic motivation that encourages users to consistently work on open-source projects.

Effective badge design plays a critical role in ensuring the effort and participation on platforms like GitHub remains constant(Burke et al., 2009., Chawla et al., 2012., Deterding et al., 2011) . Easley and Ghosh (2016) highlight how badges serve as both motivational tools and strategic signals within user-generated content platforms. They provide a game-theoretic framework for understanding how badges balance individual effort and platform goals. They argue that badges derive utility from their perceived value rather than monetary compensation, motivating participants through recognition and mastery. By acting as signaling devices badges communicate contributors' skill and effort, fostering collaboration and competition within the community. The trade-offs between inclusivity and exclusivity in badge distribution are critical,

as overly inclusive designs may dilute badge value, while exclusive designs risk alienating less active users. This distinction aligns with GitHub goal to foster sustained innovation by balancing effort-based rewards with community driven recognition. Easley and Ghosh (2016) suggest that platforms can use badges to encourage long-term engagement by structuring them around both absolute standards (achievements tied to individual effort) and relative standards (achievements tied to community comparisons). This dual approach incentivizes contributors at different levels of experience and activity which is something GitHub is currently missing in their current badge design.

While incorporating badge design into the workflow, understanding where users allocate most of their time within open source software is important. The paper “Open Source Software policy in industry” Jeff Gortmaker examines how contributors allocate their time within open source ecosystems and highlighting the economic tradeoffs involved in their decisions. One of its findings is that pull request constitute a significant portion of time an expert level user- identified through matched GitHub and LinkedIn workplace data, emphasizing the labor-intensive nature of this activity and its role in driving project success, which relates to our work by now creating more within open source which translates to digital innovation.

The paper that is most relevant for us from the economics literature is Conti et al.[2022], which analyzes the GitHub Sponsorship Program and how eligibility and receiving a sponsorship effects open source innovate output. When looking at financial rewards the intrinsic motivation can be seen to have a crowding out effect. By providing monetary compensation, the sponsor may, unintentionally, signal that they are not confident the developer will be motivated enough to exert effort in the absence of added incentives (Conti et al., 2023). The findings from Conti et al. (2022) demonstrate that sponsored developers tend to create more repositories, contribute more

frequently, and address more issues compared to their non-sponsored counterparts. These results suggest that financial incentives provide a strong short-run boost in effort, as developers respond to the immediate utility derived from monetary rewards. But this effect diminishes over time. This pattern mirrors observations in other contexts, such as the decline in effort observed among athletes after receiving a long term deal or students after gaining admission to a target program. Effort is often maximized only until the primary reward is achieved, after which engagement declines. This diminishing effect of monetary rewards on effort can be explained by the concept of diminishing marginal utility. As contributors continue to receive financial rewards, the incremental satisfaction derived from each additional payment decreases, reducing their willingness to exert additional effort. Over time, developers may recalibrate their effort allocation, prioritizing other activities or projects where intrinsic motivations play a larger role or where the financial incentives are greater .

That is where this paper looks to explore. Will badges provide enough intrinsic motivation to consistently deliver results even after receiving their first badge? By understanding the impact of different types of incentives, my research aims to build on these findings and explore how achievement-based incentives, like badges, compare to financial ones in promoting innovation within open source.

### **3 Empirical Context and Data**

My objective is to examine how the GitHub Badges program affects several facets of user behavior on the site and to determine whether community-based recognition awards can influence users' involvement within open source.

#### **3.1 The GitHub Badges Program**

The GitHub Badges program was introduced as part of GitHub's broader mission to recognize and celebrate the contributions of developers to open-source projects. With nearly 5,000 public contributions being made every minute, the program aims to acknowledge the immense effort involved in building and maintaining the diverse projects and communities hosted on GitHub (GitHub, 2023). By showcasing developer milestones, the program not only rewards individual achievements but also underscores the significant impact of software on contemporary and future innovations

Through the badges feature, developers can receive badges for completing a variety of milestones, including their first few commits, pull requests, discussion answers, and even providing code to support large-scale initiatives like NASA's Mars helicopter flight. Developers' biographies feature these accomplishments, giving a visual account of their contributions and GitHub journey and generating reputational capital. Developers derive utility not only from the act of contributing but also from the reputational gains tied to their badges. The public nature of these rewards creates positive externalities that can be seen from as that the visible successes inspire others to engage with and contribute to the growth of open-source projects.

### **3.2 Data Construction**

To assess the effect of the Badges program, I compiled a dataset using the GitHub Archive database (Grigorik, 2024) and employed a custom web scraper I developed to collect data on individual badges and the dates users obtained their first badge.

I identified users who have earned badges in the GitHub Badges program and collected their usernames. Overall, there are 4083 users who have obtained achievements, forming the

treatment group. To construct the control group, I randomly sample non-badge users from the GitHub Archive database, which continuously scrapes and stores public activity on GitHub.

While the dataset offers detailed insights, it is limited to public activity which can potentially result in omitting contributions to private repositories or organizational work. The panel structure however enables fixed-effects regression models, allowing for the control of unobservable individual characteristic.

### **3.2 Summary Statistics**

Of the selected 4083 individual users who unlocked at least one achievement between June 2022 (the start date of the program) and June 2023 (exactly one year after the introduction of the GitHub Badges program), when compared to their baseline (June 2021) characteristics, on average saw a decrease of 6.51% of new repositories, 1.78% increase in total number of commits, and a decrease in issues by 4.94%.

In Table 1, I present descriptives for the total number of repositories, commits, and issues generated during the pre-sample period, which starts in February 2022 and ends in June 2022. For context, A repository (or ‘repo’) is a centralized storage location for a projects files and version history, saved in GitHub in a dedicated folder. Such repositories can be created from scratch, known as ‘original’ repositories, or they can be derived by copying and modifying an existing repository through a process called ‘forking,’ resulting in ‘forked’ repositories. In my paper, I argue that repositories are the closest thing to innovation because they require a substantial amount of coding work, expertise, and knowledge to create. Commits are a useful indicator of the amount of work people put into changing their code because they reflect a single set of changes made to files inside a repository. Conversely, issues entail the recording of issues

in a repository, promoting cooperation by allowing several contributors to address and resolve these concerns. Issues are therefore a helpful gauge of the contributions people make to open-source projects.

Table 1 shows that users with badges create 0.6 more repositories than users without badges, with the difference significant at the 5% level. When looking at original repositories show a larger, significant difference of -0.6 at the 1% level, suggesting badge users tend to focus more on original contributions compared to non-badge users.

Standard errors in parentheses show more variability among achievement users compared to users without achievements. This may indicate different engagement levels within this group. Mean differences in new commits (1.40) and new issues (-0.10) are not significant, suggesting achievements do not uniformly increase all outputs but more so original contributions.

## 4 Empirical Methodology

This paper aims to explore two key empirical questions: 1) Did the launch of the GitHub Badges program affect users' innovative output. 2) Which specific achievements had the greatest impact on the respective innovate output.

### 4.1 The impact of the Launch of the Badges Program on Users who Obtained

To assess the effect of obtaining GitHub achievements on users' innovative activities in open source, I compare the behaviors of users who eventually unlocked achievements before and after June 2022, to that of similar users who did not earn any achievements. As I do not observe the timing at which individuals obtained their first achievement, I treat this characteristic as a time-invariant user feature. This analysis aims at assessing whether users might increase their

innovation effort incentivized by the possibility of receiving a community-based recognition reward being badges on their GitHub profile.

To make the two samples more comparable, I implement the inverse probability weighting (IPW) (Imbens, 2000). I begin by estimating, for each observation, the probability of receiving treatment, that is, obtaining an achievement, through a probit model. Specifically, I model the probability of being treated as a function of measures for user activities on GitHub observed in the pre-sample period, during the months of February 2022, April 2022, June 2022, 3 months prior to start of the program. These measures are: the total number of repositories, issues, and commits created, the natural logarithms of these counts, and the square of the logarithms.

Having estimated the probit model, I predict each users probability  $p$  of obtaining an achievement. The users that obtain an achievement are then weighted by  $1/p$ , whereas those that did not obtain an achievement are weighted by  $1/(1-p)$ . Therefore, each observation is given weight as the inverse of the probability of their status. As a result of applying this approach, treated user that resemble the control group are given more weight, and controls that resemble the treatment group are also given more weight.

Using the generated weights, I estimate the following difference-in-differences Poisson model at the level of each user  $i$  observed in month-year  $t$ :

$$E(Y_{it}|X_{it}) = \gamma_i \exp(\beta \cdot \text{Badges} \cdot I[t > \text{June 2022}]),$$

Where  $Y_{it}$  is the number of repositories, issues, or commits, created by  $i$  at time  $t$ . Achievements, is an indicator equal to 1 if the user  $i$  participates in the Badges program and 0 if not. The coefficient of interest in the model is  $\beta$ . It represents the change in outcome after the Badges

program was introduced for treated users relative to untreated users. The  $\gamma_i$  are user fixed effects that absorb differences across users.

## 4.2 The impact of Specific Badges on Innovative Output

To analyze the impact of specific GitHub badges on user activities, I estimate a series of quasi-Poisson fixed-effects regression models. These models exploit variation in the timing of badge acquisition to identify their effects on key outcomes, such as the number of new repositories, issues, or pull requests created. Each regression is designed to assess the influence of individual achievements, accounting for user-specific fixed effects and controlling for activity levels across related metrics. The models include interaction terms between post-launch periods and badge acquisition, enabling an analysis of differential effects following the introduction of the GitHub Badges program. I estimate using the following model:

$$E(Y_{it}|X_{it}) = \gamma_i \exp(\beta \cdot \text{Specific_Badges} \cdot I[t > \text{June 2022}]),$$

In this equation,  $Y_{it}$  represents the outcome variable for user  $i$  at time  $t$ , such as the number of new repositories, issues, or pull requests created. The term  $\gamma_i$  captures user-specific fixed effects, accounting for time-invariant characteristics unique to each individual. This inclusion ensures that unobservable differences across users, which remain constant over time, do not bias the estimates.

## 5 Results

### 5.1 The Impact of the Badges Program Launch on Users with Achievements

I begin by examining the effect of earning achievements after the launch of the GitHub Badges program on user conditions. The results from estimating the Poisson model described by

Eq.(1) are presented in Table 2 and 3. Where standard errors are clustered at the user level. Each observation is weighted by the inverse propensity score of earning an achievement to balance the treatment and control groups.

In Table 2, column 1, I show that the introduction of the badges program lead to an increase of 0.073 in the rewarded users' hazard of creating a new repositories each month. That is, after GitHub launched the Badges program, users who earned at least one badge created 7.3% more repositories relative to non-badge users each month. If creating a repository reflects the contribution of a new open source project, this result indicates that users who have earned a badge produced higher levels of innovative output following the launch of the Badges program. In column 2, I find that the Badges program increased the hazard of creating a new original repository for badge users by 0.082, incurring a 8.2% increase in original repository output compared to non-badge users. Showing signals that original open source creation could be an area to focus on when refining badge design to optimize open source contributions if looking at to be used as an incentive to drive participation. While both after June 2022 and having a badge individually shows significant effects on the events of forks, the interaction term between these variables is not significant. In column 3, I show that the in relation to forks the effect of having an achievement is substantial and statistically significant along with being in the post launch period of the badges program. With a coefficient of -0.051 implies that there is 5.1% decrease in the hazard of forking among achievement holders. Showing that users with achievements reallocate their efforts away from derivative contributions like forking toward more original activities. The broader reduction in forking activity, as shown by the significant post-launch term -0.098 suggests that the program encouraged a systemic shift in user priorities across the

platform and how its incentives originality but raise potential trade-offs in collaborative engagement.

Switching to Table 3, column 1 explore the effect of achievements on the number of commits. The main effect of achievements has a negative coefficient of -0.082 and is statistically significant at the 5% level. Column 2 implies that the launch of the program for users who had achievements on total number of issues is increased 7.1% increase in the hazard of creating issues. Suggesting that the program encouraged more active issue reporting or discussion among users with achievements. The main effects of Achievements and Post Launch are both positive but insignificant, showing limited baseline effects outside the interaction term. Lastly, in column 3 the coefficient for Achievements is 0.073, representing a 7.3% increase in the hazard of submitting pull requests for users with achievements. The interaction term Achievements  $\times$  Post Launch is -0.043, reflecting a 4.3% reduction in pull request submissions for achievement holders in the post-launch period. The Post Launch term is 0.106, showing a 10.6% increase in pull request submissions across all users after the program's launch. The increase in pull requests for all users' post-launch demonstrates broader engagement incentivized by the program. The observed decline among achievement holders reflects opportunity cost considerations, where resources are redirected toward activities with higher perceived marginal returns. The results highlight how the program altered the allocation of effort particularly for those recognized with achievements.

## 5.1 The Impact of Specific Badges on Open Source Output

Having shown evidence that having a badge and being in the post launch period effects the creation of new original repositories, issues, and pull requests. I now turn to the question of which specific badges related to these activities have the greatest effect, if any. For badges

associated with new original repositories two stand out: Open Sourcerer, awarded for contributing to open-source projects through creating or maintaining repositories and Starstruck, which reflects recognition through stars on repositories, signaling well-received or innovative contributions. In regard to issues, the badges that are most closely aligned with issue events are Heart on your sleeve which is rewarded for creating discussion or participating actively in community interactions and Galaxy Brain which is given when a user closes a issue or pull request within 5 minutes of opening. Pull shark and Pair Extraordinaire will be tested on pull requests to see if those badges have an impact on how many pull requests they create.

Using Eq(2) I show in Table 4, 5, and 6 obtaining a specific badge does not significantly impact any of the three dependent variables analyzed- original repositories, issues, or pull requests. Taken together, these findings tell a story that while recognition from specific badges may hold symbolic or reputational value for users, it does not translate into measurable changes in their innovative contribution to open source in terms of new repositories, issues raised, or pull requests submitted.

## 6 Conclusion

Motivating contributors in open-source platforms presents unique challenges compared to traditional innovation systems. Unlike proprietary models, open-source communities thrive on shared effort and collaboration without direct monetary incentives or ownership rights. This dynamic raises questions about how to effectively encourage participation and creativity. As open-source innovation increasingly shapes technological progress and supports infrastructure, identifying and implementing mechanisms that foster long-term engagement is critical.

In this paper, I took an economic approach to the question of how gamification via badges can be most effectively used for incentivizing participation and effort on online systems based on user contributions. I examine to what extent the GitHub Badges program aimed at recognizing and motivating developers by celebrating milestones in their coding journey, influenced their open-source contributions. To do so, I analyze the impact of the Badges program on users who earned a badge estimating difference-in-differences models with inverse probability weights. The results suggest that the introduction of the Badges program led to an increase in new original repositories and an increase in issues and pull requests delegated. Second, I examine the potential impact of obtaining specific badges related to activities evolved around original repositories, issues, and pull requests by using a Poisson quasi fixed-effect regression. The results indicate no significant impact to contribution.

I interpret these results as evidence that the launch of the program led to an increase in contributions for users who had earned a badge. However, the actual individual milestone celebration of a badge has no impact on overall user contribution when looking at badges related to the three activities listed above. This suggests that while the GitHub Badges program effectively incentivizes participation through its structure as a recognition-based system, the design of individual badges—centered on absolute standards—may lack the broader motivational pull needed to drive sustained or widespread innovation across the platform.

My methodology and approach have certain limitations. The data utilized in this analysis include only publicly available information. Including the collection of users with badges through my own scraper that collected both the badge and date earned for each individual user. Since badges are optional to display on user profiles, some users may have earned a badge

without making it publicly visible. To ensure reliability, I conducted a series of quality and consistency checks throughout the data collection and analysis process.

There are a number of interesting directions for further exploration. A very interesting family of questions concerns user behavior, understanding how users really value badges. How can badge design be optimized to incentivize open source contribution more effectively. Some users may prioritize absolute accomplishments, while others may place greater value on relative achievements. Investigating how these preferences vary across users would provide signal towards a direction to focus mechanism design on. An interesting set of questions arise when considering multiple incentives, particularly drawing parallels to systems used in video games. Many games, such as popular title *Call of Duty*, incorporate multifaceted incentive systems through mechanisms like battle passes. Players pay a fee to gain access to the opportunity to unlock various rewards, including cosmetic items, experience boosts, and exclusive content. This system combines monetary incentives for the company with user retention strategies, as players are motivated to engage consistently to maximize their rewards. This model demonstrates how combining financial investment with tiered reward structures can sustain user participation while driving platform goals. Applying such insights to open-source ecosystems could inspire innovative designs for badge systems. For example, combining non-monetary badges with additional incentives, such as reputation-building or collaborative achievements, may better align user motivations with sustained contributions. Developing such a nuanced understanding of how users value the social-psychological rewards created by badges is essential to designing effective reward mechanisms for incentivizing user contribution (Easley, 2016) . Finally, while this study focuses on the quantity of contributions, an important question for future research is how badges influence the quality of contributions. Do badges encourage users to prioritize activity volume

over meaningful engagement? Exploring this trade-off and developing methods to incorporate contribution quality into the design of badge systems represents an exciting and impactful direction for future work.

In conclusion, this thesis highlights the potential of badges as a non-monetary incentive mechanism for fostering innovation in open-source software. By examining the GitHub Badges Program, I demonstrated how badges can encourage contributions while also revealing their limitations in significantly influencing specific activities like repository creation, issues, and pull requests. These findings emphasize the need for careful badge design to align with the motivations of contributors and the goals of open-source platforms.

## References

- Azoulay, P., & Li, D. (2020). Scientific grant funding. Vol. 26889, National Bureau of Economic Research, Cambridge, MA.
- Bitzer, J., et al. (2004). *Intrinsic Motivation in Open Source Software Development*. [econstor.eu](http://econstor.eu).
- Bitzer, J., et al. (2006). *Intrinsic Motivation in Open Source Software Development*. *Journal of Comparative Economics*, Academic Press, 13 Nov. 2006. [sciencedirect.com](http://sciencedirect.com).
- Boudreau, K. J., & Jeppesen, L. B. (2015). Unpaid crowd complementors: The platform network effect mirage. *Strategic Management Journal*, 36(12), 1761–1777.
- Boudreau, K. J., Lacetera, N., & Lakhani, K. R. (2011). Incentives and problem uncertainty in innovation contests: An empirical analysis. *Management Science*, 57, 843-863.
- Brooks, Bailey. “Introducing Achievements: Recognizing the Many Stages of a Developer’s Coding Journey.” *The GitHub Blog*, 9 June 2022, [github.blog](http://github.blog).
- Burke, M., Marlow, C., & Lento, T. M. (2009). Feed me: Motivating newcomer contribution in social network sites. In *ACM Conference on Human Factors in Computing Systems*, 945–954.
- Burke, M., & Settles, B. (2011). Plugged in to the community: Social motivators in online goal-setting groups. In *International Conference on Communities and Technologies*, 1–10.

- Chawla, S., Hartline, J. D., & Sivan, B. (2012). Optimal crowdsourcing contests. In *ACM-SIAM Symposium on Discrete Algorithms*, 856–868.
- Conti, A., Peukert, C., & Roche, M. P. (2021). Beefing it up for your investor? Open sourcing and startup funding: Evidence from GitHub. *Open Sourcing and Startup Funding: Evidence from GitHub*. August 25, 2021.
- Conti, A., Gupta, V., Guzman, J., & Roche, M. (2023). Incentivizing innovation in open source: Evidence from the GitHub Sponsors Program. *Harvard Business School*, [hbs.edu](https://hbs.edu).
- Daim, T. U., Abdelmalak, M., & Basoglu, N. (2022). Adoption and diffusion of health care technologies: The case of Istanbul Metropolitan Municipality hospitals. *Mind & Society*, 21(2), 249-270. [link.springer.com](https://link.springer.com).
- Deterding, S., Sicart, M., Nacke, L., O’Hara, K., & Dixon, D. (2011). Gamification: Using game-design elements in non-gaming contexts. In *CHI Workshop on Gamification*.
- Easley, D., & Ghosh, A. (2016). Incentives, gamification, and game theory: An economic approach to badge design. *ACM Transactions on Economics and Computation*, 4(3), Article 16. [doi.org](https://doi.org/).
- Gibbons, Robert. “Transaction-Cost Economics: Past, Present, and Future?” *Scandinavian Journal of Economics*, Wiley Blackwell, 1 Jan. 2010. [ideas.repec.org](https://ideas.repec.org).
- GitHub (2023). *Octoverse: The state of open source and rise of AI in 2023*.
- Grigorik, I. (2022). *GH Archive database*. [gharchive.org](https://gharchive.org).
- Hausberg, J. P., & Spaeth, S. (2020). Why makers make what they make: Motivations to contribute to open source hardware development. *R&D Management*. [onlinelibrary.wiley.com](https://onlinelibrary.wiley.com).
- “Incentivizing Innovation in Open Source: Evidence from the GitHub Sponsors Program.” *Harvard Business School*. [hbs.edu](https://hbs.edu). Accessed 15 July 2024.
- J. Liu, S. Huang, D. Li, S. Wen, & H. Liu (2022). Addictive incentive mechanism in crowdsensing from the perspective of behavioral economics. *IEEE Transactions on Parallel and Distributed Systems*, 33(5), 1109-1127. doi: 10.1109/TPDS.2021.3104247.
- Krishnamurthy, S., et al. “Acceptance of monetary rewards in open source software development.” *Research Policy*, North-Holland, 27 Dec. 2013. [sciencedirect.com](https://www.sciencedirect.com).
- Kulakauskaite, Indre. “How AI Is Accelerating Innovation.” *HYPE Innovation Management Software and Consulting*, HYPE Innovation, 19 Jan. 2024. [hypeinnovation.com](https://hypeinnovation.com).
- M. Burke, C. Marlow, & T. M. Lento. (2009). Feed me: Motivating newcomer contribution in social network sites. In *ACM Conference on Human Factors in Computing Systems*, 945–954.
- “2023 State of Open Source Report.” *OpenLogic by Perforce*. [openlogic.com](https://openlogic.com). Accessed 15 July 2024.
- Prendergast, Canice. “The allocation of food to food banks.” *Journal of Political Economy*, University of Chicago Press, 1 Jan. 2022. [ideas.repec.org](https://ideas.repec.org).

Figure 1: New repositories over time distinguishing by Badge status

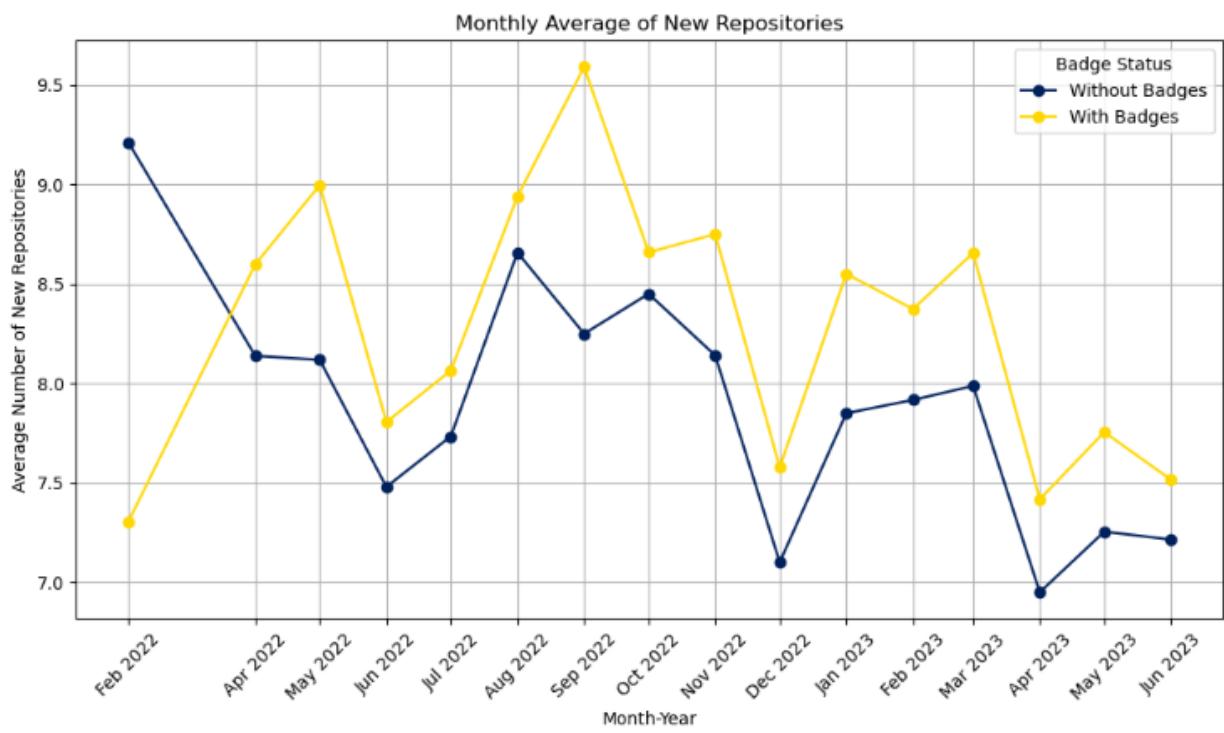


Table 1: Descriptive Statistics: Badges vs. Non-badged Users

Variable	(1)		(2)		(1)-(2)	
	Non-badged		Badged		Mean Difference	
	N	Mean/(SE)	N	Mean/(SE)	N	Value
New Repos (all)	125,975	8.2 (44.3)	8,166	8.8 (19.7)	134,141	-0.60**
New Forked Repos	125,975	0.8 (5.0)	8,166	0.8 (11.8)	134,141	0.00
New Original Repos	125,975	7.4 (43.9)	8,166	8.0 (15.6)	134,141	-0.60***
New Commits	125,975	45.1 (171.1)	8,166	43.7 (92.3)	134,141	1.40
New Issues	125,975	3.8 (32.7)	8,166	3.9 (13.8)	134,141	-0.10

Notes: *The total number of repositories ("repos"), commits, and issues are computed over the pre-sample period, which starts in February 2022 and ends in May 2022.*

Significance: \*\*\*=0.01, \*\*=0.05, \*=0.1.

Table 2: The Impact of Launching the Program on New Repos, Original Repos, and Forks

	<i>Dependent Variable:</i>		
	New Repos (1)	Original Repos (2)	Forks (3)
After June 2022	-0.045* (0.026)	-0.035 (0.029)	-0.098*** (0.018)
Badges	-0.026 (0.287)	0.0003 (0.03)	-0.296*** (0.109)
After June 2022 x Badges	0.073** (0.033)	0.082** (0.036)	-0.051 (0.064)
Observations	724,297	724,297	724,297

Notes: This table reports the results from estimating variants of Eq. (1) in the main text. In columns 1, the examined outcome is the number of new repositories created in month t. In columns 2, the examined outcome is the number of new original repositories created in t. In column 3, the examined outcome is the number of new forks created in month t. Badges is an indicator identifying users that have badges, while After June 2022 is a (0/1) indicator. Standard errors are clustered at the user level. Significance: \*\*\* = 0.01, \*\* = 0.05, \* = 0.1.

Table 3: The Impact of Launching the Program on Commits, Issues, and Pull Requests

	<i>Dependent Variable:</i>		
	Commits	Issues	Pull Requests
	(1)	(2)	(3)
After June 2022	0.086*** (0.008)	0.042*** (0.015)	0.106*** (0.006)
Badges	-0.082** (0.038)	-0.023 (0.028)	0.073** (0.029)
After June 2022 x Badges	0.009 (0.030)	0.071** (0.037)	-0.043** (0.016)
Observations	724,297	724,297	724,297

Notes: This table reports the results from estimating variants of Eq. (1) in the main text. In columns 1, the examined outcome is the number of new commits pushed in month t. In columns 2, the examined outcome is the number of issues events created in t. In column 3, the examined outcome is the number of pull requests event in month t. Badges is an indicator identifying users that have badges, while After June 2022 is a (0/1) indicator. Standard errors are clustered at the user level. Significance: \*\*\* = 0.01, \*\* = 0.05, \* = 0.1.

Table 4: The Impact of Specific Badges on Open Source Output

<i>Dependent Variable:</i>	
	Original Repos
	(1)
StarStruck x After June 2022	0.013 (0.030)
Open Sourcerer x After June 2022	0.001 (0.0009)
Observations	62,790

Table 5: The Impact of Specific Badges on Open Source Output

<i>Dependent Variable:</i>	
	Issues
	(1)
Quickdraw x After June 2022	0.057 (0.055)
Galaxy Brain x After June 2022	-0.024 (0.57)
Observations	52,363

Table 6: The Impact of Specific Badges on Open Source Output

<i>Dependent Variable:</i>	
	Pull Requests
	(1)
Pair Extraordinaire x After June 2022	0.009 (0.041)
Pull Shark x After June 2022	0.068 (0.072)
Observations	60,912